INSTITUTO TECNOLÓGICO DE ORIZABA

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

MAESTRÍA EN SISTEMAS COMPUTACIONALES

PROYECTO DE TESIS

TÍTULO: DESARROLLO DE UN MÓDULO DE MINERÍA DE CONJUNTOS DE CONTRASTE PARA LA COMPARACIÓN DE RESULTADOS DE ENCUESTAS SOBRE LA DISMINUCIÓN DE AUTOPSIAS EN HOSPITALES MEXICANOS.

PRESENTADO POR:

I.S.C. MISSAEL HERNÁNDEZ ROSADO

M13011153

PARA OBTENER EL GRADO DE:

MAESTRO EN SISTEMAS COMPUTACIONALES

DIRECTOR DEL PROYECTO DE TESIS:

DRA. LISBETH RODRÍGUEZ MAZAHUA

CODIRECTOR DEL PROYECTO DE TESIS:

M.S.C. ARACELI OLMOS VALLEJO

ORIZABA, VERACRUZ, MÉXICO. MAYO 2026

Educación
Secretaría de Educación Pública

TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Orizaba
Departamento de División de Estudios de Posgrado e Investigación

Orizaba, Veracruz, **16/abril/2026**
Dependencia: **División de Estudios de Posgrado e Investigación**
Asunto: **Autorización de Impresión**
OPCION: **I**

C. MISSAEL HERNÁNDEZ ROSADO
CANDIDATO A GRADO DE MAESTRO EN:
SISTEMAS COMPUTACIONALES
P R E S E N T E.-

De acuerdo con el Reglamento de Titulación vigente de los Centros de Enseñanza Técnica Superior, dependiente de la Dirección General de Institutos Tecnológicos de la Secretaría de Educación Pública y habiendo cumplido con todas las indicaciones que la Comisión Revisora le hizo respecto a su Trabajo Profesional titulado:

" Desarrollo de un módulo de minería de conjuntos de contraste para la comparación de resultados de encuestas sobre la disminución de autopsias en hospitales mexicanos"

comunico a Usted que este Departamento concede su autorización para que proceda a la impresión del mismo.

A T E N T A M E N T E
Excelencia en Educación Tecnológica®
CIENCIA – TÉCNICA - CULTURA®

DR. OSCAR BAÉZ SENTÍES
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

Educación
INSTITUTO TECNOLÓGICO DE ORIZABA
DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

OG-13-F06

2026 año de Margarita Maza

100% PLÁSTICO

Educación
Secretaría de Educación Pública

TECNOLÓGICO NACIONAL DE MÉXICO

Instituto Tecnológico de Orizaba
Departamento de División de Estudios de Posgrado e Investigación

C. OSCAR BAEZ SENTIES
JEFE DE LA DIVISIÓN DE ESTUDIOS
DE POSGRADO E INVESTIGACIÓN
P R E S E N T E.-

Los que suscriben, miembros del jurado, han realizado la revisión de la Tesis del (la) C.

MISSAEL HERNÁNDEZ ROSADO

La cual lleva el título de:

Desarrollo de un módulo de minería de conjuntos de contraste para la comparación de resultados de encuestas sobre la disminución de autopsias en hospitales mexicanos

Y concluyen que se acepta.

A T E N T A M E N T E
Excelencia en Educación Tecnológica®
CIENCIA – TÉCNICA - CULTURA®

PRESIDENTE:	DRA. LISBETH RODRÍGUEZ MAZAHUA	FIRMA
SECRETARIO:	DR. ULISES JUÁREZ MARTÍNEZ	FIRMA
VOCAL:	M.R.T. IGNACIO LÓPEZ MARTÍNEZ	FIRMA
VOCAL SUP:	M.S.C. ARACELI OLMOS VALLEJO	FIRMA

TA-09-18

2026 año de Margarita Maza

CARTA DE ORIGINALIDAD

En la ciudad de Orizaba, Veracruz el día 19 del mes de Mayo del año 2026 el (la) que suscribe I.S.C. Missael Hernández Rosado, alumno (a) del programa de Maestría en Sistemas Computacionales con número de control M13011153, manifiesta que es autor (a) del trabajo de tesis titulado "Desarrollo de un módulo de minería de conjuntos de contraste para la comparación de resultados de encuestas sobre la disminución de autopsias en hospitales mexicanos" y declara que el trabajo es original, ya que su contenido es producto de su directa contribución intelectual. Todos los datos y las referencias a materiales ya publicados están debidamente identificados con su respectivo crédito e incluidos en las notas bibliográficas y en las citas que se destacan como tal y, en los casos que así lo requieran, se tienen las debidas autorizaciones de quienes poseen los derechos patrimoniales. Por lo tanto, se hace responsable de cualquier litigio o reclamación relacionada con derechos de propiedad intelectual, exonerando de toda responsabilidad al Tecnológico Nacional de México / Instituto Tecnológico de Orizaba.



Missael Hernández Rosado

Nombre y firma autógrafa

CARTA DE CESIÓN DE DERECHOS

En la ciudad de Orizaba, Veracruz el día 19 del mes de Mayo del año 2026 el (la) que suscribe I.S.C. Missael Hernández Rosado, alumno (a) del programa de Maestría en Sistemas Computacionales con número de control M13011153, manifiesta que es autor (a) intelectual del trabajo de tesis bajo la dirección de la Dra. Lisbeth Rodriguez Mazahua y ceden los derechos del trabajo intitulado "Desarrollo de un módulo de minería de conjuntos de contraste para la comparación de resultados de encuestas sobre la disminución de autopsias en hospitales mexicanos" al TecNM/Instituto Tecnológico de Orizaba para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y del director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección: msc@orizaba.tecnm.mx . Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.



Missael Hernández Rosado

Nombre y firma

Agradecimientos

Agradezco profundamente a mi familia, a quienes dedico cada logro profesional alcanzado.

A mi esposa, Tania, quien me acompañó durante este periodo de entrega, fungiendo como la primera auditora de cada una de mis presentaciones de seminario desde casa. Gracias por apoyarme incluso cuando nuestra hija aún estaba dentro de ti, acompañándonos en cada despertar de la mañana y en más de un desvelo, eres mi pareja, mi equipo, te amo.

A mi hija Lía, porque durante este camino crecí profesionalmente, pero aún más al convertirme en padre mientras cursaba esta maestría. Llegaste para representar un nuevo reto que me impulsó a crecer como persona, haciéndome más racional en mis decisiones, más consciente y, sobre todo, más sensible y amoroso a lo largo de este proceso.

A mis padres, Ramón y Gloria, porque, aunque ya no eran quienes me despertaban para asistir a clases a las 7 de la mañana bajo su techo, siempre han estado presentes en mi corazón, en mi forma de hablar, de expresarme, en mi personalidad, así como en los principios y valores que inculcaron con esfuerzo, dedicación y amor a lo largo de su vida. Este logro también lleva su nombre, porque soy reflejo de lo que han sido para mí: unos padres auténticos y ejemplares.

A mi hermana Abigail, quien siempre estuvo pendiente de mis avances y progreso; sin hacer mucho ruido, pero demostrando un gran amor en cada mensaje.

Finalmente, agradezco a cada uno de los maestros y doctores que dirigieron este proyecto e impartieron clases durante mi formación. Valoro y admiro profundamente su esfuerzo, dedicación y compromiso con la enseñanza.

Índice

RESUMEN **11**

ABSTRACT **12**

INTRODUCCIÓN **13**

CAPÍTULO 1. ANTECEDENTES **14**

1.1 MARCO TEÓRICO 14

1.1.1 Autopsia *14*

1.1.2 Autopsia verbal *14*

1.1.3 Autopsia médica *14*

1.1.4 Autopsia forense *15*

1.1.5 Causa de muerte *15*

1.1.6 Minería de datos *15*

1.1.7 Aprendizaje supervisado *15*

1.1.8 Clasificación en el aprendizaje supervisado *16*

1.1.9 Regresión en el aprendizaje supervisado *16*

1.1.10 Aprendizaje no supervisado *16*

1.1.11 Agrupamiento en el aprendizaje no supervisado *17*

1.1.12 Reglas de asociación en el aprendizaje no supervisado *17*

1.1.13 Aprendizaje automático *17*

1.1.14 Descubrimiento supervisado de reglas descriptivas *17*

1.1.15 Minería de conjuntos de contraste *18*

1.1.16 Descubrimiento de subgrupos *18*

1.1.17 Minería de patrones emergentes *18*

1.2 SITUACIÓN TECNOLÓGICA, ECONÓMICA Y OPERATIVA DE LA EMPRESA 18

1.3 PLANTEAMIENTO DEL PROBLEMA 19

1.4 OBJETIVO GENERAL Y ESPECÍFICOS 20

1.4.1 Objetivo general *20*

1.4.2 Objetivos específicos *20*

1.5 JUSTIFICACIÓN 21

CAPÍTULO 2. ESTADO DE LA PRÁCTICA **23**

2.1 TRABAJOS RELACIONADOS 23

2.2 ANÁLISIS COMPARATIVO 34

2.3 PROPUESTA DE SOLUCIÓN 49

CAPÍTULO 3. APLICACIÓN DE LA METODOLOGÍA **51**

3.1 Análisis 51
3.1.1 Preguntas de Investigación *52*
3.1.2 Elegibilidad y Selección de Estudios Relevantes *52*
3.1.3 Criterios de Inclusión y Exclusión *54*
3.1.4 Análisis Comparativo *54*
3.2 Selección 57
3.2.1 STUCCO *58*
3.2.2 RuleKit2-CS *64*
3.2.3 ECSM *68*
3.2.4 Comparación de Algoritmos *70*
3.3 Diseño de la aplicación 75
3.3.1 Arquitectura del módulo *75*
3.3.2 Base de Datos *79*
3.3.2 Casos de uso *84*
3.3.3 Módulo web *86*
3.4 Desarrollo del Módulo de Minería de Conjuntos de Contraste 92
3.4.1 Kanban *93*
3.4.2 Contenerización y despliegue *98*
3.4.3 Integración con Ant y Makefile *101*
3.4.4 Construcción de Vistas y Funciones de Soporte en PostgreSQL *105*
3.4.5 Implementación de lógica del módulo en Java y ejecución de Rulekit2-CS *121*
3.4.6 Implementación de la vista con JSF y PrimeFaces *136*

CAPÍTULO 4 RESULTADOS 148

4.1 Presentación del módulo de minera de conjuntos de contraste 150
4.1.1 Pantalla Principal del Sistema *151*
4.1.2 Inicio de Sesión *152*
4.1.3 Configuración del Modelo para la Ejecución de RuleKit2-CS *153*
4.1.4 Reglas Sin Interpretar *157*
4.1.5 Interpretación de las Reglas *159*
4.1.6 Análisis Visual de las Reglas *161*
4.1.7 Generación del Reporte PDF *163*
4.2 Vista Minable: Resultados aplicando la función de calidad Lift 168
4.2.1 Reglas sin interpretar *168*
4.2.2 Reglas interpretadas *172*
4.2.3 Análisis Visual *177*
4.2.4 Evaluación Objetiva *182*

4.2.5 Evaluación Subjetiva *185*
4.3 VISTA MINABLE: RESULTADOS APLICANDO LA FUNCIÓN DE CALIDAD WRACC 186
4.3.1 Reglas sin interpretar *187*
4.3.2 Reglas interpretadas *192*
4.3.3 Análisis Visual *196*
4.3.4 Evaluación Objetiva *201*
4.3.5 Evaluación Subjetiva *203*
4.4 VISTA MINABLE: RESULTADOS APLICANDO LA FUNCIÓN DE CALIDAD *CORRELATION* 205
4.4.1 Reglas sin interpretar *205*
4.4.2 Reglas interpretadas *209*
4.4.3 Análisis Visual *213*
4.4.4 Evaluación objetiva *217*
4.4.5 Evaluación subjetiva *218*
4.5 VISTA MINABLE: RESULTADOS APLICANDO LA FUNCIÓN DE CALIDAD *KLOSGEN* 219
4.5.1 Reglas sin interpretar *219*
4.5.2 Reglas interpretadas *224*
4.5.3 Análisis Visual *227*
4.5.4 Evaluación Objetiva *233*
4.5.5 Evaluación Subjetiva *233*
4.6 VISTA MINABLE: RESULTADOS APLICANDO LA FUNCIÓN DE CALIDAD *RELATIVERISK* 235
4.6.1 REGLAS SIN INTERPRETAR 235
4.6.2 Reglas Interpretadas *237*
4.6.3 Análisis Visual *239*
4.6.4 Evaluación Objetiva *243*
4.6.5 Evaluación Subjetiva *243*
4.7 MATRIZ BINARIA: RESULTADOS APLICANDO LA FUNCIÓN DE CALIDAD *LIFT* 244
4.7.1 Reglas sin interpretar *245*
4.7.2 Reglas Interpretadas *251*
4.7.3 Análisis visual *258*
4.7.4 Evaluación objetiva *263*
4.7.5 Evaluación subjetiva *265*
4.8 MATRIZ BINARIA: RESULTADOS APLICANDO LA FUNCIÓN DE CALIDAD *WRACC* 265
4.8.1 Reglas sin interpretar *266*
4.8.2 Reglas interpretadas *273*
4.8.3 Análisis visual *282*
4.8.4 Evaluación objetiva *286*
4.8.5 Evaluación subjetiva *286*

4.9 Matriz Binaria: Resultados aplicando la función de calidad *Correlation* 287
4.9.1 Reglas sin interpretar *287*
4.9.2 Reglas interpretadas *294*
4.9.3 Análisis visual *302*
4.9.4 Evaluación objetiva *305*
4.9.5 Evaluación subjetiva *306*
4.10 Matriz Binaria: Resultados aplicando la función de calidad *Klosgen* 307
4.10.1 Reglas sin interpretar *307*
4.10.2 Reglas interpretadas *313*
4.10.3 Análisis Visual *321*
4.10.4 Evaluación objetiva *325*
4.10.5 Evaluación subjetiva *325*
4.11 Matriz Binaria: Resultados aplicando la función de calidad *RelativeRisk* 326
4.11.1 Reglas sin interpretar *326*
4.11.2 Reglas interpretadas *330*
4.11.3 Análisis visual *333*
4.11.4 Evaluación objetiva *336*
4.11.5 Evaluación subjetiva *337*
4.12 Análisis de Resultados 337
4.12.1 Comparativa Técnica y Estadística *338*
4.12.2 Evaluación Subjetiva y Criterio Experto *342*
4.12.3 Selección de la Función de Calidad *344*

CAPÍTULO 5. CONCLUSIONES Y RECOMENDACIONES 347

5.1 Conclusiones 347
5.2 Recomendaciones 349

PRODUCTOS ACADÉMICOS 351

REFERENCIAS 352

Índice de Tablas

Tabla 2.1 Análisis Comparativo De Artículos Relacionados 34
Tabla 3.1 Preguntas De Investigación Y Motivaciones 52
Tabla 3.2 Artículos Seleccionados 55
Tabla 3.3 Descripción De Parámetros STUCCO 60
Tabla 3.4 Configuraciones Aplicadas Por Fase De Análisis STUCCO 60
Tabla 3.5 Resultados Obtenidos Con STUCCO Por Fase De Análisis 63
Tabla 3.6 Descripción De Parámetros Rulekit2-CS 66
Tabla 3.7 Configuraciones Aplicadas Por Fase De Análisis Rulekit2-CS 66
Tabla 3.8 Resultados Obtenidos Con Rulekit2-CS Por Fase De Análisis 68
Tabla 3.9 Homologación De Parámetros STUCCO Vs Rulekit2-CS 71
Tabla 3.10 Resultados Comparativos – STUCCO Vs Rulekit2-CS 73
Tabla 3.11 Tablero Kanban 95
Tabla 4.1 Lift – Vista Minable: Reglas Sin Interpretar 170
Tabla 4.2 Lift – Vista Minable: Reglas Interpretadas 173
Tabla 4.3 WRAcc – Vista Minable: Reglas Sin Interpretar 189
Tabla 4.4 WRAcc – Vista Minable: Reglas Interpretadas 193
Tabla 4.5 Correlation – Vista Minable: Reglas Sin Interpretar 207
Tabla 4.6 Correlation – Vista Minable: Reglas Interpretadas 210
Tabla 4.7 Klosgen – Vista Minable: Reglas Sin Interpretar 221
Tabla 4.8 Klosgen – Vista Minable: Reglas Interpretadas 225
Tabla 4.9 RelativeRisk – Vista Minable: Reglas Sin Interpretar 237
Tabla 4.10 RelativeRisk – Vista Minable: Reglas Interpretadas 238
Tabla 4.11. Lift – Matriz Binaria: Reglas Sin Interpretar 247
Tabla 4.12. Lift – Matriz Binaria: Reglas Interpretadas 253
Tabla 4.13. WRAcc – Matriz Binaria: Reglas Sin Interpretar 268
Tabla 4.14. WRAcc – Matriz Binaria: Reglas Interpretadas 275
Tabla 4.15. Correlation – Matriz Binaria: Reglas Sin Interpretar 289
Tabla 4.16. Correlation – Matriz Binaria: Reglas Interpretadas 295
Tabla 4.17. Klosgen – Matriz Binaria: Reglas Sin Interpretar 309
Tabla 4.18. Klosgen – Matriz Binaria: Reglas Interpretadas 314
Tabla 4.19. RelativeRisk – Matriz Binaria: Reglas Sin Interpretar 328
Tabla 4.20. Klosgen – Matriz Binaria: Reglas Interpretadas 330
Tabla 4.21. Vista Minable: Tablas Comparativa de Funciones de Calidad 341
Tabla 4.22. Matriz Comparativa de Desempeño 346

Índice de Figuras

Figura 3.1 PRISMA Identificación De Estudios 53
Figura 3.2 Arquitectura Del Módulo De Minería De Conjuntos De Contraste 77
Figura 3.3 Diagrama Conceptual De La Base De Datos 80
Figura 3.4 Diagrama Lógico De La Base De Datos 81
Figura 3.5 Diagrama Físico De La Base De Datos 82
Figura 3.6 Diagrama De Caso De Uso Del Sistema 85
Figura 3.7 Diagrama De Caso De Uso Del Módulo De CSM 85
Figura 3.8 Diagrama De Flujo De Aplicación Web 87
Figura 3.9 Pantalla Principal Módulo CSM 88
Figura 3.10 Pantalla Reglas Sin Interpretar Módulo CSM 89
Figura 3.11 Pantalla Reglas Interpretadas Módulo CSM 90
Figura 3.12 Pantalla Análisis Visual Módulo CSM 90
Figura 3.13 Reporte PDF Pagina 1 Y 2 Módulo CSM 91
Figura 3.14 Reporte PDF Pagina 3 Módulo CSM 91
Figura 3.15 Tablero Kanban 94
Figura 4.1 Pantalla Principal Del Sistema 152
Figura 4.2 Ventana De Inicio De Sesión 152
Figura 4.3 Interfaz General Del Módulo 153
Figura 4.4 Selección Del Nivel De Análisis 154
Figura 4.5 Selección De La Métrica De Calidad 155
Figura 4.6 Creación Del Modelo 156
Figura 4.7 Ejecución Rulekit2-CS 157
Figura 4.8 Vista De Reglas Sin Interpretar 158
Figura 4.9 Vista De Interpretación En Lenguaje Natural De Las Reglas Generadas 160
Figura 4.10 Vista Del Análisis Visual De Métricas Asociadas A Las Reglas Generadas 162
Figura 4.11 Portada Del Reporte PDF 164
Figura 4.12 Descripción De La Métrica Y Explicación De Los Parámetros 165
Figura 4.13 Sección Del Reporte Correspondiente A Las Reglas Sin Interpretar 166
Figura 4.14 Reglas Interpretadas Incluidas Dentro Del Reporte PDF 167
Figura 4.15 Primera De Las Páginas Del Análisis Visual Incluidas En El Reporte PDF 168
Figura 4.16 Lift – Vista Minable: Soporte Por Regla 177
Figura 4.17 Lift – Vista Minable: Soporte Negativo Por Regla 178
Figura 4.18 Lift – Vista Minable: Diferencia De Soporte Por Regla 179
Figura 4.19 Lift – Vista Minable: Precisión Por Regla 180
Figura 4.20 Lift – Vista Minable: Diferencia De Precisión Por Regla 180
Figura 4.21 Lift – Vista Minable: Redundancia Por Regla 181
Figura 4.22 WRAcc – Vista Minable: Soporte Por Regla 197

Figura 4.23 WRAcc – Vista Minable: Soporte Negativo Por Regla 198
Figura 4.24 WRAcc – Vista Minable: Diferencia De Soporte Por Regla 198
Figura 4.25 WRAcc – Vista Minable: Precisión Por Regla 199
Figura 4.26 WRAcc – Vista Minable: Diferencia De Precisión Por Regla 200
Figura 4.27 WRAcc – Vista Minable: Redundancia Por Regla 200
Figura 4.28 Correlation – Vista Minable: Soporte Por Regla 214
Figura 4.29 Correlation – Vista Minable: Soporte Negativo Por Regla..... 214
Figura 4.30 Correlation – Vista Minable: Diferencia De Soporte Por Regla 215
Figura 4.31 Correlation – Vista Minable: Precisión Por Regla 216
Figura 4.32 Correlation – Vista Minable: Diferencia De Precisión Por Regla 216
Figura 4.33 Correlation – Vista Minable: Redundancia Por Regla 217
Figura 4.34 Klosgen – Vista Minable: Soporte Por Regla 228
Figura 4.35 Klosgen – Vista Minable: Soporte Negativo Por Regla 229
Figura 4.36 Klosgen – Vista Minable: Diferencia De Soporte Por Regla 230
Figura 4.37 Klosgen – Vista Minable: Precisión Por Regla 230
Figura 4.38 Klosgen – Vista Minable: Diferencia De Precisión Por Regla 231
Figura 4.39 Klosgen – Vista Minable: Redundancia Por Regla 232
Figura 4.40 RelativeRisk – Vista Minable: Soporte Por Regla 240
Figura 4.41 RelativeRisk – Vista Minable: Soporte Negativo Por Regla 240
Figura 4.42 RelativeRisk – Vista Minable: Diferencia De Soporte Por Regla 241
Figura 4.43 RelativeRisk – Vista Minable: Precisión Por Regla 241
Figura 4.44 RelativeRisk – Vista Minable: Diferencia De Precisión Por Regla 242
Figura 4.45 RelativeRisk – Vista Minable: Redundancia Por Regla 242
Figura 4.46 Lift – Matriz Binaria: Soporte Por Regla 259
Figura 4.47 Lift – Matriz Binaria: Soporte Negativo Por Regla 260
Figura 4.48 Lift – Matriz Binaria: Diferencia De Soporte Por Regla 261
Figura 4.49 Lift – Matriz Binaria: Precisión Por Regla 261
Figura 4.50 Lift – Matriz Binaria: Diferencia De Precisión Por Regla 262
Figura 4.51 Lift – Matriz Binaria: Redundancia Por Regla 263
Figura 4.52 WRAcc – Matriz Binaria: Soporte Por Regla 283
Figura 4.53 WRAcc – Matriz Binaria: Soporte Negativo Por Regla 283
Figura 4.54 WRAcc – Matriz Binaria: Diferencia De Soporte Por Regla 284
Figura 4.55 WRAcc – Matriz Binaria: Precisión Por Regla 284
Figura 4.56 WRAcc – Matriz Binaria: Diferencia De Precisión Por Regla 285
Figura 4.57 WRAcc – Matriz Binaria: Redundancia Por Regla 285
Figura 4.58 Correlation – Matriz Binaria: Soporte Por Regla 303
Figura 4.59 Correlation – Matriz Binaria: Soporte Negativo Por Regla 303
Figura 4.60 Correlation – Matriz Binaria: Diferencia De Soporte Por Regla 304
Figura 4.61 Correlation – Matriz Binaria: Precisión Por Regla 304

Figura 4.62 Correlation – Matriz Binaria: Diferencia De Precisión Por Regla 305
Figura 4.63 Correlation – Matriz Binaria: Redundancia Por Regla 305
Figura 4.64 Klosgen – Matriz Binaria: Soporte Por Regla 322
Figura 4.65 Klosgen – Matriz Binaria: Soporte Negativo Por Regla 322
Figura 4.66 Klosgen – Matriz Binaria: Diferencia De Soporte Por Regla 323
Figura 4.67 Klosgen – Matriz Binaria: Precisión Por Regla 323
Figura 4.68 Klosgen – Matriz Binaria: Diferencia De Precisión Por Regla 324
Figura 4.69 Klosgen – Matriz Binaria: Redundancia Por Regla 325
Figura 4.70 RelativeRisk – Matriz Binaria: Soporte Por Regla 334
Figura 4.71 RelativeRisk – Matriz Binaria: Soporte Negativo Por Regla 334
Figura 4.72 RelativeRisk – Matriz Binaria: Diferencia de Soporte por Regla 335
Figura 4.73 RelativeRisk – Matriz Binaria: Precisión por Regla 335
Figura 4.74 RelativeRisk – Matriz Binaria: Diferencia de Precisión por Regla 336
Figura 4.75 RelativeRisk – Matriz Binaria: Redundancia por Regla 336

Índice de Listados

Listado 3.1. Definición de servicios y orquestación de contenedores en Docker Compose 99
Listado 3.2. Dockerfile para la construcción de la imagen del servidor de aplicaciones 100
Listado 3.3. Comando de compilación y empaquetado mediante Apache Ant 101
Listado 3.4. Makefile: Declaración de variables globales y configuración del entorno 102
Listado 3.5. Makefile: Objetivo de compilación y generación del WAR 103
Listado 3.6. Makefile: Objetivo de construcción de la imagen Docker 103
Listado 3.7. Makefile: Objetivo de despliegue automatizado de contenedores 104
Listado 3.8. Makefile: Objetivos para la gestión y monitoreo de registros (logs) 104
Listado 3.9. Makefile: Automatización de respaldo y restauración de base de datos 104
Listado 3.10. Makefile: Rutina de limpieza de imágenes Docker huérfanas 105
Listado 3.11. Rutina de limpieza e inicialización de la tabla vista_minable_csm 106
Listado 3.12. Definición estructural (DDL) de la vista minable para conjuntos de contraste ... 107
Listado 3.13. Lógica de filtrado de instancias por centro hospitalario (Vista Minable) 107
Listado 3.14. Consultas de extracción y desnormalización de atributos médicos 108
Listado 3.15. Inserción parametrizada de registros en la vista minable 108
Listado 3.16. Limpieza de columnas técnicas en la vista minable 109
Listado 3.17. Rutina de reinicio para la matriz binaria de conjuntos de contraste 109
Listado 3.18. Patrón de nomenclatura para columnas dinámicas 110
Listado 3.19. Construcción dinámica de columnas dimensionales (One-Hot Encoding) 110
Listado 3.20. Ejemplos de identificadores de variables binarias generadas 110
Listado 3.21. Filtrado de instancias en la generación de la matriz binaria 111
Listado 3.22. Inicialización de registros base en la matriz binaria 111
Listado 3.23. Asignación dinámica de valores de presencia en atributos binarios 111
Listado 3.24. Limpieza final de la estructura de la matriz binaria 112
Listado 3.25. Declaración de la función NLG para traducción de reglas (Vista Minable) 113
Listado 3.26. Análisis sintáctico (Parsing) de la regla generada por RuleKit2-CS 113
Listado 3.27. Bucle de interpretación semántica de antecedentes 114
Listado 3.28. Construcción del prefijo y concatenación de condiciones 115
Listado 3.29. Generación de la conclusión y ensamblaje final de la oración 116
Listado 3.30. Función auxiliar para decodificación de variables binarias 117
Listado 3.31. Declaración de la función NLG para reglas basadas en patrones binarios 117
Listado 3.32. Separación de estructura antecedente-consecuente en reglas de matriz 118
Listado 3.33. Lógica de traducción de condiciones binarias (Afirmación/Negación) 118
Listado 3.34. Construcción del prefijo para oraciones basadas en matriz 119
Listado 3.35. Generación de la conclusión para reglas de matriz binaria 119
Listado 3.36. Declaración del controlador y atributos principales 122
Listado 3.37. Atributos de conexión y metadatos de experimento 123

Listado 3.38. Inicialización de parámetros del módulo y configuración de conexión 124
Listado 3.39. Orquestación del flujo de minería en el módulo 125
Listado 3.40. Selección del dataset, vista y nivel analítico 126
Listado 3.41. Reemplazo de placeholders en el modelo XML y configuración de 127
Listado 3.42. Persistencia de la configuración y generación del archivo ARFF 128
Listado 3.43. Ejecución del motor RuleKit2-CS y captura de bitácora 129
Listado 3.44. Manejo del código de terminación y carga del CSV 130
Listado 3.45. Recuperación de instancias desde la base de datos mediante Weka 131
Listado 3.46. Localización e interpretación del archivo CSV 132
Listado 3.47. Definición del bean de sesión y alcance de la vista JSF 133
Listado 3.48. Asociación de atributos de vista con parámetros del experimento 134
Listado 3.49. Estructura base de la vista de consulta 137
Listado 3.50. Formulario de configuración y ejecución del modelo 138
Listado 3.51. Botones de creación del modelo y ejecución de RuleKit2-CS 139
Listado 3.52. Acceso desde la vista de modelado a las vistas de resultados 140
Listado 3.53. Eventos de inicialización de la vista de consulta 141
Listado 3.54. Formulario de consulta de parámetros en la vista de resultados 141
Listado 3.55. Tabla de reglas interpretadas en la vista integrada de resultados 142
Listado 3.56. Configuración de gráficos de barras para métricas de reglas 144
Listado 3.57. Tabla de reglas sin interpretar 145
Listado 3.58. Tabla de reglas interpretadas 146

Resumen

La disminución de autopsias en el Hospital Regional Río Blanco (H.R.R.B.) ha generado interrogantes sobre las causas que la provocan, y se ha identificado la necesidad de validar si este fenómeno es específico de este hospital o si también se refleja en otros centros hospitalarios.

En este trabajo se desarrolló un módulo bajo un enfoque de minería de conjuntos de contraste para comparar los resultados de encuestas sobre la disminución de autopsias en hospitales mexicanos, basado en reglas mediante la herramienta RuleKit2-CS, integrada en un entorno arquitectónico modular compuesto por Java, Spring y PostgreSQL. La arquitectura del módulo se organiza en tres capas bajo el patrón Modelo-Vista-Controlador (MVC), permitiendo una gestión clara de las responsabilidades y facilitando la escalabilidad y el mantenimiento del sistema.

Se siguió un proceso de desarrollo iterativo basado en la metodología ágil Kanban, con un enfoque en la estructura KDD (*Knowledge Discovery in Databases,* Descubrimiento de conocimiento en bases de datos) para organizar las fases de análisis de datos, desde la recolección hasta la interpretación de los resultados. La aplicación del módulo de minería de conjuntos de contraste proporciona a las áreas de patología una herramienta robusta para el análisis comparativo, ayudando a identificar las causas específicas de la disminución de autopsias y permitiendo tomar decisiones fundamentadas que reviertan esta tendencia.

Abstract

The decrease in autopsies at the Río Blanco Regional Hospital (H.R.R.B.) has raised questions about the causes that induce it, and the need to validate whether this phenomenon is specific to this hospital or if it is also reflected in other hospitals has been identified.

In this work, it was developed a module under a contrast set mining approach to compare the results of surveys on the decrease in autopsies in Mexican hospitals, based on rules using the RuleKit2-CS tool, integrated in a modular architectural environment composed of Java, Spring and PostgreSQL. The architecture of the module is organized in three layers under the Model-View-Controller (MVC) pattern, allowing clear management of responsibilities and facilitating scalability and maintenance of the system.

An iterative development process based on the agile Kanban methodology was followed, with a focus on the KDD (Knowledge Discovery in Databases) structure to organize the data analysis phases, from the collection to the interpretation of the results. The application of the contrast set mining module provide pathology areas with a robust tool for comparative analysis, helping to identify the specific causes of the decrease in autopsies and allowing informed decisions to be made that reverse this trend.

Introducción

La notable disminución en la realización de autopsias en el Hospital Regional Río Blanco ha suscitado una creciente preocupación dentro del área de patología. Las autopsias, esenciales para el avance del conocimiento médico y la mejora en la calidad de la atención de salud, han experimentado una reducción considerable en los últimos años. Ante este panorama, surge la necesidad de analizar las causas subyacentes a esta tendencia y determinar si este fenómeno es exclusivo del H.R.R.B. o si también afecta a otros hospitales en México.

El presente proyecto consistió en el desarrollo de un módulo de minería de datos, específicamente utilizando la técnica de minería de conjuntos de contraste, para comparar las opiniones de médicos del H.R.R.B. y otros hospitales en relación con la disminución de autopsias. Mediante la identificación de patrones discriminativos en los datos, este análisis permite una comprensión más clara de las causas que han llevado a la disminución de las autopsias y contribuirá a la toma de decisiones informadas que permitan revertir esta tendencia.

El documento se encuentra estructurado en cinco capítulos. En el primer capítulo se presenta el marco teórico, el contexto general del estudio, el planteamiento del problema, los objetivos y la justificación de la investigación. El segundo capítulo expone los trabajos relacionados, incorporando un análisis comparativo de las principales aportaciones vinculadas al tema. En el tercer capítulo se describe la aplicación de la metodología empleada para el desarrollo del módulo propuesto. El cuarto capítulo presenta los resultados obtenidos a partir del análisis comparativo, así como su interpretación. Finalmente, el quinto capítulo integra las conclusiones y recomendaciones derivadas del trabajo realizado.

Capítulo 1. Antecedentes

El primer capítulo de la presente investigación expone los conceptos clave que sustentan la investigación, define el problema central a abordar, establece el objetivo general y los objetivos específicos, y proporciona una justificación detallada.

1.1 Marco teórico

A continuación, se precisan los términos fundamentales que están relacionados e intervienen en el tema de investigación.

1.1.1 Autopsia

Una autopsia es comúnmente conocida como la examinación detallada de un cadáver para identificar las causas de la muerte. Este proceso es ordenado y riguroso, involucrando tanto la observación del exterior del cuerpo como la inspección de los órganos internos [1].

1.1.2 Autopsia verbal

La autopsia verbal es un método para determinar la causa de muerte mediante entrevistas con familiares o cuidadores del difunto. Estas entrevistas utilizan un cuestionario estandarizado para recopilar información sobre los síntomas, el historial médico y las circunstancias que precedieron al fallecimiento. Posteriormente, profesionales de la salud o algoritmos analizan estos datos para identificar la causa probable de la muerte [2].

1.1.3 Autopsia médica

La autopsia médica, también conocida como autopsia clínica u hospitalaria, es un procedimiento *post mortem* altamente especializado que combina un examen externo e interno del cuerpo humano. Su objetivo principal es determinar o confirmar la causa de muerte y detectar posibles enfermedades subyacentes. Este tipo de autopsia no solo es fundamental para garantizar la calidad de la práctica médica, sino también para la recopilación de estadísticas de causas de muerte, el seguimiento de enfermedades y tratamientos, así como para la salud pública [3].

1.1.4 Autopsia forense

Una autopsia forense, también conocida como autopsia médico-legal, es un examen *post mortem* realizado con fines médico-legales. Este procedimiento se lleva a cabo bajo la dirección de una autoridad legal competente, encargada de investigar muertes que son repentinas, inesperadas, sospechosas o que presentan circunstancias poco claras o litigiosas. Además, las autopsias forenses se utilizan para examinar casos de muertes violentas, laborales o industriales, y aquellas asociadas a tratamientos médicos o quirúrgicos donde se alega negligencia, así como muertes relacionadas con anestesia. El objetivo principal es esclarecer las causas y circunstancias de la muerte, aportando evidencia relevante en investigaciones judiciales [4].

1.1.5 Causa de muerte

El término “causa de muerte” se refiere a la razón médica del fallecimiento, listando las enfermedades o lesiones que llevaron al deceso. Se describe como una condición o lesión que desencadenó la serie de eventos que resultaron directamente en la muerte [5].

1.1.6 Minería de datos

La minería de datos es el proceso de identificar correlaciones, patrones y tendencias significativas mediante el análisis de grandes volúmenes de datos almacenados. Utiliza tecnologías de reconocimiento de patrones junto con técnicas estadísticas y matemáticas [6].

1.1.7 Aprendizaje supervisado

El aprendizaje supervisado se emplea en tareas de predicción, con el objetivo de prever o clasificar un resultado específico, como la presencia o ausencia de una enfermedad mental. Ha sido utilizado en grandes conjuntos de datos como aquellos que incluyen predictores demográficos, clínicos y sociales [7].

En la minería de datos, el aprendizaje supervisado se divide en dos tipos de problemas: clasificación y regresión.

1.1.8 Clasificación en el aprendizaje supervisado

La clasificación consiste en asignar categorías a datos nuevos basándose en ejemplos previos que han sido correctamente clasificados. Este tipo de problema ocurre cuando la variable de salida es categórica, como por ejemplo los colores “rojo” o “azul”, o el estado de salud “enfermo” o “sano”. La clasificación se aplica en tareas como la recomendación de productos o la predicción de series temporales discretas. Algunos algoritmos supervisados populares para resolver problemas de clasificación incluyen los bosques aleatorios (*Random Forest*), que combinan múltiples árboles de decisión, y las máquinas de soporte vectorial (*Support Vector Machine*), que encuentran un hiperplano óptimo para separar las categorías en los datos [8].

1.1.9 Regresión en el aprendizaje supervisado

La regresión es un tipo de problema en el cual la variable de salida es un valor continuo o real, como el precio en dólares o el peso de un objeto. Se enfoca en predecir valores numéricos con base en la relación entre variables de entrada. Un ejemplo común de este tipo de problemas incluye la predicción de tendencias en series temporales o recomendaciones de productos con base en datos históricos. Los algoritmos populares para abordar problemas de regresión incluyen la regresión lineal, que modela la relación entre variables como una línea recta, y los bosques aleatorios (*Random Forest*), que utilizan múltiples árboles de decisión para mejorar la precisión en la predicción [9].

1.1.10 Aprendizaje no supervisado

El aprendizaje automático no supervisado utiliza algoritmos de *machine learning* (aprendizaje automático) para analizar y agrupar conjuntos de datos que no han sido previamente etiquetados. Su objetivo es identificar dimensiones, componentes, grupos o trayectorias subyacentes dentro de una estructura de datos, siendo útil en tareas de descripción al descubrir relaciones en los datos sin un resultado predefinido [10].

1.1.11 Agrupamiento en el aprendizaje no supervisado

El agrupamiento es una técnica de minería de datos utilizada en el aprendizaje no supervisado que organiza un conjunto de datos no etiquetados en grupos o categorías basadas en similitudes y diferencias intrínsecas. Su objetivo es identificar estructuras ocultas en los datos y clasificar objetos en función de sus características compartidas, sin la necesidad de etiquetas predefinidas [11].

1.1.12 Reglas de asociación en el aprendizaje no supervisado

Las reglas de asociación son un método basado en reglas dentro del aprendizaje no supervisado que identifica relaciones y patrones significativos entre variables dentro de un conjunto de datos. Estas reglas establecen correlaciones entre diferentes atributos, permitiendo descubrir dependencias y asociaciones que pueden utilizarse para la toma de decisiones y la optimización de procesos [12].

1.1.13 Aprendizaje automático

El aprendizaje automático es una rama de la informática que se enfoca en aprender patrones a partir de datos para mejorar el rendimiento en diversas tareas, como la predicción. Este proporciona nuevas herramientas para abordar desafíos que los métodos estadísticos tradicionales no resuelven, este campo se divide principalmente en aprendizaje supervisado y no supervisado [13].

1.1.14 Descubrimiento supervisado de reglas descriptivas

Es un conjunto de técnicas que se ubica entre la predicción y la descripción, con el objetivo de explicar los datos en relación con una variable específica mediante el uso de patrones o reglas.

El objetivo de este marco de tareas es la descripción de relaciones interesantes para los expertos supeditadas a una variable de interés o clase. Estas técnicas se centran en describir características o factores que influyen en un comportamiento, más que en predecir nuevos comportamientos [14].

1.1.15 Minería de conjuntos de contraste

La minería de conjuntos de contraste busca patrones cuya distribución varía significativamente entre diferentes grupos de objetos, los cuales son mutuamente excluyentes, por lo que una instancia únicamente pertenece a uno de ellos. Su aplicación es interesante en investigaciones de ciencias sociales, pues es muy habitual este tipo de análisis, y en el presente se aplica a otros campos como analítica de negocio o medicina, por mencionar algunos [15].

1.1.16 Descubrimiento de subgrupos

El objetivo del descubrimiento de subgrupos se enfoca en describir subconjuntos de datos cuya distribución estadística difiere notablemente del conjunto total, dada una variable de interés o etiqueta de clase. Se centra en encontrar relaciones interesantes que no necesitan ser completas, es decir, no todas las instancias de la clase cumplen con esa relación, y que suelen expresarse en forma de patrones individuales, simplificando y resumiendo los datos de manera comprensible para los expertos [16].

1.1.17 Minería de patrones emergentes

La minería de patrones emergentes se encarga de identificar ciertos patrones que sean frecuentes para una clase y poco frecuentes para el resto. Los objetivos son la búsqueda de tendencias emergentes a lo largo del tiempo y la búsqueda de características discriminativas entre las diferentes clases de un problema. Dada esta propiedad se emplea tanto con carácter descriptivo como de manera predictiva.

Al igual que el descubrimiento de subgrupos, la principal característica de esta tarea es su capacidad de simplificar y resumir los datos de manera fiable y altamente comprensible para los expertos [17].

1.2 Situación tecnológica, económica y operativa de la empresa

El Tecnológico Nacional de México Campus Orizaba, ubicado en la ciudad de Orizaba, Veracruz, es una institución educativa de renombre en la región, orientada

al desarrollo de la ciencia y la técnica. En términos tecnológicos, el campus cuenta con una infraestructura avanzada que soporta la formación académica en diversas áreas de ingeniería.

Económicamente, al ser una institución pública dependiente de la Secretaría de Educación Pública (SEP), el Tecnológico de Orizaba opera bajo un esquema de financiamiento gubernamental.

Operativamente, el Tecnológico de Orizaba tiene una estructura organizativa bien definida que facilita la gestión de sus múltiples programas académicos y proyectos de investigación. Con una oferta académica que incluye 10 carreras a nivel licenciatura, 6 maestrías, incluyendo la Maestría en Sistemas Computacionales, 2 doctorados, y una especialidad, la institución se caracteriza por su enfoque en la formación de profesionales altamente capacitados. El reconocimiento por parte de organismos certificadores como CACEI (El Consejo de Acreditación de la Enseñanza de la Ingeniería, Asociación Civil), CONAIC (Consejo Nacional de Acreditación en Informática y Computación, Asociación Civil), y SNP (Sistema Nacional de Posgrados) respalda la calidad de la educación impartida y su reputación.

1.3 Planteamiento del problema

La notable disminución en la realización de autopsias en el Hospital Regional Río Blanco (H.R.R.B.) suscitó una preocupación considerable en la dirección del área de patología. Esta situación motivó el desarrollo de un sistema que aplica técnicas de minería de datos para analizar encuestas y descubrir patrones y tendencias en las opiniones del personal médico sobre la reducción de esta práctica. Se emplearon reglas de asociación y redes bayesianas para establecer asociaciones entre las variables que podrían explicar la disminución de autopsias en el H.R.R.B. [11]. Posteriormente, se añadió un módulo de minería de patrones emergentes para comparar los motivos de esta disminución con los de otros hospitales [12], seguido de la implementación de un módulo de descubrimiento de subgrupos [13].

No obstante, persiste un problema fundamental: la falta de un análisis comparativo que valide de manera precisa el conocimiento obtenido hasta ahora. Es crucial determinar si las causas de la disminución de autopsias son comunes a otros hospitales o si son específicas del H.R.R.B. Para abordar esta cuestión, el presente proyecto consistió en el desarrollo de un módulo adicional en el sistema existente, utilizando la técnica de minería de conjuntos de contraste, una metodología de Descubrimiento Supervisado de Reglas Descriptivas (SDRD). Este módulo permite una comparación detallada entre los resultados de las encuestas del H.R.R.B. y otros hospitales, facilitando la identificación de patrones discriminativos que revelen si las causas de la reducción en la práctica de autopsias son generales o particulares.

En este trabajo se desarrolló un nuevo módulo mediante el cual se obtiene un análisis que proporciona una base sólida para que las áreas de patología puedan tomar decisiones informadas y efectivas, con el objetivo de revertir la tendencia actual y aumentar el número de autopsias realizadas.

1.4 Objetivo general y específicos

En esta sección se establecen el objetivo general y los objetivos específicos de la investigación, los cuales guiarán el desarrollo del proyecto y definirán las metas a alcanzar para abordar de manera efectiva la problemática planteada.

1.4.1 Objetivo general

Comparar los resultados de las encuestas sobre la disminución de autopsias en hospitales mexicanos usando técnicas de minería de conjuntos de contraste para la validación de las similitudes y diferencias obtenidas entre las opiniones médicas.

1.4.2 Objetivos específicos

1. Analizar el estado del arte de la aplicación de algoritmos de minería de conjuntos de contraste (CSM) para la selección preliminar de las tecnologías apropiadas para la solución del problema.

2. Realizar un análisis comparativo de las tecnologías para la implementación de los algoritmos de CSM por medio de pruebas de concepto para la selección final de las que se utilizarán para el desarrollo del módulo.

3. Generar los conjuntos de datos con las opiniones de médicos de otros hospitales sobre la disminución de autopsias.

4. Aplicar algoritmos de CSM a los conjuntos de datos para la elección de los más adecuados para su implementación.

5. Desarrollar el módulo del sistema para la comparación de los conjuntos de datos utilizando los algoritmos seleccionados de CSM.

6. Validar el nuevo módulo desarrollado usando un caso de estudio.

1.5 Justificación

A pesar de los avances obtenidos mediante técnicas de minería de datos y descubrimiento supervisado de reglas descriptivas en proyectos previos, persisten interrogantes sobre las causas subyacentes a la disminución de autopsias principalmente en el Hospital Regional Río Blanco (H.R.R.B.) en comparación con otros hospitales.

La técnica de minería de conjuntos de contraste, utilizada en este proyecto, permite identificar patrones discriminativos específicos entre diferentes clases de datos, comparando los resultados de las encuestas del H.R.R.B. con los obtenidos en el Hospital General de Zona 53, Edo. Mex. (H.G.Z.53) y el Hospital General San Juan Bautista Tuxtepec, Oaxaca y ofreciendo una nueva perspectiva para validar y complementar los hallazgos.

Este análisis es crucial para determinar si las razones que han llevado a la disminución de autopsias en el H.R.R.B. son exclusivas de este hospital o si reflejan una tendencia generalizada en otros centros de salud. Al ofrecer resultados más precisos y diferenciados, este proyecto proporciona información valiosa que permitirá a las áreas de patología tomar decisiones informadas y efectivas que

logren incrementar el número de autopsias realizadas, mejorando así la calidad de los servicios médicos y el conocimiento científico en el área de patología.

Capítulo 2. Estado de la práctica

Con el propósito de fundamentar rigurosamente el análisis del tema central de este proyecto, se realizó una revisión exhaustiva de 15 artículos académicos seleccionados por su relevancia y contribución a múltiples áreas de estudio.

El análisis de estos artículos contribuyó crucialmente a la comprensión del estado del arte en técnicas avanzadas, como el descubrimiento supervisado de reglas descriptivas (SDRD) y la minería de conjuntos de contraste, que son esenciales para enmarcar y contextualizar la investigación realizada.

Posteriormente, se sintetizan los aspectos más relevantes de cada artículo revisado.

2.1 Trabajos relacionados

Con el fin de descubrir patrones que resaltan las diferencias entre varias clases de datos, permitiendo así una comprensión más clara de lo que distingue a un grupo de otro, se emplea la minería de conjuntos de contraste, así como sus marcos de trabajo y variantes.

Jayne et al. [14] presentaron un algoritmo de minería de conjuntos de contraste STUCCO (*Search and Testing for Understandable Consistent Contrasts,* Búsqueda y prueba de contrastes consistentes y comprensibles) aplicado en la detección de apnea del sueño.

El objetivo del estudio fue desarrollar un método de cribado preciso, utilizando únicamente señales fisiológicas, tales como la frecuencia cardíaca y la concentración de oxígeno. Estas señales se miden con dispositivos portátiles de uso ordinario. En esta etapa, se buscó identificar las señales y características estadísticas que diferencian los episodios de apnea de los de no apnea. Para lograr esto, se aplicó la técnica de minería de conjuntos de contraste a un conjunto de datos abiertos sobre apnea del sueño.

Asimismo, el conjunto de datos utilizado proviene de la detección del OSASUD *(Obstructive Sleep Apnea Syndrome Detection*, Detección del Síndrome de Apnea

Obstructiva del Sueño), recolectado de 30 pacientes con ictus en una ICU (*Intensive Care Unit,* Unidad de Cuidados Intensivos) del Hospital Universitario de Udine, Italia. Se seleccionaron señales fisiológicas que se miden con dispositivos portátiles para limpiarlas y derivar características estadísticas antes de convertir el conjunto de datos en un formato compatible con el algoritmo de minería de conjuntos de contraste STUCCO.

Los resultados preliminares mostraron reglas de contraste interesantes capaces de guiar investigaciones futuras y el desarrollo de modelos predictivos. Estos hallazgos indicaron que es posible desarrollar métodos menos invasivos y más accesibles para la detección de apnea del sueño utilizando dispositivos portátiles.

Como extensión de la minería de conjuntos de contraste y con relación al área médica, la principal contribución del artículo realizado por Ahmed et al. [15] fue un modelo basado en conjuntos de contraste difusos que utiliza una red de atención profunda para clasificar textos escritos por pacientes mentales en distintos síntomas.

Dicho modelo abordó la problemática relacionada con la salud mental, especialmente en el contexto de la pandemia de COVID-19, provocando el aumento de la necesidad de IDPT (*Internet-Delivered Psychological Treatments*, Tratamientos Psicológicos Entregados por Internet).

Este modelo propuesto empleó técnicas de aprendizaje profundo y NLP (*Natural Language Processing*, Procesamiento de Lenguaje Natural) para extraer y clasificar síntomas de ansiedad y depresión a partir de textos. Adicionalmente, se utilizó *transfer learning* para adaptar y expandir los léxicos de atención posicional ponderada. Las tecnologías empleadas incluyen sistemas basados en aprendizaje profundo y conjuntos de contraste difusos, con una implementación comparada con técnicas tradicionales de clasificación sin incrustación.

Asimismo, los resultados experimentales demostraron que el modelo logró alcanzar una curva ROC (*Receiver Operating Characteristic Curve,* Curva Característica de Funcionamiento del Receptor) alta de 0.82 al clasificar datos de salud mental en nueve síntomas distintos. Esto mejoró significativamente la precisión y

generalización del sistema de aprendizaje. A través de este enfoque, el modelo fue capaz de ofrecer una mejor herramienta para el diagnóstico y tratamiento de problemas de salud mental, aprovechando el análisis de texto y el aprendizaje profundo.

A través de múltiples estudios y subrayando nuevamente la aplicación crucial de la CSM (*Contrast Set Mining,* Minería de Conjuntos de Contraste), Trasierras et al. [16] abordaron el problema de la detección de características comunes y únicas entre diferentes subtipos de cáncer, con el objetivo de mejorar las terapias personalizadas a través de la propuesta de una metodología basada en CSM para capturar relaciones de alto orden en datos transcriptómicos de cáncer.

Esta técnica permitió extraer relaciones genéticas específicas relacionadas con las vías funcionales afectadas por la enfermedad, sin requerir conocimientos previos para su validación.

Las tecnologías empleadas en este estudio incluyeron análisis de expresión génica mediante RNA-Seq (*Sequencing of RNA molecules*, Secuenciación de moléculas de ARN), análisis de supervivencia mediante el método de Kaplan-Meier, y bases de datos como TCGA (*The Cancer Genome Atlas,* El atlas del genoma del cáncer) para la obtención de datos genómicos y clínicos.

En conclusión, los resultados demostraron que el uso de CSM permite identificar relaciones genéticas robustas y sin sesgo.

La aplicación de minería de conjuntos de contraste ha desencadenado su enfoque en múltiples áreas y Qian et al. [17] lo demostraron en su estudio, implementando la variable de CCSM (*Continuous Contrast Set Mining,* Minería de Conjuntos de Contraste Continuo), adoptando la estructura STUCCO incluyendo un conjunto de heurísticas de poda, correcciones de Bonferroni y Python. Se resaltó su distintivo como una metodología que permitió identificar patrones significativos y distinguir entre diferentes grupos de fallos sin necesidad de discretización de datos.

Su aplicación examinó el problema de la depuración de fallos, especialmente en un entorno donde grandes volúmenes de datos de fallos debían analizarse para identificar y resolver problemas de calidad del software.

Se demostró que CCSM generó en promedio mil ciento veinte conjuntos de contraste en cuatrocientos cincuenta y ocho segundos, logrando una aceleración de cuarenta veces en comparación con los enfoques de discretización tradicionales. La metodología se evaluó utilizando un conjunto de datos de producción de Facebook y los resultados indicaron que CCSM fue más rápido y preciso, permitiendo a los desarrolladores identificar y solucionar problemas de manera más eficiente.

Con el propósito de descubrir patrones en conjuntos de datos agrícolas mediante la aplicación del algoritmo STUCCO se realizó el estudio por Priyadarshini *et al.* [18]. Se utilizaron datos de cultivo, para identificar conjuntos de contraste que revelaran diferencias significativas en las distribuciones de atributos entre diferentes grupos de datos.

En primer lugar, se utilizaron técnicas de minería de datos y análisis estadístico para procesar y analizar los datos. El algoritmo STUCCO se empleó para encontrar patrones contrastantes, se implementó utilizando el lenguaje de programación Python, usando Jupyter, Notebook, con el apoyo de las bibliotecas NumPy y Pandas para el procesamiento y análisis de datos.

Del mismo modo, los resultados obtenidos demostraron que el enfoque fue efectivo para caracterizar enfermedades y variedades de cultivos, p. ej., se encontraron patrones específicos que diferenciaban entre cultivos sanos y aquellos afectados por ciertas enfermedades.

Finalmente, los resultados resaltaron la capacidad del algoritmo para descubrir patrones relevantes en conjuntos de datos agrícolas y su eficacia en la caracterización de enfermedades de la soja y las características de las variedades iris. Estos hallazgos subrayaron su potencial para aplicarse en otros dominios que

requieran el análisis de grandes volúmenes de datos para la identificación de patrones significativos.

La aplicación del propio algoritmo STUCCO se empleó de igual forma dentro del análisis del comportamiento de prescripción de los médicos abordado por Varlamis [19]. Se presentó el contexto de la hipertensión y la hiperlipidemia aplicando minería de conjuntos de contraste para encontrar las diferencias en las prescripciones entre pacientes.

Se utilizaron datos de diagnóstico y prescripción de cuatro mil cuatrocientos setenta y tres pacientes a lo largo de tres años, estudiando el cambio en la prescripción médica tras el diagnóstico de reflujo gastroesofágico (K21) o diabetes mellitus insulinodependiente (E10).

El estudio utilizó bases de datos clínicas, técnicas avanzadas de minería de datos, algoritmos de asociación para detectar patrones frecuentes en las prescripciones y herramientas de software incluyendo lenguajes de programación como Java y Python, que fueron esenciales para el procesamiento y análisis de grandes conjuntos de datos médicos.

Se destacaron medicamentos específicos en los resultados del análisis, el enfoque permitió descubrir comorbilidades frecuentes y asociaciones ocultas entre diagnósticos y prescripciones, proporcionando una herramienta útil para los profesionales de la salud y una metodología valiosa para estudiar los patrones de prescripción médica y sus variaciones en respuesta a diagnósticos específicos.

El ámbito médico contiene diferentes enfoques, en este caso el enfoque al problema de la subexplotación de narrativas y la identificación de términos significativos se presentó por Diouf *et al.* [20], con el objetivo de especificar la causa de muerte a partir de autopsias verbales narrativas.

La principal contribución se posicionó en desarrollar un enfoque innovador que combinó la técnica TF-IDF (*Term Frequency-Inverse Document Frequency,* Término-Frecuencia Inversa de Documento) y el modelo de temas LDA (*Latent Dirichlet Allocation,* Asignación Latente de Dirichlet*)*, lo que permitió analizar

grandes volúmenes de datos textuales. Utilizando un enfoque no supervisado, se integró el algoritmo de minería CoARTEX (*Corpus Augmented by a Term Rewarding Algorithm from an Indexation Technique,* Corpus aumentado por un algoritmo de recompensa de términos a partir de una técnica de indexación) y la métrica MePTIC (*Metric for Weighting Informative Terms in a Corpus*, Métrica para ponderar términos informativos en un corpus), apoyados en bibliotecas como scikit-learn, gensim, spaCy y NLTK *(Natural Language Toolkit,* Kit de Herramientas de Lenguaje Natural). Los resultados demostraron que la combinación de TF-IDF y LDA fue efectiva para descubrir términos significativos, mejorando la precisión y eficiencia en la identificación de causas de muerte, reduciendo tiempo y recursos, así como validando su coherencia con diagnósticos médicos.

Con un enfoque alineado a la minería de conjuntos de contraste, en la investigación de Machorro-Cano et al. [21] se empleó una metodología basada en EPM (*Emerging Pattern Mining*, Minería de Patrones Emergentes) para analizar un conjunto de datos recopilados de expertos médicos sobre la disminución de autopsias en el Hospital Regional Río Blanco, Veracruz, México, con el propósito de comprender sus implicaciones en la práctica médica.

El análisis se realizó mediante once algoritmos de EPM integrados en el EPM-Framework desarrollado en Java, entre los que se incluyen iEPMiner (*Interesting Emerging Pattern Miner*, Minería de Patrones Emergentes Interesantes), LCMine (*Logical Complex Miner*, Minería de Complejos Lógicos), EP-Random Forest, CEPMine (*Crisp Emerging Pattern Mining*, Minería de Patrones Emergentes Definidos), EvAEP (*Evolutionary Algorithm for Extracting Emerging Patterns*, Algoritmo Evolutivo para la Extracción de Patrones Emergentes), DGCP-Tree (*Dynamically Growing Contrast Pattern Tree*, Árbol de Patrones de Contraste de Crecimiento Dinámico), SJEP-C (*Strong Jumping Emerging Patterns Classifier*, Clasificador de Patrones Emergentes de Salto Fuerte), Top-k Minimal SJEPs (*Top-k Minimal Strong Jumping Emerging Patterns*), Tree-based JEP-C (*Jumping Emerging Pattern Classifier*, Clasificador de Patrones Emergentes de Salto), DeEPS (*Decision-making by Emerging Patterns*, Toma de Decisiones mediante

Patrones Emergentes) y BCEP (*Bayesian Classification based on Emerging Patterns*, Clasificación Bayesiana basada en Patrones Emergentes).

El desempeño de los algoritmos fue evaluado mediante métricas como WRAcc (*Weighted Relative Accuracy*, Precisión Relativa Ponderada), TPR (True Positive Rate, Tasa de Verdaderos Positivos) y FPR (False Positive Rate, Tasa de Falsos Positivos), destacándose particularmente iEPMiner y LCMine por su capacidad para identificar patrones altamente discriminativos y estadísticamente relevantes. Los resultados permitieron identificar patrones significativos asociados a la tendencia decreciente en la realización de autopsias, aportando evidencia empírica sobre factores institucionales, profesionales y perceptuales involucrados en dicho fenómeno.

En conjunto, el estudio introdujo una aplicación innovadora de técnicas de minería de patrones emergentes en el ámbito médico, abordando las razones para solicitar o rechazar autopsias, las causas de su disminución y los perfiles del personal hospitalario responsable de su solicitud, constituyéndose como un antecedente metodológico fundamental para la investigación presente.A través de la minería basada en patrones de contraste se abordó la caracterización de los visitantes de un sitio web en términos de conversiones, es decir, acciones con un objetivo comercial, por Huidobro *et al.* [22].

Su principal contribución fue desarrollar una metodología que agrupó las actividades de los visitantes en sesiones y asoció cada conversión a un paso del embudo de ventas, generando patrones de contraste únicos. Se utilizaron herramientas WAS (*Web Analytics Solutions,* Soluciones de Análisis Web) como Google Analytics y Matomo, junto con algoritmos de minería de patrones, para identificar y ordenar patrones relevantes que permitieron diferenciar a los visitantes. En los resultados se destacó que pocos patrones eran suficientes para caracterizar eficazmente a los visitantes, facilitando el diseño de estrategias de *marketing* más efectivas, usando cada conversión como un atributo binario para crear dos clases: visitantes que realizaron la conversión y visitantes que no. Esta metodología ofreció una solución innovadora para el análisis del comportamiento del usuario en línea y

proporcionó una base sólida para la toma de decisiones estratégicas en *marketing*, con planes de incorporar datos de diferentes canales de venta en el futuro.

En contraste con el trabajo de Diouf et al. [20], Mapundu et al. [23] abordaron la determinación de la causa de muerte desde una perspectiva diferente. Su enfoque abordó la introducción de un marco explicable de SEDL *(Stacked Ensemble Deep Learning*, Aprendizaje Profundo de Conjunto Apilado) para revisar narrativas VA (*Verbal Autopsies*, Autopsias Verbales) y generar predicciones precisas y fácilmente interpretables de CoD *(Cause of Death*, Causa de Muerte). Este enfoque no solo fortaleció la capacidad predictiva del modelo, sino que además integró principios de aprendizaje automático y procesamiento del lenguaje natural para mejorar la comprensión de las narrativas clínicas y facilitar la toma de decisiones médicas.

Por tal motivo, se emplearon tecnologías avanzadas de NLP y arquitecturas de aprendizaje profundo como LSTM (*Long Short-Term Memory*, Memoria a Corto y Largo Plazo) y CNN *(Convolutional Neural Network,* Red Neuronal Convolucional). La combinación de estos modelos permitió superar el rendimiento individual de LSTM y CNN, ofreciendo un enfoque más robusto para analizar narrativas de VA. Se alcanzó una precisión del 82% en los conjuntos de entrenamiento y prueba, y se mejoró la interpretabilidad del sistema mediante LIME *(Local Interpretable Model-agnostic Explanations,* Explicaciones Locales Interpretables Independientes del Modelo*)*, facilitando la comprensión de las predicciones por parte de los profesionales de la salud. Este enfoque incrementó la precisión y eficiencia en la predicción de CoD, redujo costos y minimizó errores, representando un avance significativo en la salud pública y en la aplicación de autopsias verbales.

La relación entre el tema de autopsias y aprendizaje automático, tratada por Mapundu et al. [23], encuentra un enfoque complementario en el trabajo de Shahid et al. [24], donde se abordó la clasificación automática de informes de autopsias. En este contexto, se destacó como principal contribución el desarrollo de un sistema de aprendizaje automático que emplea técnicas avanzadas de NLP, mejorando así

la eficiencia y precisión en la identificación de CoD a través de la categorización y análisis automatizado de dichos informes.

Se diseñó una arquitectura que combinó algoritmos de clasificación de texto como SVM *(Support Vector Machines, Máquinas de Vectores de Soporte*) y *Naive Bayes*, seleccionados por su capacidad para manejar grandes volúmenes de datos y su efectividad en la clasificación. Se implementaron técnicas de extracción de características como TF-IDF y análisis de palabras clave para mejorar la precisión. Python fue el lenguaje de programación utilizado, apoyado por bibliotecas como scikit-learn, NLTK y spaCy. La metodología ágil permitió ajustar el modelo iterativamente y se demostró una alta precisión en la clasificación de textos y se destacó su aplicación en el ámbito forense para identificar CoD. En resumen, el estudio presentó una solución innovadora para el análisis de informes de autopsia mediante técnicas de aprendizaje automático y procesamiento del lenguaje natural, ofreciendo una herramienta valiosa para la práctica forense.

El cáncer, tratado en el estudio de Trasierras et al. [16], también se identificó como el enfoque principal en la investigación de Chang et al. [25]. Sin embargo, en esta investigación se introdujo un marco avanzado de minería de patrones de contraste aplicado a los datos de ClinicalTrials.gov, con el objetivo de identificar los factores determinantes en el éxito o fracaso de los ensayos clínicos de medicamentos contra el cáncer.

El estudio se centró en la alta tasa de fracasos en ensayos clínicos de medicamentos contra el cáncer, atribuidos principalmente a la eficacia y seguridad insuficientes, así como a problemas en el diseño del ensayo y razones comerciales. El marco propuesto utilizó técnicas avanzadas como el procesamiento del lenguaje natural y la minería de patrones de contraste para identificar factores clave. Implementado con Python y sistemas gestores de bases de datos, este enfoque alcanzó una precisión del 80% en la predicción de resultados, proporcionando una herramienta valiosa para optimizar el diseño de futuros ensayos clínicos y mejorar el desarrollo de medicamentos contra el cáncer.

Las técnicas de aprendizaje automático se aplicaron en la investigación de Chen et al. [26], a un tema crítico relacionado con los factores que influyeron en el suicidio entre los ancianos en áreas rurales de China.

Se utilizaron datos de autopsias psicológicas y técnicas de aprendizaje automático para identificar asociaciones entre el suicidio y factores como la depresión, la calidad de vida, el apoyo social y el estado civil. Se emplearon cuatro modelos CART (*Classification and Regression Tree*, Árboles de Clasificación y Regresión) para jerarquizar estos factores, destacando la depresión como el predictor más significativo. El enfoque permitió revelar relaciones complejas no identificadas por métodos tradicionales, resaltando la importancia de enfoques analíticos innovadores para comprender el suicidio. Los hallazgos subrayan la necesidad de intervenciones personalizadas para prevenir el suicidio en ancianos de áreas rurales de China.

El tema de la clasificación de CoD a partir de informes de VA en regiones sin un registro confiable de defunciones, común en países de ingresos bajos y medianos se estudió por Manaka et al. [26]. La principal contribución del trabajo se centró en el desarrollo de una arquitectura basada en aprendizaje por transferencia, utilizando BERT (*Bidirectional Encoder Representations of Transformers,* Representaciones de Codificador Bidireccional de Transformadores*)* y *ELMo* (*Embeddings from Language Models*, Incrustaciones de *Modelos de Lenguaje)* previamente entrenados en dominios de inglés general y salud para mejorar la clasificación mediante técnicas NLP. El sistema combinó representaciones textuales con características binarias de datos estructurados, mejorando significativamente la precisión.

Finalmente, se destacó la importancia de adaptar estas técnicas para mejorar la calidad de los datos de salud en regiones con recursos limitados, ofreciendo una base más sólida para la toma de decisiones en políticas de salud pública y transformando la clasificación de causas de muerte en contextos con datos limitados.

Partiendo del propio tema de narrativas de VA, en el artículo de Fan *et al.* [27] se estudió la inferencia válida utilizando predicciones de modelos de lenguaje basadas en dichas narrativas en contextos donde la mayoría de las muertes ocurren fuera del sistema de salud.

El estudio presentó el método llamado multiPPI++ (*Prediction-powered inference for Multinomial Classification*, Inferencia basada en predicciones para la clasificación multinomial), que extiende la inferencia a la clasificación multinomial utilizando técnicas avanzadas de NLP, métodos basados en BoW (*Bag-of-Words*, Bolsa de Palabras) usando herramientas de la biblioteca de Python, algoritmos, y modelos como BERT y LLM (*Large Language Model*, Modelo de Lenguaje Grande), incluyendo el modelo GPT (*Generative Pre-trained Transformer*, Transformador Generativo Preentrenado).

Además, se abordaron problemas de transportabilidad y sesgo en datos de autopsias verbales, mejorando la precisión de las estimaciones de causas de muerte. Se subrayó la importancia de datos etiquetados de alta calidad y la revisión de inferencias en la toma de decisiones de salud pública. En conclusión, multiPPI++ ofreció una solución innovadora para mejorar la precisión de estimaciones en contextos de salud pública con recursos limitados.

2.2 Análisis comparativo

La Tabla 2.1 presenta un análisis comparativo de los artículos revisados en relación con este proyecto. En ella se detallan la problemática abordada, la contribución principal, las tecnologías empleadas, los resultados obtenidos y el estado actual de cada estudio, lo que permite establecer un marco de referencia sólido para contextualizar la investigación en curso.

Tabla 2.1 Análisis comparativo de artículos relacionados

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Jayne et al. [14]	Dificultad de acceso y alto costo de métodos actuales de detección y diagnóstico de la apnea del sueño.	Aplicación de minería de conjuntos de contraste para la comprensión de los tipos de señales y características estadísticas específicas para diferenciar el estado de apnea del sueño y el estado de no apnea.	Algoritmo de minería de conjuntos de contraste STUCCO.	Regla única relevante que indica a pacientes susceptibles cuando la frecuencia cardíaca se encuentre entre ciertos parámetros de LPM (Latidos Por Minuto), al igual que la concentración de oxígeno. Menor apoyo a reglas en nivel individual, pero mayor confianza y elevación en comparación con las reglas a nivel de cohorte.	Investigación finalizada. **Desarrollo futuro:** Se presentó como base para un modelo predictivo futuro.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Ahmed et al. [15]	Identificación y clasificación de síntomas de depresión a partir de textos generados por pacientes, provocado por el aumento de ansiedad y depresión tras el confinamiento, COVID-19 y situaciones actuales.	Modelo basado en contraste difuso usando una red de atención para palabras ponderadas posicionalmente, para detectar y clasificar con mayor precisión síntomas depresivos a través del texto escrito por pacientes mentales.	PHQ-9, Método CSEP Repositorio de base de datos de léxico: WordNet. Estándar de codificación UTF-8. Modelo preentrenado para el vector global GloVe. Servicio Amazon *Mechanical Turk*. Diseño LSTM bidireccional.	Modelo con un ROC en el conjunto de desarrollo del 0.82, y en el conjunto de entrenamiento del 0.91, alta tasa positiva, reconocimiento de la palabra objetivo en la tarea, aprendizaje de síntomas a través del texto y visualización del peso de los conjuntos de contraste para las palabras ponderadas.	Investigación finalizada. **Desarrollo futuro:** Continuación de investigación para mejorar el rendimiento del modelo y la evolución de un algoritmo más adaptativo para clasificación de texto y ajustes.
Trasierras et al. [16]	Detección de características comunes y únicas entre subtipos de cáncer.	Enfoque basado en minería de conjuntos de contraste para capturar relaciones genéticas específicas.	Minería de conjuntos de contraste. RNA-Seq. Análisis de vías funcionales.	Identificación de grupos de genes relacionados con subtipos específicos de cáncer, proporcionando información para terapias personalizadas.	Investigación finalizada.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Qian et al. [17]	Análisis de características de grupos de fallos conjuntos para asistir a los desarrolladores en la clasificación y resolución de fallos.	Se propuso la minería continua de conjuntos de contraste como el primer enfoque aplicado sin la necesidad de discretización. Se descubrió la detección de patrones anómalos a través del análisis automatizado de accidentes.	Algoritmo de minería de conjuntos de contraste continuo con modificaciones de la estructura STUCCO (conjunto de heurísticas de poda y correcciones de Bonferroni).	La minería de conjuntos de contraste continuo ofreció una aceleración fundamental cuarenta veces más rápida que la discretización con diez contenedores y diez veces más rápida que la discretización con tres contenedores. Generó conjuntos de contraste de mayor calidad. La nueva puntuación de anomalía favoreció la detección de anomalías con bajo soporte esperado.	Propuesta de un algoritmo finalizado para la minería de patrones en datos continuos. **Desarrollo futuro:** Experimentos con métodos no paramétricos para validar la significancia estadística de los resultados.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Priyadarshini et al. [18]	Descubrimiento de patrones significativos para describir enfermedades en cultivos de soja y las características de las variedades iris mediante la aplicación del algoritmo STUCCO.	Se demostró la aplicabilidad de la minería de conjuntos de contraste para la detección de patrones utilizando conjuntos de datos agrícolas. Se identificaron características que provocan las enfermedades de la soja y las características distintivas de las variedades iris.	Algoritmo STUCCO implementado en Python. Jupyter Notebook con el apoyo de las bibliotecas NumPy y Pandas para el procesamiento y análisis de datos.	Apoyo en el descubrimiento de patrones y en la selección e identificación de características que causan las enfermedades posteriores a un diagnóstico.	Investigación avanzada con resultados preliminares concluyentes. Desarrollo futuro: Oportunidades para trabajos futuros y mejoras en áreas para el desarrollo continuo y aplicaciones futuras de los hallazgos.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Varlamis [19]	Identificación y detección de patrones en el cambio de comportamiento de prescripción médica después de la detección de trastornos específicos.	Demostración de la correlación entre trastornos de hipertensión e hiperlipidemia, en relación con el reflujo gastroesofágico y diabetes mellitus insulinodependiente en el cambio de patrones de prescripción. Localización de patrones frecuentes e identificación de los compuestos de sustancias.	Minería de conjuntos de contraste implementado en Java y Python. Scripts en Python para el procesamiento inicial de datos y la extracción de conocimiento. Algoritmo FP-Growth (*Frequent Pattern Growth*, Crecimiento de patrón frecuente) de la biblioteca ML-Extend y el algoritmo STUCCO, para CSM.	Se localizaron patrones frecuentes en los compuestos de sustancias agregados, eliminados o continuados para los pacientes después del diagnóstico de los trastornos estudiados.	El estudio se encontraba en una fase de desarrollo inicial a intermedia. **Desarrollo futuro:** Recomendaciones para realizar más investigaciones con conjuntos de datos más amplios y diversos, provenientes de múltiples médicos, para verificar y ampliar los hallazgos presentados.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Diouf *et al.* [20]	Subexplotación de narrativas de autopsias verbales no estructuradas debido a mecanismos de procesamiento inadecuados y técnicas computacionales inapropiadas.	Se desarrolló el algoritmo CoARTEX, el cual generó un corpus aumentado y mejoró la precisión en la extracción de causas de muerte.	Algoritmo Coartex utilizando la métrica MePTIC y el lenguaje de programación Python. Bibliotecas scikit-learn, gensim, spaCy y NLTK para la aplicación del modelo TF-IDF y LDA. Herramienta GridSearchCV para optimizar los parámetros del modelo.	Se reveló que las mediciones realizadas mostraron el mismo patrón. Al utilizar el algoritmo CoARTEX se destacó el aspecto de demora en su ejecución, se identificó debilidad en ciertos valores y se alcanzó el punto máximo en lo relacionado con causas de muerte. Se identificaron las existencias de coocurrencias de términos.	El artículo llegó a una fase avanzada de investigación y desarrollo. **Desarrollo futuro:** Necesidad de realizar investigaciones futuras y pruebas adicionales para mejorar y comparar el método con otros enfoques.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Machorro—Cano *et al.* [21]	Necesidad de identificar las causas de la reducción de autopsias en el Hospital Regional Río Blanco en los últimos años.	Se obtuvieron conocimientos acerca de las razones, motivos y circunstancias que tuvieron los médicos del Hospital Regional Río Blanco para solicitar una autopsia, las causas de la negativa de la autopsia por parte de los familiares de los pacientes fallecidos, el porqué no se realizaron suficientes autopsias en el hospital y el personal hospitalario correspondiente para solicitar una autopsia según los médicos.	Se emplearon once algoritmos EPM, incluidos en EPM-Framework de Java: EP-Random Forest, LCMine, CEPMine, EvAEP, DGCP-Tree, SJEP-C, Top-k minimal SJEPs, Tree-based JEP-C, iEPMiner, DeEPS.	Se demostró un rendimiento más rápido a través de iEPMiner y LCMine, se produjeron patrones emergentes superiores, considerando métricas como confianza, WRACC, FPR y TPR. Se identificaron con base en LCMine, patrones emergentes que relacionan las opiniones médicas sobre el personal adecuado para solicitar las autopsias y sus motivos para este estudio.	Una parte significativa del análisis de datos se completó en el estudio. **Desarrollo futuro:** Aplicar la investigación a otros centros y la implementación de herramientas tecnológicas para facilitar la comparación y el análisis continuo.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Huidobro *et al.* [22]	Reconocimiento de características de los visitantes de un sitio web en términos de conversiones.	Enfoque de minería de patrones para caracterizar diferentes clases de visitantes en función de los objetivos comerciales, a través de un conjunto de patrones de contraste para los visitantes. Metodología replicable utilizando blogs de cualquier sitio web como datos de entrada.	WAS, tales como, *Google Analytics* y *Matomo*. Minería basada en patrones de contraste: *Bagging-RandomMiner* para detección de anomalías, el algoritmo DTB (*Decision Tree-Based, Basado en arboles de decisión*) y la implementación Weka de PBC4cip (*Pattern-Based Classification for Classification and Clustering with Integrated Pattern mining,* Clasificación Basada en Patrones para Clasificación y Agrupamiento con Minería de Patrones Integrada).	Se presentó una solución innovadora para el análisis de datos de comportamiento de usuarios en línea, utilizando técnicas avanzadas de minería de patrones y herramientas de análisis web.	Investigación finalizada. **Desarrollo futuro:** El estudio presentó oportunidades de mejora que requieren más estudios y validaciones adicionales. Se propusieron trabajos que incluyan el consenso de más expertos, el desarrollo de herramientas de software y la incorporación de análisis en línea y datos de diferentes canales de venta.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Mapundu *et al.* [23]	Determinación de la causa de muerte mediante autopsias verbales con brechas menores a través de la aplicación de arquitecturas SEDL.	Se amplió la investigación de XAI (*Explainable Artificial Intelligence,* Inteligencia Artificial Explicable*)* y su aplicación en la atención médica para construir modelos DL (*Deep Learning*, Aprendizaje Profundo) confiables con la finalidad de mejorar informes VA. Se determinó que el marco SEDL se incluye en el proceso para determinar las causas de mortalidad, junto con la anotación e interpretación humana.	Métodos DL. Arquitecturas SEDL: LIME, CNN y LSTM. Curvas de aprendizaje automático: ML. Variante LSTM de RNN (*Recurrent Neural Network*, Recurrente Red Neuronal) y CNN. *Python Spyder.* Biblioteca TextBlob, NLTK, Python StandarScaler y Keras. PostgreSQL versión 4.2. Vector global GloVe.	Se demostró que la arquitectura SEDL superó a LSTM y CNN en clasificación de enfermedades, con una precisión de entrenamiento del 82.1%. Se obtuvieron diferentes predicciones e interpretabilidad de CoD en los modelos, siendo LIME la mejor opción explicativa. Se identificó que el marco SEDL se posiciona como una alternativa alentadora al diagnóstico médico.	Investigación finalizada. **Desarrollo futuro:** Futura aplicación de vinculación de registros para identificar casos de morbilidad no notificados mediante la conexión de datos del HDSS (*Health and Demographic Surveillance System*, Sistema de Vigilancia Demográfica y de Salud) y NHLS (*National Health Laboratory Service*, Servicio Nacional de Laboratorio de Salud).

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Shahid *et al.* [24]	Reducción de tiempo e irregularidades a través de la clasificación de los informes de autopsias médicas en texto plano.	Se propuso una técnica que contribuyó al campo de la clasificación automática de textos médicos al demostrar mejoras en la precisión mediante la reducción del tamaño de las características y al utilizar un conjunto de datos más amplio y variado en términos de enfermedades.	Técnicas de ATC (*Automatic Text Classification*, Clasificación Automática de Texto). Algoritmo de extracción de características: TF (*Term Frequency*, Frecuencia del término) y TF-IDF. Anaconda de Python y R. Paquetes de ciencia de datos. Biblioteca de Python Sklearn. Algoritmo SVM. Método clasificador NB.	Se identificó que la metodología propuesta funcionó mejor en comparación con los artículos base en términos de precisión con más clases y menos tamaño de conjunto de características.	Investigación finalizada. **Desarrollo futuro:** Técnica como referencia para próximos trabajos y punto de comparación para sugerencias futuras con los métodos ATC actuales, con el objetivo de clasificar un número más significativo de enfermedades con mayor precisión y al mismo tiempo reducir el tamaño del conjunto de características.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Chang *et al.* [25]	Comprensión de las características comunes de los ensayos clínicos exitosos y no exitosos de medicamentos contra el cáncer.	Se desarrolló un marco informativo a través del cual se permitió a los investigadores de ensayos identificar patrones interpretables que estén asociados con los resultados del ensayo, así como los patrones a seguir y evitar.	Minería de patrones de contraste. MetaMap. Tecnología NLP. Base de datos DrugBank con ayuda de la paquetería *dbparser* y *Criteria2Query*. Software *Clinical-Trial-Parser* NLP.	Se sugirió basándose en los resultados de la clasificación que la minería de contraste es capaz de extraer características del ensayo asociadas significativamente con los resultados del ensayo, siendo los patrones de contraste, el clasificador que condujo a un mejor diseño de las pruebas y ayudó a los inversores en las pruebas a gestionar los riesgos.	Investigación finalizada. **Desarrollo futuro:** Consideración de un mayor desarrollo hasta convertirse en un sistema de recomendación de patrones que ofrezca retroalimentación en tiempo real a los investigadores de los ensayos.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Manaka et al. [26]	Clasificación de causas de muerte a partir de informes de autopsias verbales en regiones donde no existe un registro de defunciones confiable, tal como países de ingresos bajos y medianos.	Arquitectura con base en dos paradigmas de aprendizaje por transferencia: el aprendizaje monolingüe y la adaptación de dominio de múltiples fuentes para mejorar las narrativas de *autopsia verbal*, para clasificar la causa de muerte a través de la extracción de características.	Modelos de *BERT* y *ELMo* preentrenados en dominios de inglés y salud. Método SMOTE-Tomek Links (*Synthetic Minority Over-sampling Technique*, Técnica de Sobremuestreo Sintético de la Minoría y Enlaces de Tomek).	El sistema de aprendizaje por transferencia mejoró las tareas de clasificación de *CoD.* La combinación de características binarias de VA y características de texto narrativo aprendidas mejoró la tarea de clasificación de CoD.	Investigación finalizada. **Desarrollo futuro:** Próximo trabajo para evaluar la efectividad de esta arquitectura en otros subdominios del inglés y analizar diversas técnicas de incrustación a nivel de caracteres y palabras.

Tabla 2.1 Análisis comparativo de artículos relacionados – continuación

Artículo	Problema	Contribución	Tecnologías	Resultados	Estado
Fan *et al.* [27]	Convertir las VA en conocimientos prácticos, logrando así la inferencia válida, empleando predicciones de modelos de lenguaje para permitir la conversión.	Desarrollo del método "multiPPI++" para la inferencia válida, empleando CoD predichas a partir de texto de formato libre utilizando técnicas de NLP de última generación. Contribución al avance de los modelos predictivos en salud pública.	Técnicas de NLP: métodos basados en BoW. Herramientas de la biblioteca de Python "scikit-learn": CountVectorizer, LabelEncoder y TfidVectorizer, word_tokenize de NLTK (*Natural Language Toolkit*, kit de herramientas de lenguaje natural). Algoritmos SVM, KNN (*K Nearest Neighbors*, K Vecinos más Cercanos) y NB (*Naive Bayes*). Modelo BERT y LLM: modelo GPT-4-32. Técnica PPI (*Prediction-based Inference*, Inferencia basada en predicción) y su extensión PPI++.	GPT-4 obtuvo la mayor precisión (0.75), lo cual indicó que es ligeramente mejor o casi equivalente a los otros métodos de NLP probados, e identificó correctamente una falla en la limpieza de datos en el conjunto de datos PHMRC (*Population Health Metrics Research Consortium*, Consorcio de Investigación sobre Métricas de Salud de la Población). Se determinó que los coeficientes de regresión ingenuos de KNN y GPT-4 no están tan lejos de la verdad fundamental como los métodos de NB y SVM.	Investigación finalizada. **Desarrollo futuro:** Necesidad de más estudios para mejorar la precisión de la estimación de parámetros, abordar el desbalance de clases en CoD, y optimizar el rendimiento de los modelos en idiomas no ingleses.

A partir del análisis de los artículos relacionados con el tema principal de esta investigación, se identificaron múltiples perspectivas sobre la aplicación de técnicas avanzadas de minería de datos, incluyendo la minería de conjuntos de contraste, así como diversas arquitecturas, algoritmos y metodologías en contextos variados. Estos estudios demuestran la adaptabilidad y aplicabilidad de estas metodologías a distintos problemas y conjuntos de datos, obteniendo resultados significativos y prácticos.

La incorporación de estos conocimientos en el presente trabajo subraya la versatilidad y potencia de la minería de conjuntos de contraste como una herramienta para el análisis de datos complejos.

La minería de conjuntos de contraste se abordó en Trasierras et al. [16], aplicando dicho enfoque para capturar relaciones genéticas específicas.

La aplicación de minería de conjuntos de contraste con enfoque en el algoritmo STUCCO presenta aplicación en el artículo de Jayne et al. [14], para la comprensión de los tipos de señales y características estadísticas específicas para diferenciar el estado de apnea del sueño y el estado de no apnea. Priyadarshini et al. [18] demostraron la aplicabilidad del mismo algoritmo para la detección de patrones en conjuntos de datos agrícolas.

En relación con los trabajos anteriores, Varlamis [19] localizó patrones frecuentes en los compuestos de sustancias agregados, eliminados o continuados para los pacientes después del diagnóstico de los trastornos estudiados, implementando minería de conjuntos de contraste y el algoritmo STUCCO.

Asimismo, Qian et al. [17] propusieron un algoritmo de minería de conjuntos de contraste continuo, modificando la estructura STUCCO, como el primer enfoque aplicado sin la necesidad de discretización al análisis de características de fallos conjuntos.

En un enfoque relacionado, Ahmed et al. [15] extendieron el contraste difuso con una red de atención para mejorar la detección y clasificación de síntomas depresivos en textos escritos por pacientes.

En contraste, Diouf et al. [20], propusieron el algoritmo de minería de conjuntos de contraste CoARTEX que generó un corpus aumentado y mejoró la precisión en la extracción de causas de muerte.

Con relación crucial en cuanto al tema de la presente investigación, la minería de patrones emergentes se aplicó en Machorro-Cano et al. [21] para conocer los motivos de reducción de autopsias en el Hospital Regional Río Blanco en los últimos años.

Como un enfoque híbrido que combinó la minería de patrones emergentes con el aprendizaje profundo, se presentó la arquitectura SEDL para determinar las causas de mortalidad, junto con la anotación e interpretación humana por Mapundu et al. [23].

El enfoque de la minería de patrones de contraste, también se exploró en Huidobro et al. [22], quienes utilizaron tecnologías WAS para caracterizar clases de visitantes según objetivos comerciales. Chang et al. [25] desarrollaron un marco informativo que permitió a investigadores de ensayos identificar patrones interpretables asociados con resultados de ensayos clínicos.

En Shahid et al. [24], a través de técnicas de ATC, se propuso una técnica que contribuyó al campo de la clasificación automática de textos médicos al demostrar mejoras en la precisión mediante la reducción del tamaño de las características y al utilizar un conjunto de datos más amplio y variado en términos de enfermedades.

En función al aprendizaje automático, se abordó en Chen et al. [26], para identificar los mecanismos que sustentan las relaciones complejas entre los factores influyentes y el riesgo de suicidio con datos de autopsias psicológicas y métodos de aprendizaje automático. En Manaka et al. [27], se emplearon los modelos de BERT y ELMo para mejorar las narrativas de autopsia verbal, para clasificar la causa de muerte a través de la extracción de características, y en Fan et al. [28], se desarrolló el método “multiPPI++” para la inferencia válida, empleando CoD predichas a partir de texto de formato libre utilizando técnicas de NLP de última generación.

En conclusión, el objetivo de este proyecto es aplicar minería de conjuntos de contraste para comparar resultados de encuestas sobre la disminución de autopistas en hospitales mexicanos.

2.3 Propuesta de Solución

Este proyecto contempla el desarrollo de un módulo de minería de conjuntos de contraste que permite comparar los resultados de encuestas aplicadas a médicos de distintos hospitales mexicanos sobre la disminución de autopsias. El módulo analiza e interpreta los datos mediante modelos basados en reglas explicables, lo que facilita la identificación de diferencias entre instituciones y apoya la toma de decisiones clínicas.

Para implementar el módulo, se utilizó RuleKit2-CS como herramienta de minería de datos, ya que esta tecnología basada en Java permite generar modelos comprensibles, adecuados para contextos clínicos donde la transparencia en los resultados es esencial [29, 30, 31, 32, 33]. En el desarrollo de la interfaz se empleó JSF (JavaServer Faces), marco de trabajo presente en el sistema actual, que sigue el patrón de diseño Modelo-Vista-Controlador (MVC) y facilita la organización de los componentes, mejora el mantenimiento y permite una escalabilidad eficiente. Asimismo, se usó Java como lenguaje de programación, por su robustez, compatibilidad y adopción en entornos empresariales; y NetBeans como entorno de desarrollo, el cual permite trabajar de forma modular y se integra de manera directa con Apache Tomcat, servidor que aloja la aplicación. Para la gestión de datos, se empleó PostgreSQL como sistema gestor de base de datos, debido a su capacidad para manejar grandes volúmenes de información, su eficiencia en consultas complejas y su disponibilidad como software libre. Durante el desarrollo se aplicó la metodología ágil Kanban, la cual organiza visualmente el flujo de trabajo, permite identificar cuellos de botella y facilita la incorporación de mejoras de forma continua. En la fase de análisis de datos, se aplicó la metodología KDD que estructura el proceso desde la selección y transformación de los datos hasta la evaluación e interpretación de los modelos obtenidos. Esta propuesta plantea una solución integral, alineada con la infraestructura tecnológica existente, y fortalece las

capacidades analíticas del sistema para abordar la disminución de autopsias mediante técnicas avanzadas de minería de datos.

Capítulo 3. Aplicación de la Metodología

Este capítulo describe el proceso propuesto para el diseño y desarrollo del módulo de análisis comparativo mediante minería de conjuntos de contraste. Se estructura en cuatro apartados. En el primero, se presenta una revisión de alcance que fundamenta teóricamente el proyecto, centrada en los enfoques de minería de conjuntos de contraste (CSM) y minería de patrones de contraste (CPM), lo cual permite identificar tendencias, vacíos y áreas de oportunidad relevantes. En la segunda sección, se lleva a cabo la selección de los algoritmos de CSM más adecuados, a partir de un análisis comparativo que considera criterios técnicos y metodológicos. Posteriormente, se detalla el diseño del sistema propuesto, incluyendo su arquitectura, componentes y funcionamiento previsto. Finalmente, se describe la estrategia de desarrollo que guía la implementación e integración del módulo al sistema existente.

3.1 Análisis

Con el objetivo de fundamentar teóricamente el desarrollo del módulo de análisis comparativo de encuestas médicas mediante minería de conjuntos de contraste, se realizó una revisión de alcance centrada en dos enfoques principales: la minería de conjuntos de contraste (*Contrast Set Mining*, CSM) y la minería de patrones de contraste (*Contrast Pattern Mining*, CPM). Esta revisión permitió identificar tendencias, vacíos y oportunidades de investigación en el uso de estas técnicas para analizar percepciones médicas en torno a fenómenos complejos como la reducción de las autopsias en hospitales.

La revisión de alcance sigue el marco metodológico de Arksey y O'Malley [34], diseñado para estudios de este tipo, junto con las recomendaciones de Levac sobre su aplicación [35]. Además, se adopta el modelo PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*, Elementos de Informe Preferidos para Revisiones Sistemáticas y Metaanálisis) de Moher et al. [36] para la presentación de revisiones sistemáticas y metaanálisis, incluyendo su extensión PRISMA-ScR (*Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews, Elementos de informe preferidos para revisiones sistemáticas y extensión de metanálisis*

para revisiones de alcance). La organización y presentación de los hallazgos se basa en el trabajo de Tricco et al. [37]. La revisión de alcance se desarrolla en cinco fases: (1) identificación de preguntas de investigación, (2) búsqueda y selección de estudios relevantes, (3) criterios de inclusión y exclusión para la selección de estudios, (4) extracción y análisis de datos, y (5) síntesis y presentación de los hallazgos.

3.1.1 Preguntas de Investigación

Para alcanzar los objetivos de la revisión, se formulan tres preguntas de investigación, listadas en la Tabla 3.1. Estas preguntas sirven como base para el proceso de revisión, aclaran la razón fundamental del estudio y proporcionan una estructura para presentar los hallazgos.

Tabla 3.1. Preguntas de investigación y motivaciones

Identificador	Pregunta	Motivación
PI1	¿Cuáles son los algoritmos más utilizados en la literatura para la aplicación de CSM y CPM?	Identificar los algoritmos más empleados en la literatura, analizando sus principios fundamentales, ventajas, limitaciones y las métricas de calidad utilizadas para evaluar su desempeño en distintos contextos.
PI2	¿Qué lenguajes de programación son los más usados para la implementación de técnicas de CSM y CPM?	Explorar los lenguajes de programación más empleados en la implementación de CSM y CPM, evaluando su impacto en la eficiencia del procesamiento de datos, la escalabilidad y la facilidad de integración con herramientas de análisis.
PI3	¿Cuáles son los dominios de aplicación donde se ha utilizado con mayor frecuencia CSM y CPM?	Analizar los sectores donde estas técnicas han sido aplicadas con mayor frecuencia, determinando los principales desafíos y beneficios que han aportado.

3.1.2 Elegibilidad y Selección de Estudios Relevantes

En la segunda etapa de la revisión, los hallazgos se organizan y presentan siguiendo la declaración PRISMA, utilizando la herramienta PRISMA Flow Diagram de Haddaway et al. [38] para visualizar el proceso de selección de estudios mostrado en la Figura 3.1. La

estrategia de búsqueda define los repositorios digitales donde se consultan estudios primarios sobre Minería de Conjuntos de Contraste (CSM) y Minería de Patrones de Contraste (CPM), incluyendo ScienceDirect (Elsevier), IEEE Xplore, ACM Digital Library, SpringerLink, MDPI, Wiley Online Library y Google Académico. La selección de fuentes se basa en el volumen de publicaciones relevantes, destacando ACM Digital Library (267), ScienceDirect (130) e IEEE Xplore (76) como los repositorios con mayor cantidad de literatura, en contraste con SpringerLink (28), MDPI (12) y Wiley Online Library (6). Adicionalmente, se incluye Google Scholar (773) para ampliar la búsqueda y capturar estudios pertinentes.

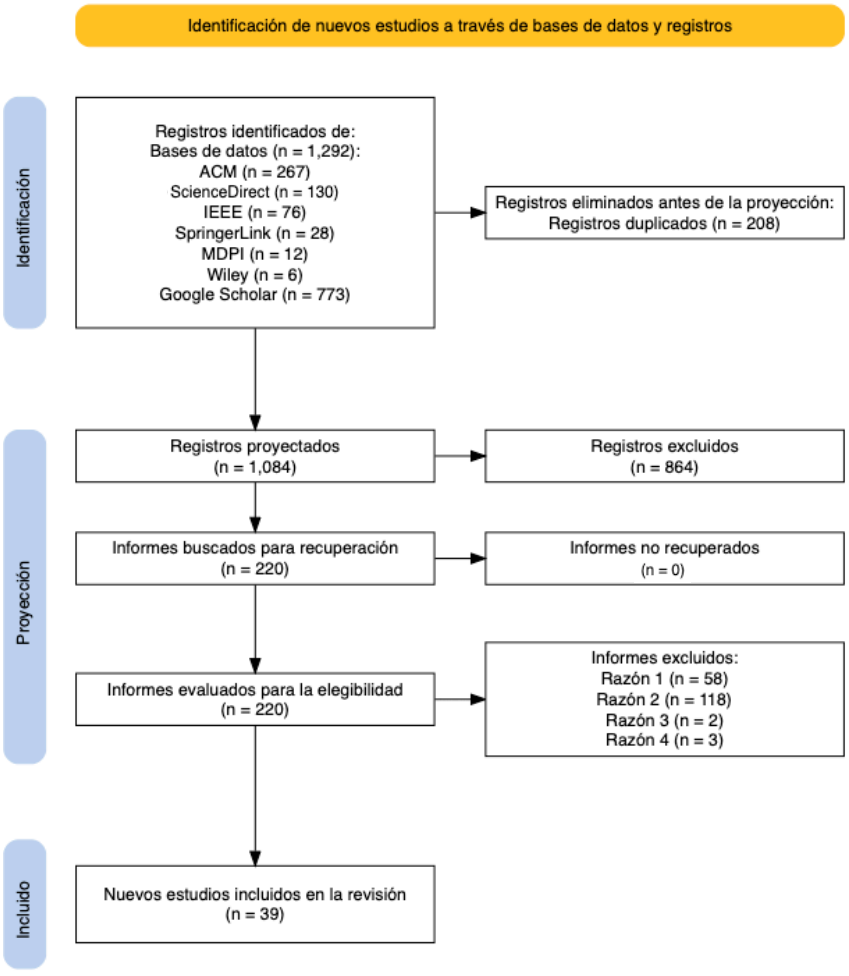


Figura 3.1 PRISMA Identificación de estudios

3.1.3 Criterios de Inclusión y Exclusión

En este trabajo se consideraron los siguientes criterios de inclusión:

1. Estudios publicados en revistas y conferencias indexadas entre 2018 y 2024.
2. Investigaciones que presentan aplicaciones de CSM y CPM en distintos dominios.
3. Trabajos que comparan distintos algoritmos de CSM y CPM.
4. Publicaciones disponibles en inglés, con acceso al texto completo.

Los criterios de exclusión son:

1. Estudios que no aborden específicamente CSM o CPM o que solo los mencionen sin aplicarlos.
2. Revisiones bibliográficas sin análisis empírico o una descripción metodológica clara.
3. Estudios con información incompleta o inaccesible.
4. Tesis de cualquier nivel y documentos técnicos que no hayan sido sometidos a revisión por pares.

3.1.4 Análisis Comparativo

Esta sección presenta los hallazgos obtenidos tras analizar los 39 artículos, organizados en la Tabla 3.2. Cada estudio se clasifica según diversas dimensiones representadas en cada columna: (1) Artículo, que cita el trabajo analizado; (2) CSM, que indica el uso de minería de conjuntos de contraste; (3) CPM, que señala la aplicación de minería de patrones de contraste; (4) Área de Aplicación, que describe los dominios donde se implementan estas técnicas; (5) Algoritmos, que especifica los métodos computacionales empleados; (6) Lenguajes, que menciona los lenguajes de programación utilizados en la implementación; y (7) Comparación de Opiniones, que identifica si los estudios incluyen análisis de encuestas o enfoques similares.

La estructura de esta clasificación facilita la identificación de tendencias, enfoques predominantes y vacíos en la literatura, ofreciendo una visión organizada sobre el estado del arte en minería de conjuntos y patrones de contraste.

Tabla 3.2 Artículos seleccionados

1	2	3	4	5	6	7
[39]	[ ]	[X]	Educación	HRFm (Hellinger Random Forest Miner)	Python	[ ]
[16]	[X]	[ ]	Salud	LCM (Linear Closed Itemset Mining)	Python, R	[ ]
[40]	[ ]	[X]	IA	Answer Set Programming (ASP)	Programming (ASP) Python	[ ]
[41]	[X]	[ ]	Minería de datos	STUCCO	Python	[ ]
[18]	[X]	[ ]	Agricultura	STUCCO	Python	[ ]
[42]	[X]	[ ]	Minería de datos	Con2Vec	Python	[ ]
[43]	[ ]	[X]	Ciencia de datos	Contrast dynamic feature dependency pattern mining for one paired multivariate time series (CMP), Group CMP (GCMP)	Python	[ ]
[44]	[ ]	[X]	Ciberseguridad	PBC4cip (Pattern-Based Classification for Class Imbalance Problems)	Python	[ ]
[17]	[X]	[ ]	Ing. Software	Continuous Contrast Set Mining (CCSM)	Python	[ ]
[45]	[X]	[ ]	Transporte	Deep Fuzzy Contrast-Set Mining	Python	[ ]
[46]	[ ]	[X]	Ing. Software	Contrast Motif Discovery	Python	[ ]
[47]	[ ]	[X]	Minería de datos	CHPMiner	Java	[ ]
[48]	[X]	[ ]	Ing. Software	STUCCO, RIPPER	Python, Scala	[ ]
[49]	[X]	[ ]	Salud	FP-Growth Algorithm	C++	[ ]
[50]	[ ]	[X]	Minería de datos	Frequent Subgraph Mining (FSM)	Python	[ ]
[51]	[X]	[ ]	Minería de datos	SDAD-CS (Supervised Dynamic and Adaptive Discretization for Contrast Sets)	Python	[ ]
[52]	[X]	[ ]	Salud	Frequent Pattern Mining (FPM)	Python	[ ]
[53]	[ ]	[X]	Ciberseguridad	EPClose	Python	[ ]
[54]	[X]	[ ]	Salud	STUCCO	Python	[ ]
[55]	[ ]	[X]	Minería de datos	TDCSP-tree (Time Distribution Contrast Sequential Pattern Tree)	Java	[ ]
[56]	[ ]	[X]	Ciberseguridad	RCEP (Rare recurring Change detection by Emerging Patterns) EPClose	Java	[ ]

Tabla 3.2 Artículos seleccionados - continuación

1	2	3	4	5	6	7
[58]	[ ]	[X]	Analítica web	Random Forest Miner	Python	[ ]
[59]	[ ]	[X]	Ciberseguridad	EPStream+	Java	[ ]
[60]	[X]	[ ]	Minería de datos	RuleKit2-CS	Java	[ ]
[61]	[X]	[ ]	Salud	CS-Miner	Python	[ ]
[62]	[X]	[ ]	Educación	STUCCO	Python	[ ]
[63]	[ ]	[X]	Salud	Random Forest	Python	[ ]
[22]	[ ]	[X]	*Marketing*	PBC4cip (Pattern-Based Classification for Class Imbalance Problems)	Java	[ ]
[64]	[X]	[ ]	Minería de datos	DiffNRDP-Miner	Python	[ ]
[65]	[X]	[ ]	Salud	Fuzzy contrast-based model	Python	[X]
[66]	[X]	[ ]	Semiconductores	STUCCO	Python	[ ]
[67]	[ ]	[X]	Ing. Software	CONAN	Python	[ ]
[68]	[ ]	[X]	Videojuegos	Modified SnippetFinder Algorithm	Python	[ ]
[69]	[ ]	[X]	Series de tiempo	COPP-Miner Algorithm	Python C++	[ ]
[70]	[ ]	[X]	Minería de datos	CSPm (Cost-Sensitive Pattern Miner)	Python	[ ]
[71]	[X]	[ ]	Minería de datos	Adaptive Discretization for frequent itemsets (SDAD-FI)	Otros	[ ]
[15]	[X]	[ ]	Salud	Fuzzy Contrast Set Mining	Python	[ ]
[72]	[X]	[ ]	Computación	Spatio-Temporal Event Detection Algorithm	Java	[ ]

El análisis de los artículos revisados muestra una amplia aplicación de la Minería de Conjuntos de Contraste (CSM, por sus siglas en inglés) y la Minería de Patrones de Contraste (CPM, por sus siglas en inglés) en distintos dominios. Los estudios que emplean CSM se enfocan principalmente en salud, minería de datos e ingeniería de software, mientras que CPM aparece con mayor frecuencia en ciberseguridad, analítica web y ciencia de datos. Entre los lenguajes de implementación, Python sobresale como el más utilizado, seguido por Java y C++.

Aunque estas técnicas demuestran su utilidad en diversas áreas, pocos estudios las aplican en la comparación de opiniones, lo que resalta la necesidad de enfoques

innovadores en el análisis de percepciones médicas y educativas. Detectar estos vacíos metodológicos abre oportunidades para investigar el uso de CSM en encuestas médicas, con el propósito de comprender mejor los factores que influyen en la reducción de autopsias y en la toma de decisiones hospitalarias fundamentadas en datos.

3.2 Selección

Como resultado del análisis exploratorio presentado en la sección anterior, donde se compararon diferentes enfoques para la identificación de diferencias significativas entre subconjuntos de datos específicamente, CSM y CPM, se concluyó avanzar sobre la línea de CSM. Esta decisión se fundamenta en la capacidad que ofrece CSM para generar reglas discriminativas fácilmente interpretables, que permiten entender con claridad las variables que distinguen a un grupo de otro dentro de contextos médicos, como es el caso de la disminución de autopsias en hospitales.

Una vez establecida esta línea metodológica, el siguiente paso consiste en comparar algoritmos representativos y actualizados dentro del enfoque de CSM, con el objetivo de seleccionar el más adecuado para su implementación en el módulo de análisis del sistema propuesto. La comparación se centrará en aspectos clave como la calidad de las reglas generadas, la interpretabilidad de los resultados, la eficiencia en el procesamiento y la adecuación al contexto específico del conjunto de datos.

En este marco, se seleccionaron tres algoritmos ampliamente reconocidos y utilizados para CSM. El primero es STUCCO (*Structured Understanding of Contrastive Concepts*), considerado un referente clásico en el área por su enfoque estructurado para identificar diferencias significativas entre subconjuntos de datos. El segundo es RuleKit2-CS, una extensión especializada del sistema RuleKit, diseñada para generar reglas contrastivas de alta calidad a partir de grandes volúmenes de datos categóricos. Finalmente, se incorpora ECSM (*Exceptional Contrast Set Mining*), un algoritmo que introduce mejoras en términos de eficiencia computacional y escalabilidad, mediante técnicas avanzadas de poda y optimización.

La evaluación de estos algoritmos se llevó a cabo utilizando el mismo conjunto de datos minables derivado de encuestas hospitalarias, y sus resultados se analizaron comparativamente para determinar cuál de ellos ofrece el mejor desempeño en este contexto. La elección final sirvió como base para la implementación del módulo analítico dentro del sistema.

3.2.1 STUCCO

El algoritmo STUCCO constituye una de las propuestas más reconocidas dentro del campo de CSM al enfocarse en la identificación de reglas contrastivas que reflejan diferencias estadísticas significativas entre grupos mutuamente excluyentes. STUCCO no se limita a encontrar asociaciones frecuentes, sino que busca específicamente patrones que presenten un comportamiento diferencial entre segmentos, permitiendo resaltar atributos cuya presencia o ausencia distingue claramente a un grupo con respecto a otros dentro del conjunto de datos.

Desde un enfoque técnico, STUCCO trabaja con datos categóricos y requiere como mínimo dos elementos fundamentales: un atributo de clase que defina los grupos de comparación, y un conjunto de variables predictoras (atributos) que alimenten el proceso de generación de reglas. El algoritmo identifica combinaciones de valores de atributos que aparecen con una frecuencia desproporcionada en un grupo determinado, lo que se traduce en una señal estadísticamente relevante de contraste.

STUCCO se aplicó sobre el conjunto de datos *vista_minable*, el cual contiene información estructurada obtenida de encuestas aplicadas a médicos en diferentes hospitales mexicanos como se observa en la Figura 3.2. El atributo utilizado como variable de agrupación fue *centro_hospitalario*, delimitando así los subgrupos analizados a partir de sus percepciones institucionales. Este enfoque permitió contrastar explícitamente las diferencias en la percepción sobre la disminución de autopsias entre unidades médicas seleccionadas, concretamente entre los hospitales etiquetados como H1 (Hospital Regional Río Blanco), H11 (Hospital General San Juan Bautista Tuxtepec, Oaxaca) y H12 (Hospital General de Zona 53, Estado de México).

La implementación seleccionada para este estudio corresponde a una versión optimizada en Python, que se encuentra disponible públicamente como software de código abierto. Esta versión permite configurar con precisión diversos parámetros que controlan la sensibilidad y especificidad de los análisis mostrados en la Tabla 3.3

Con base en la configuración de parámetros descrita, se diseñaron tres niveles de análisis: Exploratorio, Refinado y Específico, cada uno con una combinación distinta de valores para controlar el rigor estadístico y la complejidad de las reglas generadas. La Tabla 3.4 resume las configuraciones utilizadas y su propósito metodológico dentro del análisis contrastivo.

Proceso KDD

El proceso de descubrimiento de conocimiento en bases de datos guiado permite transformar los datos provenientes de encuestas médicas en patrones contrastivos útiles mediante técnicas formales de minería de datos. Este proceso siguió rigurosamente las etapas de selección, preprocesamiento, transformación, minería de datos y evaluación, teniendo como insumo principal el conjunto de datos denominado *vista_minable*.

- **Selección de los datos**

  La tabla *vista_minable* se generó a partir de un esquema hospitalario previamente modelado, que incluyó diversas entidades relacionadas con médicos, hospitales, atributos profesionales, y respuestas categóricas a preguntas cerradas, entre otros. El procedimiento para generar esta vista implicó combinar múltiples tablas almacenadas en una vista generada en PostgreSQL que da como resultado la tabla con sus atributos y valores, utilizando claves foráneas para establecer relaciones entre médicos, hospitales y sus respectivas respuestas.

Tabla 3.3 Descripción de parámetros STUCCO

Parámetro	Descripción	Recomendación
min_support	Frecuencia mínima que una regla debe tener dentro de un grupo, expresada como proporción (0-1).	0.01 - 0.1
min_support_count	Cantidad mínima absoluta de ocurrencias requeridas para considerar una regla.	≥ 10
min_difference	Diferencia mínima entre ocurrencias de una regla en un grupo y otros.	≥ 2
min_lift	Cociente entre el soporte observado de una regla en el grupo y el esperado por azar. Mayor indica mayor contraste.	≥ 3
min_confidence	Probabilidad de que una regla se cumpla en el grupo dado que los atributos están presentes.	0.75 - 0.95

Tabla 3.4 Configuraciones aplicadas por fase de análisis STUCCO

Tipo de Análisis	max_length	min_lift	min_support	min_difference	min_confidence	Objetivo
Exploratorio	1 – 2	1 – 2	0.01 – 0.03	1	0.6 – 0.7	Hallazgo inicial, revisión de variables
Intermedio / Refinado	2 – 3	3 – 5	0.03 – 0.07	2	0.75 – 0.85	Análisis más robusto y depurado
Específico / Final	3 – 4	≥ 5	≥ 0.07	≥ 3	≥ 0.85	Análisis final, conclusiones significativas

Para el presente caso de estudio, el análisis se aplicó exclusivamente a tres hospitales específicos: H1, H11 y H12. Estos centros hospitalarios se seleccionaron por su relevancia directa dentro del alcance de la investigación, ya que forman parte del subconjunto institucional sobre el cual se pretende comparar percepciones clínicas respecto a la disminución de autopsias. Además, el volumen y calidad de los registros correspondientes a estos hospitales aseguraron condiciones óptimas para la ejecución y validación de los algoritmos contrastivos.

- **Preprocesamiento de los datos**

  Durante esta etapa se aplicaron técnicas de limpieza estructurada sobre la vista consolidada. Se eliminaron registros con campos incompletos en atributos clave, se normalizaron las respuestas categóricas mediante codificación alfanumérica estandarizada (p. ej., a1, b2, entre otros), y se verificó que cada atributo presentara una distribución suficientemente informativa.

  Asimismo, se descartaron atributos con alta cardinalidad, alta redundancia o escasa presencia (bajo soporte), ya que no aportaban valor significativo al análisis contrastivo y podían introducir ruido o complejidad innecesaria. El resultado fue un subconjunto de atributos óptimamente seleccionados para la generación de reglas explicables y contrastivas.

- **Transformación de los datos**

  Una vez depurados, los datos se transformaron para cumplir con los requisitos estructurales de las herramientas de minería. Se exportó la vista *vista_minable* en formato *.csv*, codificada en UTF-8, asegurando que todas las columnas fueran de tipo categórico y que no existieran valores numéricos continuos, fechas ni campos vacíos.

  Se validó la unicidad de los valores por atributo, así como la consistencia del dominio de *centro_hospitalario*, conservando exclusivamente los registros correspondientes a los hospitales H1, H11 y H12 mediante filtros aplicados antes de la ejecución de los algoritmos. Esta transformación permitió garantizar

compatibilidad directa con las implementaciones de STUCCO, así como control sobre la integridad y distribución de los datos.

- **Minería de datos**

  En esta fase se aplicó el algoritmo STUCCO sobre el archivo *vista_minable.csv*, con el objetivo de identificar conjuntos de atributos que presentaran diferencias significativas entre los grupos hospitalarios definidos.

  Para STUCCO, se utilizó una implementación en Python que permitió configurar parámetros como *min_support, min_difference*, *min_lift*, *min_confidence* y *max_length*. Se ejecutaron tres configuraciones distintas: exploratoria, refinada y específica ajustando la profundidad del análisis en cada etapa. En todos los casos, se asignó *centro_hospitalario* como atributo de agrupación, permitiendo contrastar directamente los patrones encontrados entre H1, H11 y H12. Para la ejecución del algoritmo se utilizó una computadora portátil MacBook Pro de 16 pulgadas (modelo 2019), equipada con un procesador Intel Core i7 de seis núcleos a 2.6 GHz, una GPU AMD Radeon Pro 5300M con 4 GB de memoria dedicada, así como gráficos integrados Intel UHD Graphics 630 con 1536 MB. La memoria RAM instalada es de 32 GB DDR4 a 2667 MHz, y el sistema operativo utilizado fue macOS Sequoia versión 15.4.1

- **Evaluación e interpretación**

  STUCCO fue ejecutado bajo cada una de estas configuraciones sobre el conjunto de datos *vista_minable,* aplicando los filtros correspondientes para trabajar exclusivamente con los hospitales H1, H11 y H12. Los resultados reflejan cómo la variación en los parámetros afectó directamente tanto la cantidad como la calidad de las reglas encontradas:

  - **Fase Exploratoria**

    Se generaron 1,058 reglas, con un *lift* promedio de 3.41. Este nivel mostró asociaciones moderadas entre atributos y grupos hospitalarios, con una regla destacada que alcanzó un *lift* máximo de 33.07, aunque con bajo

soporte. El tiempo de ejecución fue de 47.24 segundos, lo cual sugiere que esta fase es adecuada para la detección inicial de relaciones relevantes.

- **Fase Refinada**

  Se detectaron 1,403 reglas, lo que representó un incremento respecto a la fase anterior, producto de una mayor profundidad combinatoria (hasta 3 condiciones por regla). El *lift* promedio se elevó a 6.39, evidenciando reglas más sólidas. El tiempo de ejecución aumentó a 363.25 segundos, lo que refleja el costo computacional asociado a configuraciones más exigentes.

- **Fase Específica**

  Esta etapa arrojó 324 reglas, enfocándose en las más significativas estadísticamente. El *lift* promedio alcanzó 7.94, con un máximo de 21.50. El tiempo de ejecución fue de 389.09 segundos. Aunque la cantidad de reglas se redujo, su valor explicativo y generalidad fueron notablemente superiores.

A continuación, se detallan los resultados generados por STUCCO tras su ejecución en las tres configuraciones propuestas. Cada fase refleja un nivel creciente de exigencia estadística y complejidad combinatoria. Los resultados consolidados se presentan en la Tabla 3.5.

Tabla 3.5 Resultados obtenidos con STUCCO por fase de análisis

Tipo de análisis	**Número de reglas**	**Lift promedio**	**Lift máximo**	**Lift mínimo**	**Tiempo de ejecución (s)**
Exploratorio	1,058	3.41	33.07	1.50	47.24
Intermedio / Refinado	1,403	6.39	12.43	4.00	363.25
Específico / Final	324	7.94	21.50	5.00	389.09

Los resultados confirman que, a mayor rigor en los parámetros, menor es el número de reglas, pero mayor es su calidad estadística, lo que valida el uso escalonado del análisis. La fase exploratoria cumplió su objetivo de proporcionar una visión general y guía para la calibración de parámetros. La fase refinada permitió detectar patrones con mayor

representatividad, y la fase específica reveló reglas altamente contrastantes, útiles para interpretación clínica e informes científicos.
Cabe destacar que el tiempo de ejecución aumentó exponencialmente en función del rigor de los parámetros. En conjunto, el análisis con STUCCO demostró ser eficaz para detectar diferencias significativas entre hospitales en cuanto a percepciones médicas sobre la disminución de autopsias, aportando evidencia clara y explicable desde el punto de vista técnico y clínico.

3.2.2 RuleKit2-CS

El algoritmo RuleKit2-CS constituye una de las herramientas más flexibles y eficientes dentro del campo de CSM, al enfocarse en la generación de reglas contrastivas mediante mecanismos de penalización, refinamiento iterativo y métricas múltiples de calidad. A diferencia de enfoques más directos, como STUCCO, RuleKit2-CS emplea un motor configurable que permite adaptar su comportamiento a diferentes niveles de profundidad analítica y control de complejidad, manteniendo la interpretabilidad de las reglas generadas [29, 30, 31, 32, 33].

Desde un enfoque técnico, RuleKit2-CS trabaja con datos categóricos y requiere como mínimo una columna que represente los grupos de comparación, junto con un conjunto de atributos que permitan inferir patrones discriminativos. El algoritmo construye reglas en múltiples iteraciones sobre los datos, filtrando aquellas con soporte bajo, alta proporción de falsos positivos o redundancia entre condiciones. La arquitectura modular de RuleKit2-CS permite definir criterios precisos para la inducción, poda y votación de reglas, utilizando medidas como *Lift, WRAcc* o *Correlation.*

RuleKit2-CS se aplicó sobre el conjunto de datos *vista_minable*, el cual contiene información estructurada obtenida de encuestas aplicadas a médicos en distintos hospitales mexicanos. El atributo utilizado como variable de agrupación fue *centro_hospitalario*, delimitando así los subgrupos analizados a partir de sus percepciones institucionales. Este enfoque permitió contrastar explícitamente las

diferencias en la percepción sobre la disminución de autopsias entre unidades médicas seleccionadas, concretamente entre los hospitales etiquetados como H1, H11 y H12.

La implementación utilizada para este estudio corresponde a la versión RuleKit2-CS en Java, distribuida como parte del entorno RuleKit desarrollado por ADaA-POLSL. Esta versión permite controlar una serie de parámetros clave que influyen directamente en la generación, evaluación y selección de reglas, como se muestra en la Tabla 3.6.
Con base en la configuración de estos parámetros, se definieron tres niveles de análisis: Exploratorio, Refinado y Específico. Cada fase fue diseñada para ajustar progresivamente la profundidad y rigor de las reglas generadas, tal como se resume en la Tabla 3.7.

Proceso KDD

El proceso de descubrimiento de conocimiento en bases de datos aplicado en este algoritmo permitió transformar los datos provenientes de encuestas médicas en patrones contrastivos útiles mediante técnicas formales de minería de datos. El flujo metodológico siguió exactamente el mismo aplicado para STUCCO definido en la sección anterior dentro del apartado 3.2.1. dejando solo como diferencia los siguientes puntos:

- **Minería de datos con RuleKit2-CS**
  La etapa de minería de datos se llevó a cabo utilizando RuleKit2-CS, una implementación del enfoque de CSM desarrollada en Java, la cual opera mediante configuraciones definidas en archivos XML. Esta herramienta permite especificar parámetros avanzados para controlar el proceso de inducción, poda, penalización y evaluación de reglas, facilitando su ajuste a distintos niveles de rigor analítico.
  Para este algoritmo se definieron tres configuraciones analíticas: exploratoria, refinada y específica, cada una con diferentes umbrales y estrategias de control.

Tabla 3.6 Descripción de parámetros Rulekit2-CS

Parámetro	Descripción	Recomendación
minsupp_all	Soporte mínimo que una regla debe tener dentro del grupo objetivo	0.01 – 0.10
max_neg2pos	Proporción máxima entre falsos positivos y verdaderos positivos	≤ 0.5
max_passes_count	Número máximo de iteraciones para construir reglas complejas	1 – 5
penalty_strength	Penalización por uso de atributos poco informativos	0.0 en exploratorio, ≥ 0.5 en refinado
penalty_saturation	Activación de penalización ante repetición de atributos redundantes	1.0 general, 0.8 en específico
induction_measure	Métrica de calidad usada para construir reglas	*Lift*
pruning_measure	Métrica para eliminar reglas poco relevantes	*Lift o WRAcc*
voting_measure	Métrica para determinar la calidad final de una regla	*Lift*

Tabla 3.7 Configuraciones aplicadas por fase de análisis Rulekit2-CS

Tipo de Análisis	max_passes_count	minsupp_all	penalty_strength	penalty_saturation	Objetivo
Exploratorio	1	0.02 – 0.03	0.0	1.0	Hallazgo inicial, revisión general
Intermedio / Refinado	2 – 3	0.05 – 0.08	0.5	1.0	Análisis más robusto y ajustado
Específico / Final	3 – 5	≥ 0.09	≥ 0.8	0.8	Análisis enfocado, conclusiones significativas

Los parámetros empleados fueron los siguientes:

- *minsupp_all*
- *minsupp_new*
- *max_neg2pos*
- *max_passes_count*
- *penalty_strength*
- *penalty_saturation*
- *induction_measure, pruning_measure y voting_measure*

En las tres configuraciones, se utilizó el atributo *centro_hospitalario* como variable de agrupación, permitiendo generar reglas contrastivas que identificaran diferencias estructurales entre las percepciones médicas de los hospitales H1, H11 y H12. La correcta configuración de estos parámetros fue clave para obtener reglas significativas desde un punto de vista clínico, manteniendo un equilibrio entre calidad estadística, interpretabilidad y eficiencia computacional.

- **Evaluación e interpretación**

  Los resultados obtenidos con RuleKit2-CS se agruparon en tres fases:

  - **Fase Exploratoria**

    Se obtuvieron 32 reglas con un *lift* promedio de 4.92 y un *lift* máximo de 8.33. Estas reglas reflejan una asociación clara entre condiciones de atributos y hospital de pertenencia, con tiempo de ejecución muy reducido (0.38 segundos), ideal para pruebas iniciales o análisis iterativo.

  - **Fase Refinada**

    Se generaron 34 reglas, con un *lift* promedio de 4.06. Aunque ligeramente menor que en la fase exploratoria, las reglas presentan mayor estabilidad estadística. El tiempo de ejecución fue de 0.78 segundos, manteniéndose bajo incluso con parámetros más exigentes.

  - **Fase Específica / Final**

    Se identificaron 49 reglas con un *lift* promedio de 4.84. Esta fase logró el mayor número de reglas, todas generadas bajo criterios estrictos. El

soporte promedio más alto de todas las fases (0.391) sugiere una alta representatividad, y el tiempo de ejecución fue de solo 1.36 segundos.

A continuación, se presentan los resultados obtenidos tras la ejecución de RuleKit2-CS en las tres fases de análisis configuradas: exploratoria, refinada y específica. Cada configuración varió en sus parámetros de soporte, penalización y número de iteraciones, lo que permitió observar el impacto del rigor analítico sobre la cantidad, calidad y eficiencia de las reglas generadas. La Tabla 3.8 resume el desempeño del algoritmo en términos del número de reglas contrastivas obtenidas, el valor promedio y extremo del *lift*, así como el tiempo de ejecución requerido en cada fase.

Tabla 3.8 Resultados obtenidos con Rulekit2-CS por fase de análisis

Tipo de análisis	N.º de reglas	Lift promedio	Lift máx.	Lift mín.	Tiempo (s)
Exploratorio	32	4.92	8.33	1.53	0.38
Refinado	34	4.06	8.33	1.45	0.78
Específico / Final	49	4.84	8.33	1.32	1.36

El algoritmo RuleKit2-CS demostró ser altamente eficiente para la identificación de reglas contrastivas. En todas las fases, el tiempo de ejecución fue mínimo en comparación con alternativas más complejas, y las reglas generadas mostraron un equilibrio entre significancia estadística y claridad interpretativa.

La estabilidad del *lift*, junto con el aumento progresivo del soporte promedio, sugiere que RuleKit2-CS favorece la generación de reglas representativas incluso bajo configuraciones exigentes. Su arquitectura basada en penalizaciones y evaluación iterativa lo hace especialmente útil para módulos que requieran respuesta rápida, escalabilidad y resultados interpretables.

3.2.3 ECSM

El algoritmo ECSM (*Exceptional Contrast Set Mining,* Minería de conjuntos de contraste excepcional) forma parte de un conjunto limitado de enfoques orientados a la identificación de reglas contrastivas *excepcionales*, es decir, aquellas que no solo

distinguen grupos entre sí, sino que además resaltan patrones que se desvían significativamente del comportamiento general esperado. Este algoritmo fue desarrollado como una propuesta académica centrada en la generación de reglas altamente específicas, con el objetivo de facilitar explicaciones diferenciadas sobre subconjuntos de interés [74].

Desde el punto de vista técnico, ECSM trabaja exclusivamente con atributos categóricos y requiere un conjunto de datos etiquetado con clases mutuamente excluyentes. A través de un procedimiento de evaluación basado en métricas como Chi-squared y Soporte. El algoritmo fue originalmente implementado como una aplicación de escritorio sobre el entorno .NET Framework, con un enfoque orientado a la visualización interactiva de reglas.

Entre sus características principales se incluyen:

- Enfoque en la identificación de contrastes estadísticamente significativos que representen comportamientos fuera de lo común en un subconjunto.
- Generación de reglas interpretables, adecuadas para informes técnicos o decisiones operativas.
- Aplicación orientada a análisis discretos, no integrados en flujos automáticos de minería de datos.

Si bien ECSM resultó conceptualmente atractivo para este proyecto, en especial por su enfoque en reglas contrastivas altamente específicas, durante la fase de evaluación técnica se identificaron múltiples limitaciones que imposibilitaron su aplicación práctica dentro del flujo de análisis planteado.

Entre los principales obstáculos se encuentran:

- **Incompatibilidad tecnológica:** ECSM fue implementado como una aplicación de escritorio sobre .NET Framework 4.x, lo que restringe su ejecución a sistemas Windows específicos. No existe versión disponible como biblioteca programable ni entorno de ejecución independiente que pueda integrarse en plataformas modernas de análisis como Python, Java o entornos basados en PostgreSQL.
- **Ausencia de soporte programático o por línea de comandos:** A diferencia de STUCCO y RuleKit2-CS, ECSM no permite la ejecución automatizada ni admite

la carga dinámica de parámetros a través de archivos de configuración. Esta falta de interfaz programática limita su uso a sesiones interactivas, incompatibles con procesos de minería reproducibles o escalables.

- **Formato de entrada inflexible:** El algoritmo exige estructuras de entrada codificadas manualmente, sin soporte directo para archivos estándar como .csv. Esta rigidez complica la incorporación del conjunto *vista_minable*, el cual contiene más de cien atributos categóricos provenientes de múltiples fuentes interrelacionadas.
- **Ausencia de documentación y mantenimiento activo:** No se identificó documentación técnica actualizada ni evidencia de mantenimiento del proyecto. Tampoco se encontró una comunidad activa de usuarios ni un repositorio de código fuente confiable, lo que impidió modificar, compilar o ajustar el comportamiento del algoritmo ante errores de ejecución.

A pesar de que ECSM presenta una propuesta metodológica válida y valiosa desde el punto de vista teórico, su implementación técnica resultó obsoleta e incompatible con el ecosistema actual de análisis de datos. La imposibilidad de integrarlo en el flujo automatizado del proceso KDD, su rigidez estructural, y la falta de soporte funcional hicieron inviable su aplicación en este estudio.
Por lo tanto, se decidió no aplicar ECSM dentro de las pruebas de minería de conjuntos de contraste, enfocando los esfuerzos en algoritmos operativos y reproducibles como STUCCO y RuleKit2-CS, los cuales no solo ofrecen resultados contrastivos de alta calidad, sino que además cumplen con los estándares de integración técnica, escalabilidad y trazabilidad requeridos por el presente proyecto.

Esta decisión no desestima el valor metodológico de ECSM, sino que responde estrictamente a criterios de viabilidad tecnológica y sostenibilidad dentro del entorno de análisis diseñado para este sistema e integración del módulo de CSM.

3.2.4 Comparación de Algoritmos

Con el objetivo de establecer una comparación rigurosa entre los algoritmos STUCCO y RuleKit2-CS, se diseñó un marco metodológico que permitiera evaluar ambos enfoques

bajo condiciones lo más homogéneas posibles. Esta comparación se realizó sobre el mismo conjunto de datos (*vista_minable*), utilizando el atributo *centro_hospitalario* como variable de agrupación, y replicando tres fases analíticas: exploratoria**,** refinada y específica**.**

Dado que cada algoritmo emplea parámetros propios, fue necesario aplicar un proceso de homologación de criterios de calidad y normalización de métricas para permitir una comparación justa y significativa. Esto incluyó tanto la alineación conceptual de los parámetros de configuración como la aplicación de fórmulas compartidas para el cálculo de métricas como *lift*, soporte y tiempo de ejecución.

Homologación de Parámetros

Ambos algoritmos utilizan distintos esquemas para controlar la generación y evaluación de reglas. La Tabla 3.9 muestra el mapeo entre parámetros equivalentes o funcionalmente similares, permitiendo establecer un punto de partida común para el análisis comparativo:

Tabla 3.9 Homologación de parámetros STUCCO Vs Rulekit2-CS

Parámetro en STUCCO	**Parámetro en RuleKit2-CS**	**Equivalencia / Comentario**
min_support	minsupp_all, minsupp_new	Controlan el soporte mínimo como proporción
min_support_count	—	RuleKit no usa soporte absoluto, solo proporción
min_difference	— (implícito)	Reflejado mediante delta_support y métricas como Lift o Chi²
min_lift	induction_measure = Lift	Coinciden funcionalmente al usarse Lift como medida de inducción
min_confidence	—	No configurable en RuleKit; WRAcc puede ser una aproximación indirecta.
max_length	— (indirecto)	En RuleKit se regula por max_passes_count, penalty_strength, entre otros.
—	max_neg2pos	Control de falsos positivos relativo; no existe equivalente en STUCCO

Para garantizar una comparación justa y rigurosa entre STUCCO y RuleKit2-CS, se definieron y aplicaron de manera uniforme fórmulas compartidas para evaluar la calidad de las reglas generadas. Si bien cada algoritmo maneja internamente distintas métricas, la comparación se concentró en dos medidas fundamentales: *lift* y soporte relativo.
Cálculo del *Lift* (usado para ambos algoritmos):

$$\text{Lift} = \frac{\text{precisión}}{\text{proporción de la clase}} = \frac{p/(p+n)}{P/(P+N)}$$

donde:

- *p*: número de registros del grupo objetivo cubiertos por la regla (verdaderos positivos)
- *n*: número de registros de otros grupos cubiertos por la regla (falsos positivos)
- *P*: total de registros del grupo objetivo
- *N*: total de registros del resto de los grupos

En el caso de RuleKit2-CS, esta métrica no forma parte de las salidas estándar del algoritmo. Por ello, el valor de *lift* fue calculado de manera externa, a partir de las estadísticas generadas por la ejecución (support, neg_support, delta_support, entre otras). Esta medida fue esencial para poder comparar directamente con STUCCO, que sí incluye *lift* como parte integral de su evaluación interna.

Cálculo del Soporte Relativo (usado particularmente en STUCCO):

$$\text{Soporte (support)} = \frac{\text{Número de registros del grupo objetivo que cumplen la regla}}{\text{Total de registros del grupo objetivo}}$$

Esta medida fue empleada en STUCCO para aplicar filtros mediante *min_support* y *min_support_count*, descartando reglas con baja frecuencia relativa dentro del grupo.
En contraste, RuleKit2-CS no trabaja con soporte absoluto (*min_support_count*), sino con proporciones (*minsupp_all*, *minsupp_new*) y otras métricas internas como

delta_support, *neg2pos_support*, *precision, redundancy*, entre otras. Sin embargo, dichas métricas no fueron consideradas para la comparación final, dado que no están disponibles o no tienen equivalentes funcionales en STUCCO. De este modo, se evitó introducir sesgos al análisis.

La homologación metodológica mediante estas fórmulas permitió evaluar con precisión el grado de contraste (*lift*) y la representatividad interna (soporte) de las reglas obtenidas, proporcionando así una base sólida y justa para la interpretación comparativa entre ambos algoritmos.

Comparación de resultados

La ejecución de ambos algoritmos bajo configuraciones alineadas permitió identificar diferencias claras en comportamiento, rendimiento y aplicabilidad. La Tabla 3.10 resume los principales resultados cuantitativos por fase:

Tabla 3.10 Resultados comparativos – STUCCO Vs Rulekit2-CS

Tipo de análisis	Algoritmo	Nº de reglas	Lift promedio	Lift máximo	Lift mínimo	Soporte mínimo	Soporte promedio	Tiempo ejecución (s)
Exploratorio	STUCCO	**1,058**	3.41	**33.07**	1.50	0.0047	0.1177	47.24
	RuleKit2-CS	32	**4.92**	8.33	**1.53**	**0.0250**	**0.1640**	**0.38**
Refinado	STUCCO	**1,403**	**6.39**	**12.43**	**4.00**	0.0026	0.0461	363.25
	RuleKit2-CS	34	4.06	8.33	1.45	**0.0900**	**0.3650**	**0.78**
Específico / Final	STUCCO	**324**	**7.94**	**21.50**	**5.00**	0.0035	0.0274	389.09
	RuleKit2-CS	49	4.84	8.33	1.32	**0.1010**	**0.3910**	**1.36**

La comparación técnica y empírica entre los algoritmos STUCCO y RuleKit2-CS permite establecer una clara distinción en términos de rendimiento, eficiencia computacional, claridad interpretativa y compatibilidad con el entorno tecnológico del presente proyecto. Aunque ambos algoritmos son válidos para la minería de conjuntos de contraste, RuleKit2-CS sobresale como la opción más robusta y estratégica para ser integrada en el módulo de análisis comparativo del sistema actual.

Entre los aspectos más destacados se encuentran:

- **Compatibilidad total con el sistema existente:** RuleKit2-CS está desarrollado en Java, lo que permite su integración directa con la arquitectura tecnológica ya empleada en el sistema, sin requerir adaptaciones adicionales ni entornos externos. En contraste, STUCCO exige un entorno Python separado y no nativamente interoperable con el sistema principal.
- **Eficiencia computacional superior:** En todas las fases evaluadas, RuleKit2-CS completó sus análisis en menos de 2 segundos, mientras que STUCCO superó los 360 segundos en configuraciones avanzadas. Esta diferencia convierte a RuleKit2-CS en una alternativa ideal para sistemas en tiempo real o análisis iterativo con recursos limitados.
- **Cobertura y representatividad de reglas:** Las reglas generadas por RuleKit2-CS mantuvieron un soporte promedio más alto en todas las fases, lo que indica una mayor representatividad y aplicabilidad práctica. Esto favorece su utilidad en entornos clínicos donde se prioriza la comprensión y confiabilidad de los patrones extraídos.
- **Claridad y menor redundancia:** Las reglas de RuleKit2-CS son más limpias, menos redundantes y más comprensibles, facilitando su presentación ante usuarios finales como personal médico o autoridades administrativas.
- **Facilidad de configuración y control de precisión:** RuleKit2-CS permite controlar directamente la proporción de falsos positivos mediante el parámetro *max_neg2pos*, además de contar con mecanismos de penalización por atributos poco informativos. Estas capacidades no están disponibles en STUCCO.
- **Capacidad de integración modular:** Gracias a su diseño basado en archivos XML y arquitectura modular, RuleKit2-CS puede acoplarse como componente analítico dentro del sistema general sin reestructuración, respetando principios de reutilización, escalabilidad y mantenibilidad del software.

Por todo lo anterior, RuleKit2-CS se perfila como la opción más completa, eficiente y estratégicamente alineada con los objetivos y requisitos técnicos del proyecto, ofreciendo

un equilibrio óptimo entre calidad de reglas, rendimiento estadístico y viabilidad de implementación.

3.3 Diseño de la aplicación

En esta fase, se define la estructura técnica y visual del sistema propuesto, con el propósito de garantizar su funcionamiento eficiente, modular y escalable. El diseño constituye un puente entre el análisis y la implementación, al traducir los requerimientos funcionales y no funcionales en componentes arquitectónicos, modelos de datos y elementos de interfaz gráfica.

El desarrollo del diseño considera una arquitectura en capas y basada en el patrón Modelo-Vista-Controlador (MVC), lo que permite separar la lógica de negocio, la gestión de datos y la presentación visual, facilitando el mantenimiento y la evolución del sistema. Asimismo, se presenta el modelo de base de datos, en el cual se estructura la información necesaria para el almacenamiento, consulta y procesamiento de los datos utilizados en la minería de conjuntos de contraste. Este modelo garantiza la integridad, consistencia y eficiencia en las operaciones realizadas por el módulo propuesto.

Finalmente, se describe el diseño de la interfaz gráfica de usuario (GUI), centrado en ofrecer una experiencia interactiva, clara y accesible. Su propósito es permitir que los usuarios realicen tareas como la carga de datos, la ejecución de algoritmos, la visualización de resultados y la interpretación de reglas, mediante una representación visual intuitiva y alineada con la funcionalidad del sistema.

3.3.1 Arquitectura del módulo

La arquitectura para el módulo de minería de conjuntos de contraste se basa en una integración modular entre el patrón MVC y el entorno analítico RuleKit2-CS, lo que permite una comunicación eficiente entre la interfaz de usuario, la lógica de negocio y los procesos de minería de datos. Esta estructura favorece la escalabilidad, la mantenibilidad

y la separación de responsabilidades, asegurando una organización clara de los componentes funcionales del sistema.

La Figura 3.2 muestra la arquitectura general del módulo, donde se aprecia la interacción entre las capas del modelo MVC y la integración con el motor analítico RuleKit2-CS, utilizado para la ejecución de los algoritmos de minería de conjuntos de contraste y la generación de resultados comparativos.

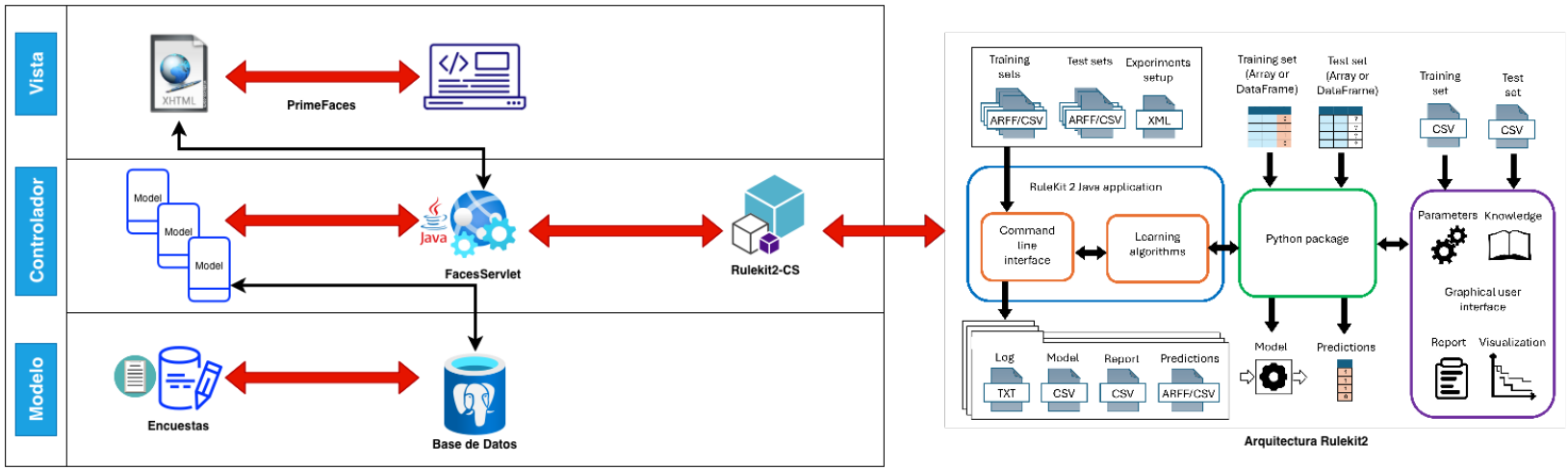


Figura 3.2 Arquitectura del módulo de minería de conjuntos de contraste

En el lado izquierdo del diagrama se observa la estructura MVC:

- **Modelo**: encapsula la lógica de acceso a los datos y la estructura de información utilizada en el proceso de minería. En esta capa se gestionan las encuestas y las conexiones con la base de datos de PostgreSQL, donde se almacenan las respuestas, atributos y resultados intermedios derivados de los análisis.
- **Controlador**: implementada en Java, se encarga de orquestar la comunicación entre la vista y el modelo mediante el uso de FacesServlet. Aquí se manejan las peticiones del usuario, la invocación de los procesos de análisis y la preparación de los datos para ser enviados al módulo RuleKit2-CS.
- **Vista**: desarrollada en XHTML utilizando PrimeFaces, proporciona la interfaz gráfica que permite la interacción con el usuario. A través de esta capa se realizan las operaciones de carga de datos, ejecución de algoritmos y visualización de resultados mediante componentes interactivos.

En la parte derecha se detalla la arquitectura interna de RuleKit2-CS, compuesta por tres secciones principales: la aplicación Java, el paquete Python y la interfaz gráfica de resultados.

- **Aplicación Java RuleKit2**: administra la configuración de los experimentos, la lectura de los conjuntos de entrenamiento y prueba en formato ARFF/CSV y la ejecución de los algoritmos de aprendizaje.
- **Paquete Python**: ofrece un entorno alternativo para la ejecución de experimentos desde un enfoque más analítico, permitiendo manipular los conjuntos de datos en estructuras tipo DataFrame y generar modelos predictivos.
- **Interfaz gráfica de RuleKit2**: facilita la configuración de parámetros, la carga de datos y la visualización de resultados mediante reportes, gráficas y resúmenes de conocimiento, sin embargo, esta integración es externa a la aplicación utilizada en el módulo de minería de conjuntos de contraste.

La comunicación entre el sistema principal y RuleKit2-CS se realiza mediante el intercambio de archivos en formato ARFF y CSV, garantizando interoperabilidad y consistencia en el flujo de información. Este diseño híbrido asegura la continuidad desde

la captura de encuestas hasta la interpretación de las reglas contrastantes, integrando la usabilidad del entorno web con el poder analítico del motor RuleKit2-CS.

3.3.2 Base de Datos

El diseño de la base de datos constituye un componente esencial dentro del módulo de minería de conjuntos de contraste, ya que garantiza la correcta estructuración, integridad y disponibilidad de los datos utilizados durante el análisis comparativo de encuestas.
Para ello, se realizó una representación progresiva del modelo en tres niveles: conceptual, lógico y físico, los cuales permiten pasar de una visión abstracta de las entidades del dominio a una implementación precisa dentro del sistema gestor de bases de datos PostgreSQL.

El modelo conceptual representado en la Figura 3.3 representa la visión general del sistema de información y sus relaciones principales.
En este modelo se identifican las entidades fundamentales del dominio, tales como *Encuesta, Pregunta, Respuesta, Medico, CentroHospitalario, Área, Institución, Ciudad* y *Pais*, así como las asociaciones que permiten vincularlas entre sí.
El propósito de este modelo es describir de forma abstracta la organización de los datos y las dependencias lógicas entre ellos, estableciendo la base semántica para la estructura posterior de la base de datos.

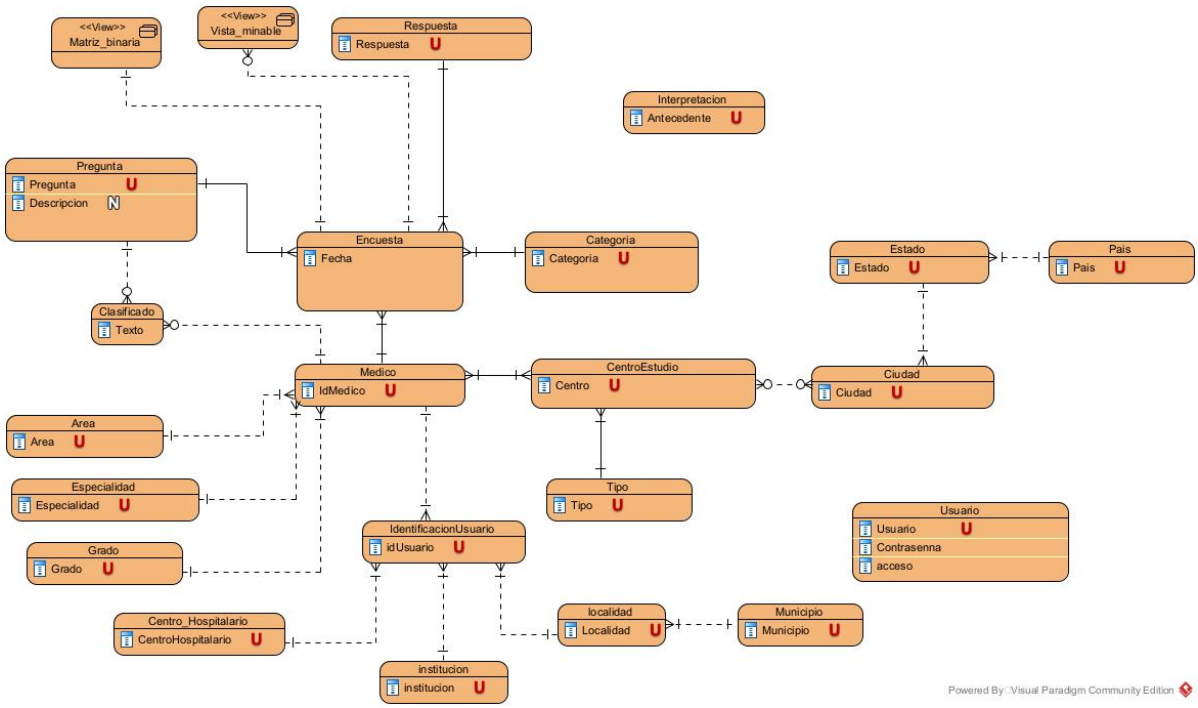


Figura 3.3 Diagrama conceptual de la base de datos

El modelo lógico representado en la Figura 3.4 traduce la estructura conceptual en un esquema compatible con el enfoque relacional, incorporando claves primarias, claves foráneas y reglas de integridad referencial.

En este nivel se definen las tablas derivadas de las entidades conceptuales, las relaciones de cardinalidad y los atributos que describen los objetos de estudio.

Por ejemplo, las tablas: *pregunta*, *respuesta* y *encuesta* mantienen una relación directa, mientras que *centrohospitalario* y *medico* se vinculan a través de claves foráneas que permiten identificar la procedencia de cada respuesta.

Este modelo garantiza la consistencia de los datos y sienta las bases para la normalización y optimización del rendimiento del sistema.

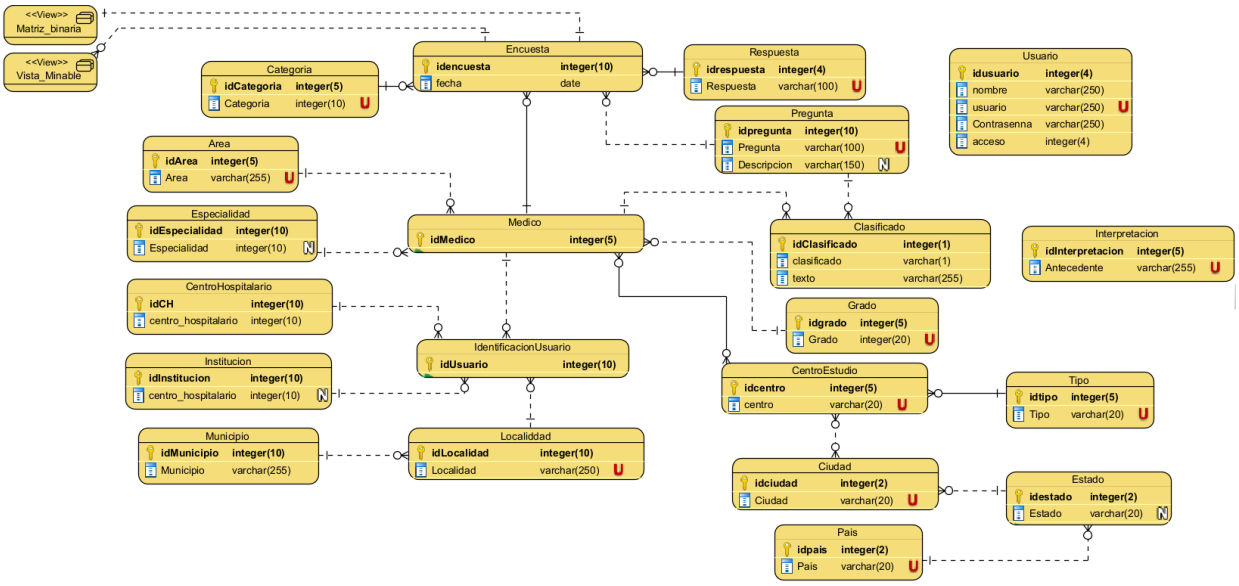


Figura 3.4 Diagrama lógico de la base de datos

El modelo físico representando en la Figura 3.5 muestra la implementación concreta del diseño dentro del sistema gestor PostgreSQL, empleando tipos de datos, índices y restricciones propias del motor de base de datos.

Cada entidad se materializa en una tabla relacional con sus respectivas columnas, tipos de datos, claves primarias, foráneas y restricciones de integridad.

Entre los elementos más relevantes se incluyen las vistas *vista_minable* y *matriz_binaria*, utilizadas como fuentes de datos para la ejecución de los algoritmos de minería de conjuntos de contraste, así como las tablas *interpretacion*, *clasificado* y *entrenamiento_texto*, que complementan el proceso de análisis y generación de interpretaciones en lenguaje natural.

Este modelo constituye la estructura final empleada por el módulo, asegurando la trazabilidad de la información desde la captura de encuestas hasta la fase analítica.

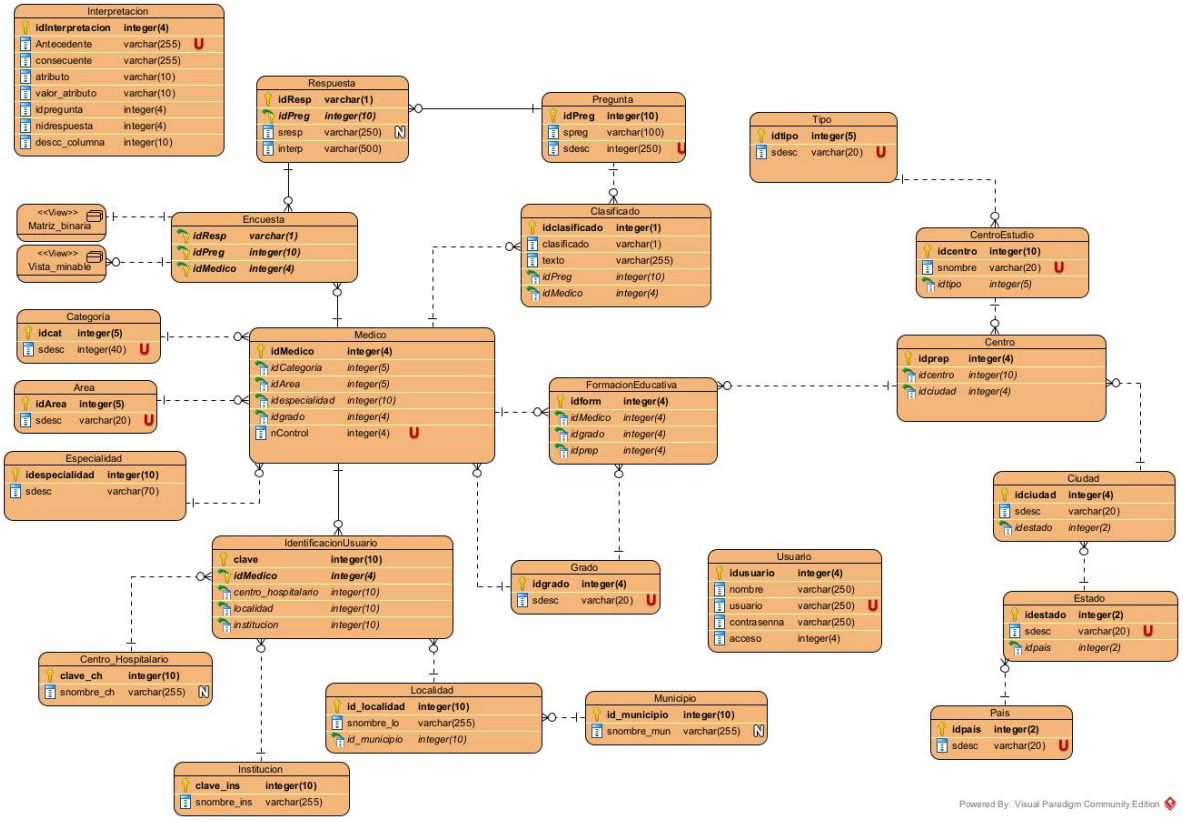


Figura 3.5 Diagrama físico de la base de datos

Estructura de Tablas:

- **Categoría**: Guarda toda la información relacionada con la manera como el encuestado aceptó contestar la encuesta, ya sea a través de una "Invitación interna" o una "Invitación accionada fuera del servicio".
- **Especialidad**: Almacena el valor que indicará a qué campo corresponde un encuestado, es decir, la referencia al sector médico será "Medicina".
- **Área**: Almacena la información que indica qué servicio presta el encuestado al hospital, es decir, si es "Adscrito", "Residente" o "Interno".
- **Grado**: Se refiere al valor que se utilizará para indicar el grado de especialización. del encuestado dentro del área médica, dichos valores son "Práctica Universitaria", "Medicina General" o "Especialidad".
- **Centro Estudio**: Almacena los nombres de los centros de estudio relacionados con la enseñanza de medicina. Por ejemplo; "Facultad de Medicina", "Universidad Veracruzana" y "Hospital Regional Río Blanco", por mencionar algunos.

- **Tipo**: Esta tabla contiene la catalogación de las instituciones de preparación médica, ya sea "Escuela" u "Hospital".
- **Ciudad, Estado, País**: Estas tablas y sus relaciones posibilitan una localización geográfica de los centros educativos, p. ej., ya que la Universidad Veracruzana cuenta con campus en distintas zonas del país será posible conocer exactamente a qué campus se hace referencia de la "Universidad Veracruzana" si se refiere a la del estado "Veracruz" en la ciudad de "Mendoza", o del estado de "Veracruz" en la ciudad "Xalapa".
- **Pregunta**: En esta tabla se almacena cada una de las preguntas que forman parte de la encuesta.
- **Respuesta**: Guarda todas las posibles respuestas en la encuesta y fija la relación de estas respuestas con la pregunta que corresponde. Por ejemplo, en el caso de respuestas como "Menos de 5" y "Más de 20", que son opciones posibles para la pregunta "Años de práctica". Esta relación se establece a través de la clave externa "idPreg" en la tabla “Respuesta”, que hace referencia a la clave primaria "idPreg" en la tabla “Pregunta”.
- **Centro**: La ubicación geográfica de cada centro de preparación médica se asigna en función de las relaciones entre las tablas "Centro Estudio" y "Ciudad".
- **Médico**: En esta tabla se guardan las encuestas, estableciendo la relación entre el número de control de cada encuesta y su categoría, especialidad, grado y área correspondiente.
- **Formación Educativa**: Almacena las relaciones con las tablas “Médico”, “Grado” y “Centro”; es así como se registra la información de los centros médicos donde se formó el médico encuestado.
- **Encuesta**: Las respuestas para cada pregunta de cada médico encuestado se registran en esta tabla, basándose en las relaciones entre las tablas "Médico", "Pregunta" y "Respuesta".
- **Clasificado**: Almacena las respuestas de las preguntas abiertas en conjunto con su correspondiente clasificación, el cual es resultado del proceso de minería de texto para la clasificación.

- **Interpretación**: Con el propósito de proporcionar una descripción lógica de los resultados obtenidos a partir de los algoritmos de minería empleados en este estudio, en esta tabla se guarda el significado y la descripción de cada atributo presente en los conjuntos de datos que se han analizado
- **Matriz_binaria** y **Vista_minable**: Contienen la información de las encuestas que han sido guardadas en la base de datos relacionadas con el centro hospitalario al que pertenecen y se crean por funciones SQL automáticas.
- **IdentificacionUsuario**: Aquí se contiene toda la información del usuario, incluyendo el centro hospitalario, localidad, institución y médico. Debido a que está asociada con varias tablas, esta tabla permite identificar el centro hospitalario desde el cual se contestó la encuesta.
- **Municipio**: Guarda los nombres de los municipios.
- **Localidad**: En esta tabla se almacenan los nombres de las localidades y se establece a qué municipio pertenecen.
- **CentroHospitalario**: Contiene el nombre de todos los centros hospitalarios existentes en la base de datos.
- **Institución**: Almacena el nombre de todas las instituciones.

La representación progresiva de los modelos conceptual, lógico y físico permite garantizar la coherencia estructural y semántica del sistema. Este enfoque favorece la interoperabilidad entre los módulos del sistema, la eficiencia en las consultas y la posibilidad de escalar el modelo hacia otros tipos de encuestas o dominios analíticos.
De esta manera, la base de datos constituye la columna vertebral del módulo de minería de conjuntos de contraste, proporcionando soporte robusto y estructurado para los procesos de descubrimiento de conocimiento y comparación de resultados hospitalarios.

3.3.2 Casos de uso

La Figura 3.6 muestra el diagrama de casos de uso general del sistema desarrollado, el cual integra las funcionalidades disponibles para los tres actores principales: Administrador, Especialista y Encuestado. Cada actor interactúa con los distintos

módulos del sistema de acuerdo con sus privilegios y responsabilidades, manteniendo la trazabilidad completa desde la gestión de usuarios hasta la ejecución de procesos analíticos y el llenado de encuestas.

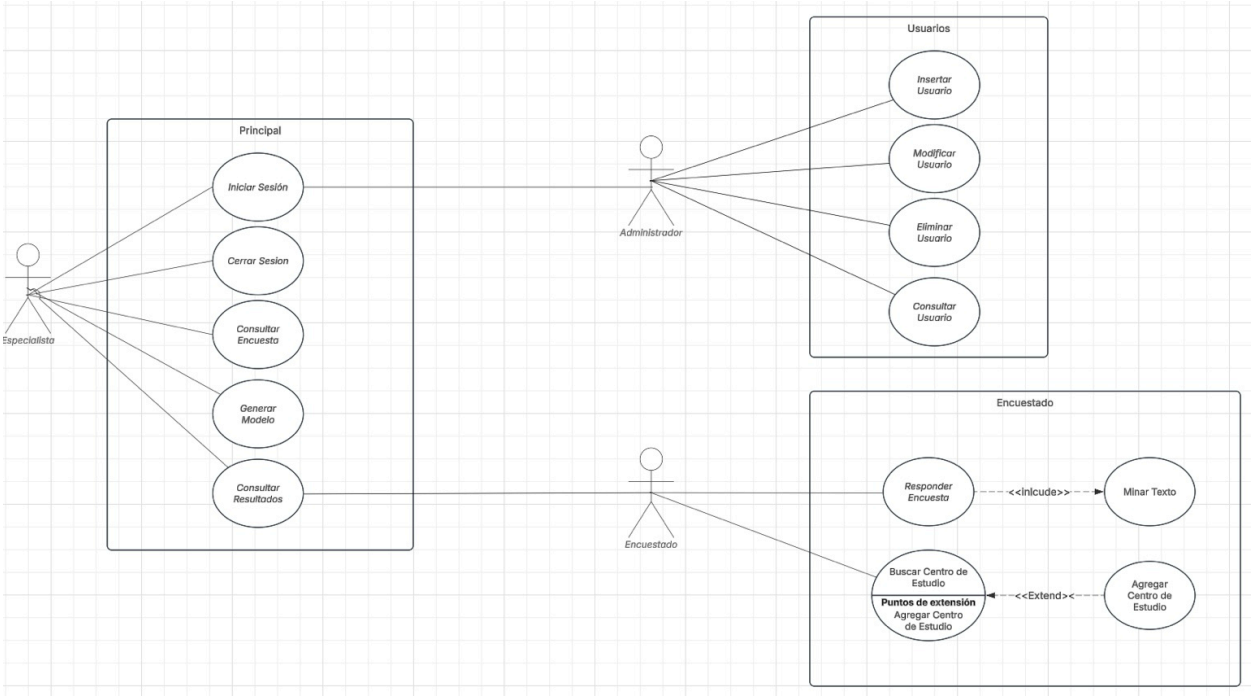


Figura 3.6 Diagrama de caso de uso del sistema

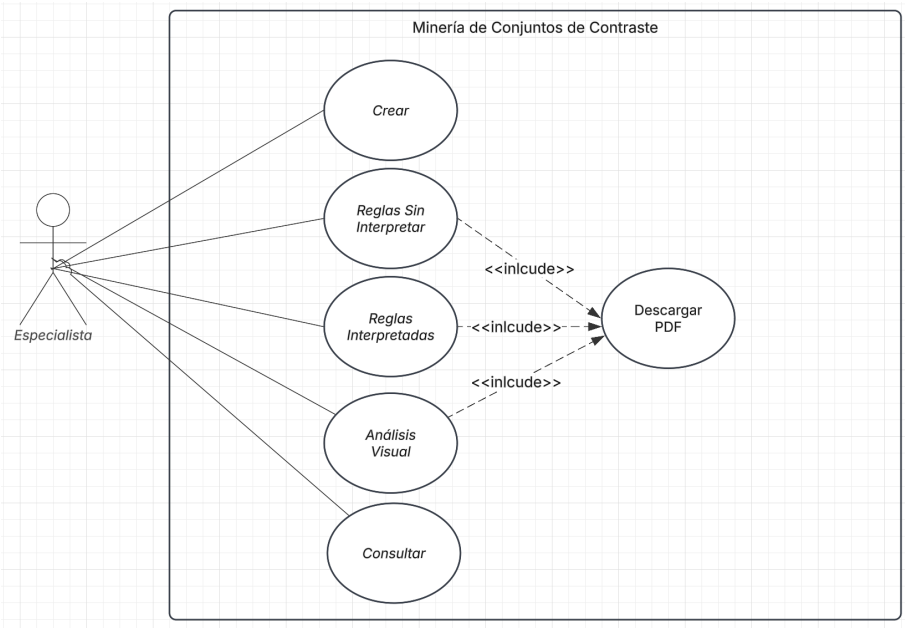


Figura 3.7 Diagrama de caso de uso del módulo de CSM

La Figura 3.7 presenta los casos de uso que habilita el módulo para el actor Especialista. El objetivo es cubrir el ciclo completo: creación del experimento, consulta de resultados, análisis visual, obtención de reglas (técnicas e interpretadas) y su exportación en formato PDF.

La representación de los casos de uso permite definir de manera estructurada las interacciones entre los distintos actores del sistema y las funcionalidades que cada uno ejecuta según su perfil. A través de los diagramas elaborados se establece una visión integral del comportamiento del sistema, identificando las operaciones críticas, las dependencias funcionales y los puntos de extensión entre los módulos.
En particular, los casos de uso del módulo de Minería de Conjuntos de Contraste describen el flujo completo desde la creación del experimento hasta la interpretación y exportación de resultados, mientras que el diagrama general abarca las actividades de gestión administrativa, aplicación de encuestas y análisis de datos, garantizando la trazabilidad en todo el proceso.

3.3.3 Módulo web

El flujo de la aplicación web organiza la navegación de forma orientada a tareas para reducir fricción cognitiva y asegurar trazabilidad desde la gestión de encuestas hasta la Minería de Conjuntos de Contraste (CSM). El diagrama mostrado en la Figura 3.8 sintetiza los puntos de entrada, rutas principales, subflujos analíticos y acciones de exportación, alineados con la arquitectura MVC y los permisos de usuario.

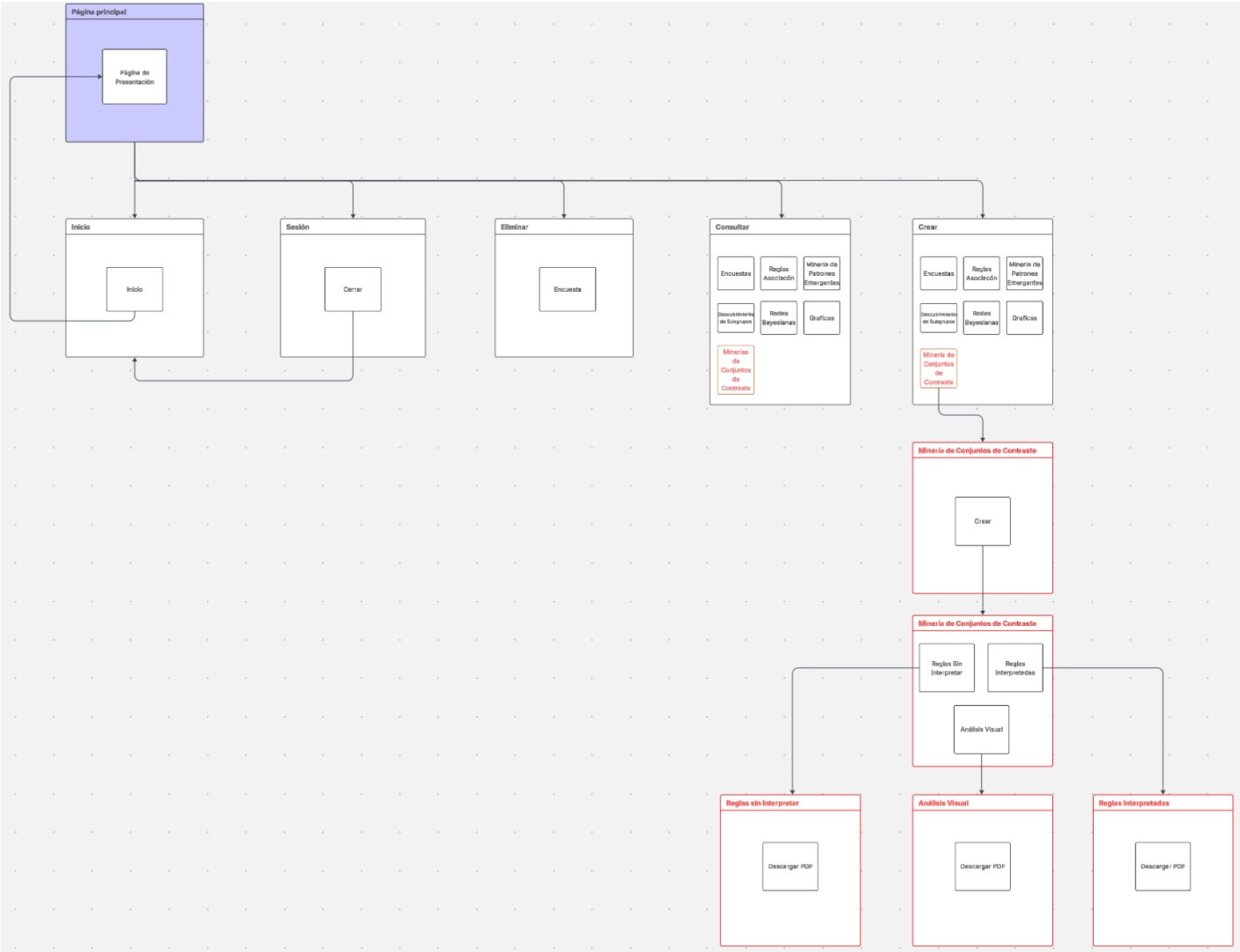


Figura 3.8 Diagrama de flujo de aplicación web

La Figura 3.9 presenta la interfaz desarrollada para el módulo de Minería de Conjuntos de Contraste. Este módulo se integra dentro del sistema web del Hospital Regional Río Blanco, constituyendo el entorno visual en el que los especialistas ejecutan, analizan e interpretan los resultados obtenidos por el algoritmo RuleKit2-CS.

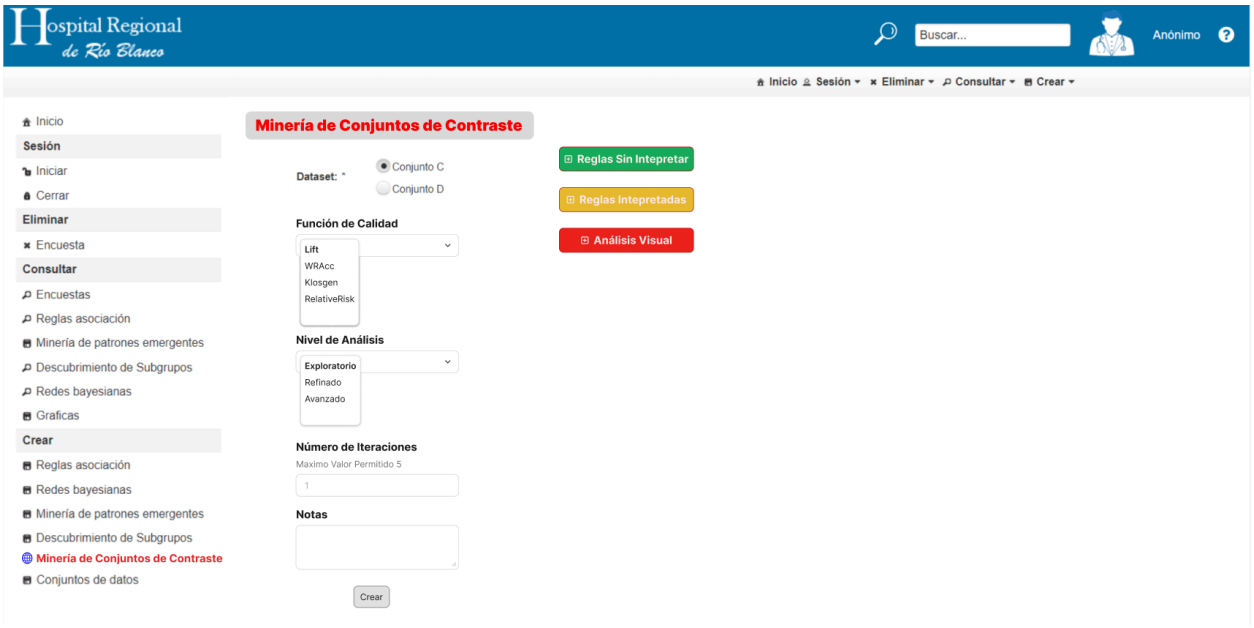


Figura 3.9 Pantalla principal módulo CSM

El diseño mantiene una distribución visual clara y responsiva, con una barra lateral izquierda que agrupa las operaciones del sistema y un panel central que despliega los formularios y resultados del módulo activo.

Las secciones principales son:

- Selección de parámetros de ejecución

  El especialista selecciona los conjuntos de datos a comparar (Conjunto C o Conjunto D) que hacen referencia a la *vista_minable* y *matriz_binaria* respectivamente., define la función de calidad (*Lift, WRAcc, Klosgen o RelativeRisk*), el nivel de análisis (Exploratorio, Refinado o Avanzado), el número máximo de iteraciones y notas descriptivas del experimento. Esta etapa garantiza reproducibilidad y control sobre los parámetros de minería.
- Generación de resultados

Al presionar el botón Crear, el sistema ejecuta el algoritmo RuleKit2-CS en el servidor y habilita las tres vistas disponibles:

- Reglas sin interpretar (Figura 3.10): muestra el listado técnico de reglas generadas, junto con métricas como soporte, confianza, *lift* y tamaño del conjunto de contraste.
- Reglas interpretadas (Figura 3.11): despliega las reglas traducidas a lenguaje natural, indicando el hospital correspondiente y una interpretación textual de los patrones encontrados.
- Análisis visual (Figura 3.12): ofrece una representación gráfica de las reglas mediante diagramas de barras, facilitando la comparación de resultados por hospital o por métrica.
- Exportación de resultados

  Todas las vistas incluyen la opción Descargar PDF, que genera automáticamente un reporte técnico (Figura 3.13, 3.14).

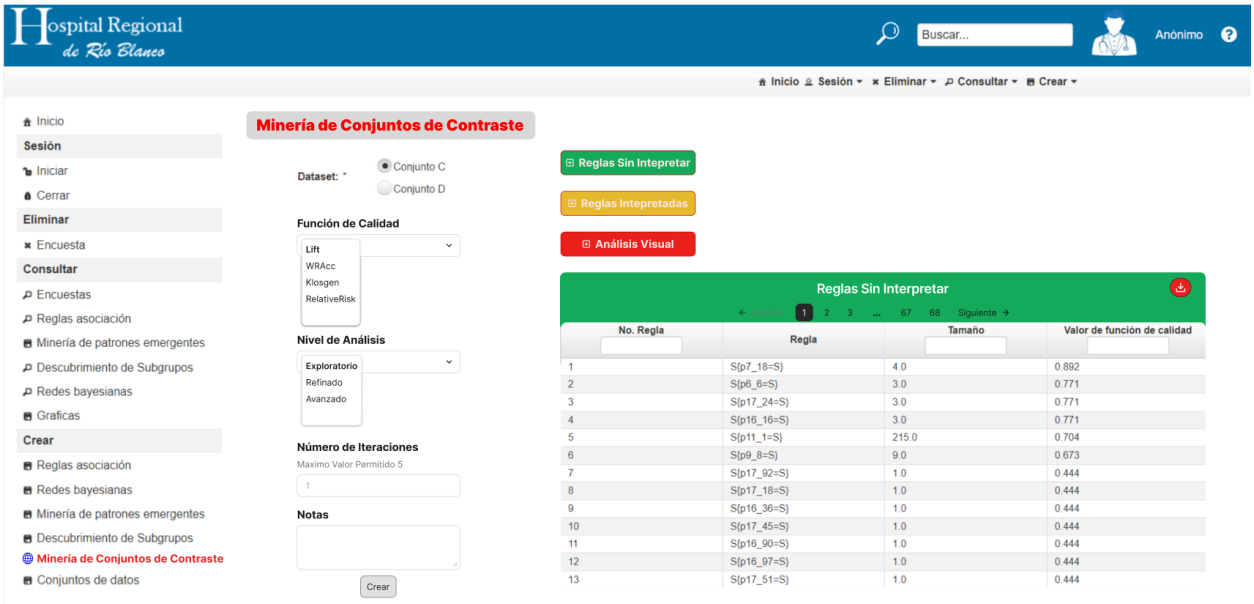


Figura 3.10 Pantalla reglas sin interpretar módulo CSM

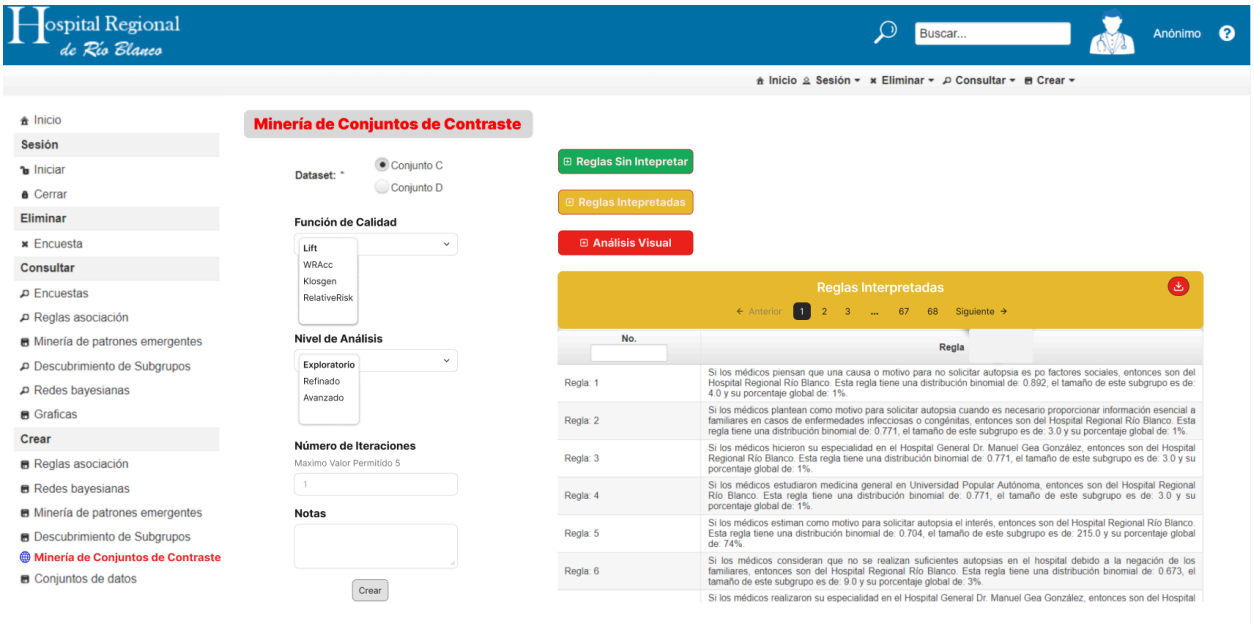


Figura 3.11 Pantalla reglas interpretadas módulo CSM

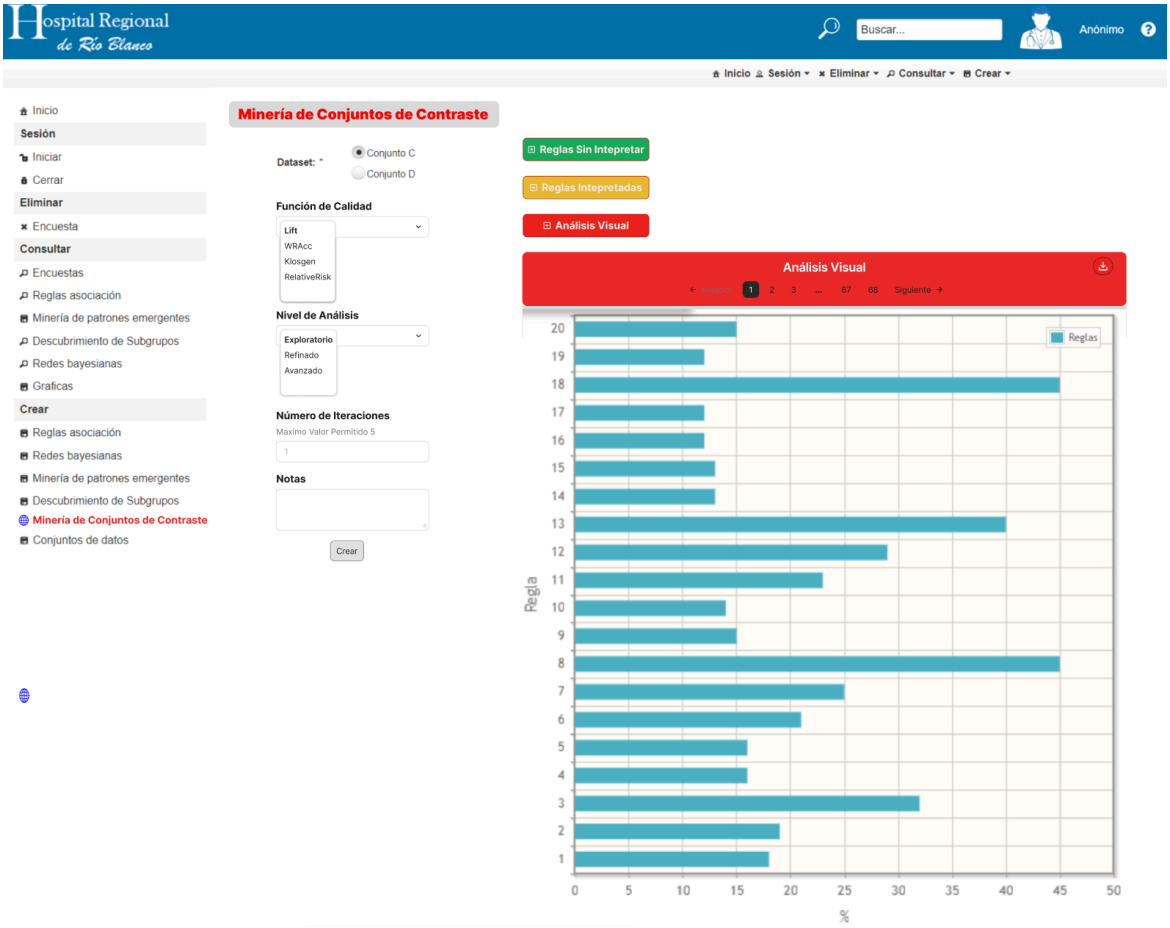
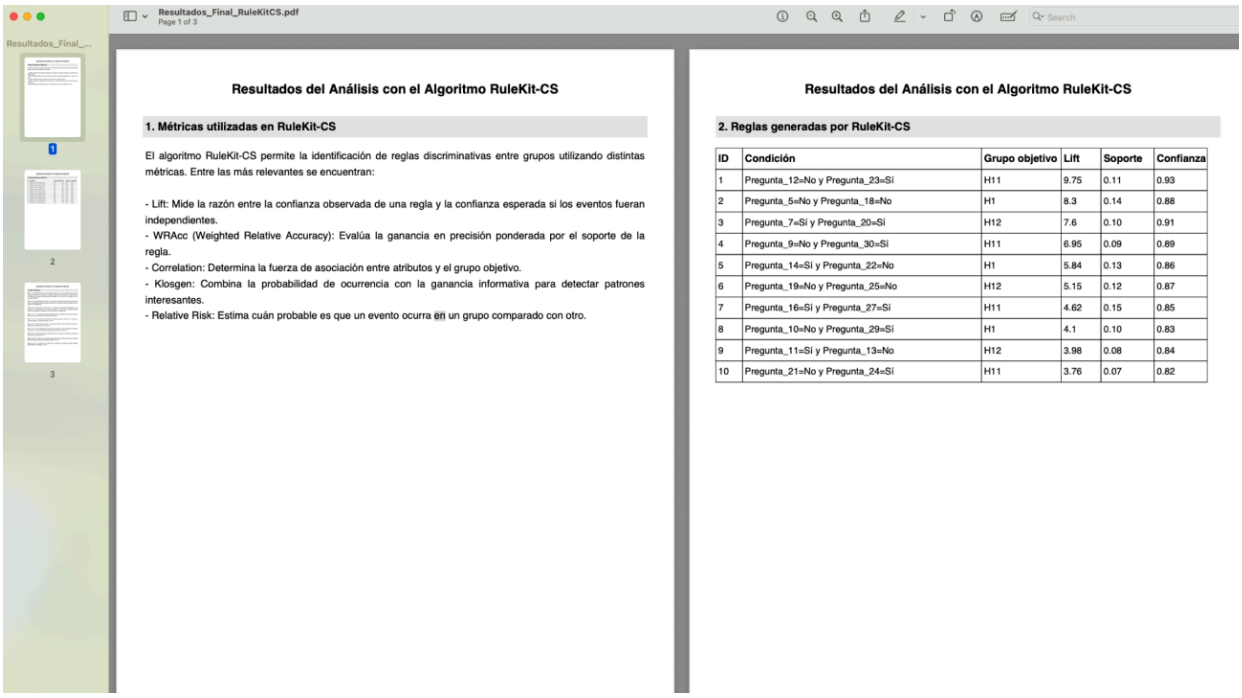


Figura 3.12 Pantalla análisis visual módulo CSM



Resultados del Análisis con el Algoritmo RuleKit-CS

1. Métricas utilizadas en RuleKit-CS

El algoritmo RuleKit-CS permite la identificación de reglas discriminativas entre grupos utilizando distintas métricas. Entre las más relevantes se encuentran:

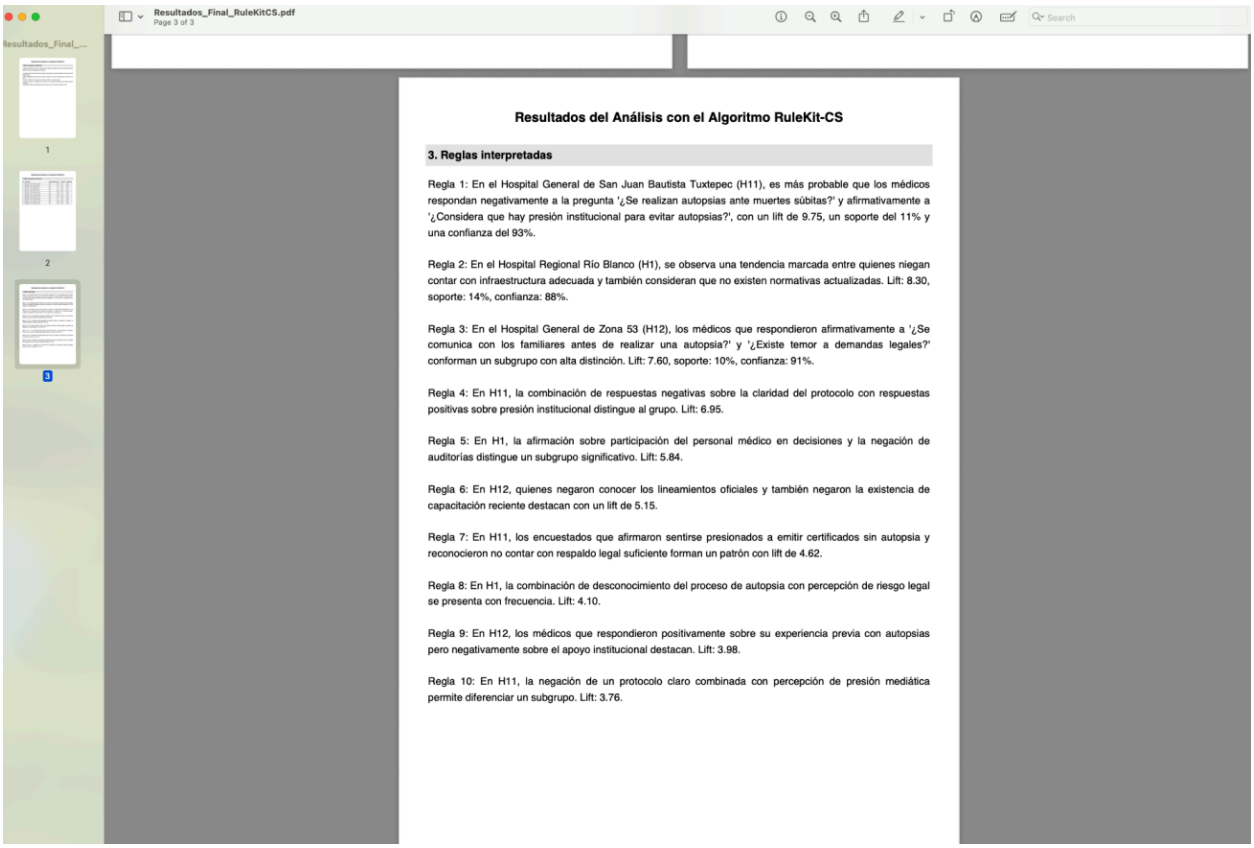
- Lift: Mide la razón entre la confianza observada de una regla y la confianza esperada si los eventos fueran independientes.
- WRAcc (Weighted Relative Accuracy): Evalúa la ganancia en precisión ponderada por el soporte de la regla.
- Correlation: Determina la fuerza de asociación entre atributos y el grupo objetivo.
- Klosgen: Combina la probabilidad de ocurrencia con la ganancia informativa para detectar patrones interesantes.
- Relative Risk: Estima cuán probable es que un evento ocurra en un grupo comparado con otro.

Resultados del Análisis con el Algoritmo RuleKit-CS

2. Reglas generadas por RuleKit-CS

ID	Condición	Grupo objetivo	Lift	Soporte	Confianza
1	Pregunta_12=No y Pregunta_23=Sí	H11	9.75	0.11	0.93
2	Pregunta_5=No y Pregunta_18=No	H1	8.3	0.14	0.88
3	Pregunta_7=Sí y Pregunta_20=Sí	H12	7.6	0.10	0.91
4	Pregunta_9=No y Pregunta_30=Sí	H11	6.95	0.09	0.89
5	Pregunta_14=Sí y Pregunta_22=No	H1	5.84	0.13	0.86
6	Pregunta_19=No y Pregunta_25=No	H12	5.15	0.12	0.87
7	Pregunta_16=Sí y Pregunta_27=Sí	H11	4.62	0.15	0.85
8	Pregunta_10=No y Pregunta_29=Sí	H1	4.1	0.10	0.83
9	Pregunta_11=Sí y Pregunta_13=No	H12	3.98	0.08	0.84
10	Pregunta_21=No y Pregunta_24=Sí	H11	3.76	0.07	0.82

Figura 3.13 Reporte PDF página 1 y 2 módulo CSM



Resultados del Análisis con el Algoritmo RuleKit-CS

3. Reglas interpretadas

Regla 1: En el Hospital General de San Juan Bautista Tuxtepec (H11), es más probable que los médicos respondan negativamente a la pregunta '¿Se realizan autopsias ante muertes súbitas?' y afirmativamente a '¿Considera que hay presión institucional para evitar autopsias?', con un lift de 9.75, un soporte del 11% y una confianza del 93%.

Regla 2: En el Hospital Regional Río Blanco (H1), se observa una tendencia marcada entre quienes niegan contar con infraestructura adecuada y también consideran que no existen normativas actualizadas. Lift: 8.30, soporte: 14%, confianza: 88%.

Regla 3: En el Hospital General de Zona 53 (H12), los médicos que respondieron afirmativamente a '¿Se comunica con los familiares antes de realizar una autopsia?' y '¿Existe temor a demandas legales?' conforman un subgrupo con alta distinción. Lift: 7.60, soporte: 10%, confianza: 91%.

Regla 4: En H11, la combinación de respuestas negativas sobre la claridad del protocolo con respuestas positivas sobre presión institucional distingue al grupo. Lift: 6.95.

Regla 5: En H1, la afirmación sobre participación del personal médico en decisiones y la negación de auditorías distingue un subgrupo significativo. Lift: 5.84.

Regla 6: En H12, quienes negaron conocer los lineamientos oficiales y también negaron la existencia de capacitación reciente destacan con un lift de 5.15.

Regla 7: En H11, los encuestados que afirmaron sentirse presionados a emitir certificados sin autopsia y reconocieron no contar con respaldo legal suficiente forman un patrón con lift de 4.62.

Regla 8: En H1, la combinación de desconocimiento del proceso de autopsia con percepción de riesgo legal se presenta con frecuencia. Lift: 4.10.

Regla 9: En H12, los médicos que respondieron positivamente sobre su experiencia previa con autopsias pero negativamente sobre el apoyo institucional destacan. Lift: 3.98.

Regla 10: En H11, la negación de un protocolo claro combinada con percepción de presión mediática permite diferenciar un subgrupo. Lift: 3.76.

Figura 3.14 Reporte PDF página 3 módulo CSM

El diseño privilegia la simplicidad y fluidez del flujo de trabajo:

- Los campos obligatorios están marcados con asterisco y las opciones de selección se presentan mediante listas desplegables.
- Las validaciones se ejecutan del lado del cliente utilizando PrimeFaces, mostrando mensajes contextuales en caso de errores o formatos inválidos.
- Los botones de acción (Reglas sin interpretar, Reglas interpretadas, Análisis visual) están codificados por colores verde, amarillo y rojo para reforzar la identificación visual de las etapas del análisis.
- El módulo mantiene la persistencia de parámetros y resultados en sesión, permitiendo reanálisis o descargas posteriores sin repetir la configuración.

El diseño de la interfaz gráfica del módulo web integra usabilidad, claridad visual y eficiencia funcional. La combinación de formularios, tablas interactivas, visualizaciones y reportes asegura que el especialista pueda controlar todo el proceso analítico desde una misma vista, alineando el entorno visual con los objetivos de la arquitectura del sistema y el modelo de minería de conjuntos de contraste.

3.4 Desarrollo del Módulo de Minería de Conjuntos de Contraste

Este apartado presenta el desarrollo e integración del Módulo de Minería de Conjuntos de Contraste dentro del sistema existente siguiendo Kanban como metodología de desarrollo siendo fundamental porque permitió gestionar el trabajo del módulo de manera visual, flexible y continua, adaptándose a la naturaleza iterativa y técnica del proyecto.
A diferencia de un modelo rígido o basado en fases cerradas, Kanban facilitó priorizar tareas en tiempo real, controlar el flujo de trabajo y evitar sobrecarga en el desarrollo del módulo de Minería de Conjuntos de Contraste.

Gracias a su implementación en GitHub Projects, fue posible monitorear el avance, identificar bloqueos y mantener trazabilidad directa entre cada tarea, *commit* y entrega, garantizando una integración progresiva, ordenada y eficiente del módulo dentro del sistema., El módulo se implementó bajo una arquitectura Modelo-Vista-Controlador y se

extendió con un entorno contenerizado que garantiza la reproducibilidad, portabilidad y despliegue estandarizado de todo el sistema.

El módulo fue diseñado para identificar diferencias sistemáticas entre hospitales mediante el análisis comparativo de encuestas, utilizando técnicas de CSM que permiten descubrir subpoblaciones contrastivas que distinguen las opiniones del personal médico respecto a la disminución de autopsias.

Desde el punto de vista arquitectónico, el módulo se acopla a la capa de aplicación existente en Java EE, donde la capa de presentación (páginas JSF) permite configurar y ejecutar experimentos; la capa de control (*conjuntosContrasteControler*) orquesta la extracción de datos, la ejecución del motor *RuleKit2-CS*, y la interpretación de resultados; mientras que la capa de persistencia (PostgreSQL) proporciona las vistas analíticas (*vista_minable_csm* y *matriz_binaria_csm*) y funciones PL/pgSQL para la traducción NLG (*Natural Language Generation*).

Esta integración mantiene la coherencia con la infraestructura del sistema, asegurando que la nueva funcionalidad se incorpore sin afectar los módulos existentes y bajo los mismos principios de escalabilidad y modularidad definidos en el diseño general. La Figura 3.2 muestra la arquitectura extendida del módulo CSM dentro del sistema principal.

3.4.1 Kanban

La Figura 3.15 muestra el tablero Kanban utilizado para la gestión del desarrollo del módulo de Minería de Conjuntos de Contraste, alojado en GitHub Projects bajo el repositorio estudio-autopsia.

En este tablero se organizaron las tareas técnicas y de integración correspondientes a las distintas fases del proyecto: contenerización, ajuste de funciones de traducción NLG, generación de reportes, construcción de vistas analíticas y depuración de métricas.

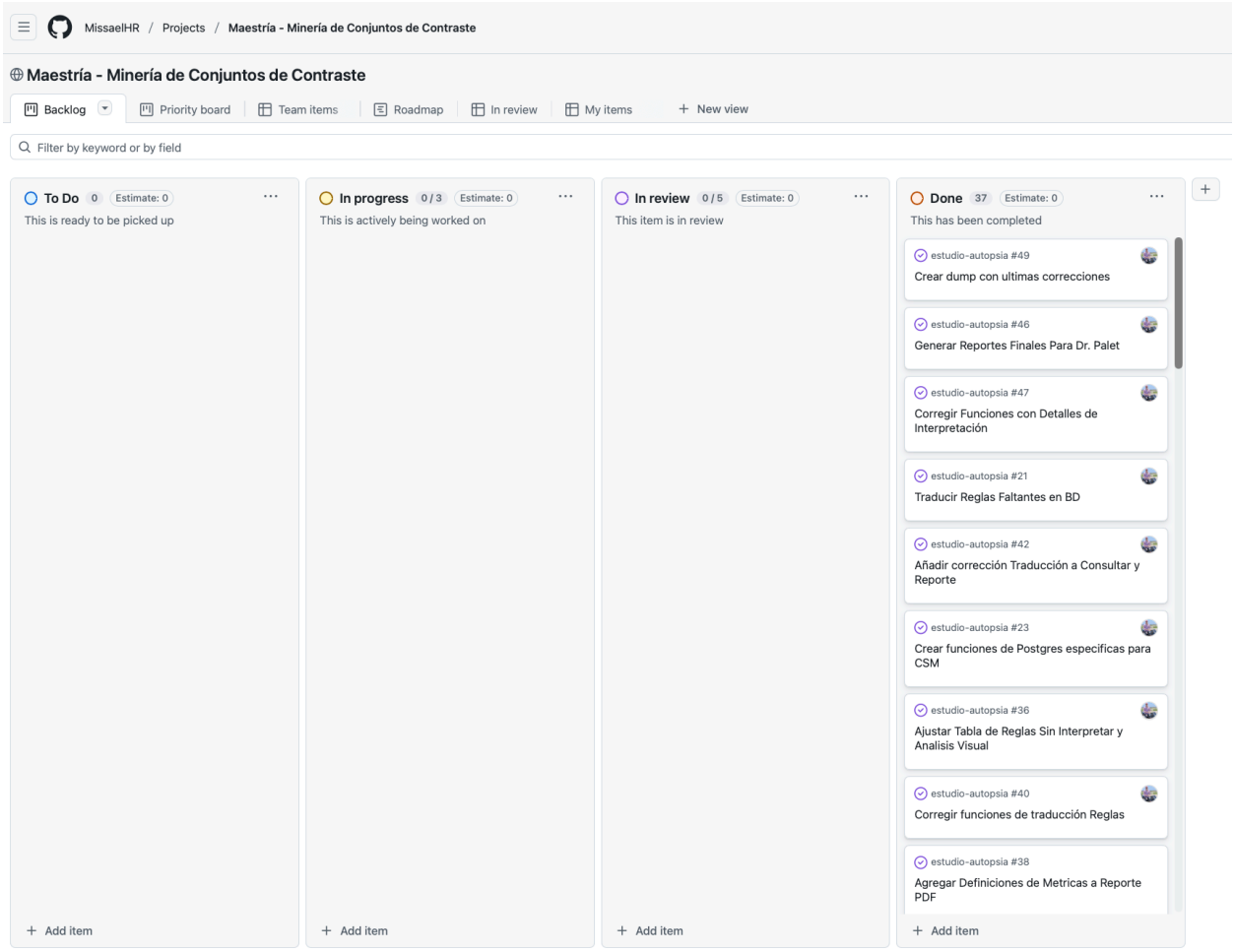


Figura 3.15 Tablero Kanban

Cada tarjeta del tablero representaba una tarea con trazabilidad directa al repositorio, identificada por un número de *issue* (por ejemplo, #49 Crear *dump* con últimas correcciones o #21 Traducir Reglas Faltantes en BD). El tablero se estructuró en las columnas *To Do*, *In Progress*, *In Review* y *Done*, lo que permitió seguir el flujo completo desde la planificación hasta la finalización de cada funcionalidad. Gracias a esta organización, fue posible medir la carga de trabajo activa, los tiempos de ciclo por tarea y mantener un registro visual del avance acumulado del módulo.

La Tabla 3.11 muestra un extracto representativo del tablero gestionado. Cada tarea refleja una acción técnica o funcional desarrollada durante el ciclo de implementación, con su estado final, tipo de actividad y categoría del módulo al que pertenece.

Tabla 3.11 Tablero Kanban

Título	URL	Asignación	Estatus
Crear Respaldo Con Últimas Correcciones	https://github.com/MissaelHR/estudio-autopsia/issues/49	MissaelHR	Completado
Generar Reportes Finales Para Dr. Palet	https://github.com/MissaelHR/estudio-autopsia/issues/46	MissaelHR	Completado
Corregir Funciones Con Detalles De Interpretación	https://github.com/MissaelHR/estudio-autopsia/issues/47	MissaelHR	Completado
Traducir Reglas Faltantes En BD	https://github.com/MissaelHR/estudio-autopsia/issues/21	MissaelHR	Completado
Añadir Corrección Traducción A Consultar Y Reporte	https://github.com/MissaelHR/estudio-autopsia/issues/42	MissaelHR	Completado
Crear Funciones De PostgreSQL Específicas Para CSM	https://github.com/MissaelHR/estudio-autopsia/issues/23	MissaelHR	Completado
Ajustar Tabla De Reglas Sin Interpretar Y Análisis Visual	https://github.com/MissaelHR/estudio-autopsia/issues/36	MissaelHR	Completado
Corregir Funciones De Traducción Reglas	https://github.com/MissaelHR/estudio-autopsia/issues/40	MissaelHR	Completado
Agregar Definiciones De Métricas A Reporte PDF	https://github.com/MissaelHR/estudio-autopsia/issues/38		Completado
Agregar Funcionalidad De Módulo A Sección Consultar	https://github.com/MissaelHR/estudio-autopsia/issues/33	MissaelHR	Completado
Generar Reporte PDF	https://github.com/MissaelHR/estudio-autopsia/issues/28	MissaelHR	Completado
Generar Gráficos	https://github.com/MissaelHR/estudio-autopsia/issues/22	MissaelHR	Completado
Implementar Lógica Algoritmo	https://github.com/MissaelHR/estudio-autopsia/issues/20	MissaelHR	Completado
Funciones Del Módulo	https://github.com/MissaelHR/estudio-autopsia/issues/14	MissaelHR	Completado

Tabla 3.11 Tablero Kanban – Continuación

Título	URL	Asignación	Estatus
Desarrollar Módulo UI	https://github.com/MissaelHR/estudio-autopsia/issues/19	MissaelHR	Completado
MAESTRIA-3 Agregar *Epm-Framework-Master* A Contenedor	https://github.com/MissaelHR/estudio-autopsia/pull/4	MissaelHR	Completado
MAESTRIA-5 Crear Modelos De Configuración XML	https://github.com/MissaelHR/estudio-autopsia/pull/18	MissaelHR	Completado
MAESTRIA-4 Cambios Iniciales Y Configuración De Modelos XML	https://github.com/MissaelHR/estudio-autopsia/pull/16	MissaelHR	Completado
Desarrollar Prueba Con *Release* Rulekit2-CS	https://github.com/MissaelHR/estudio-autopsia/issues/9	MissaelHR	Completado
MAESTRIA-6 Primeros Cambios Y Ejecución De Rulekit2-CS	https://github.com/MissaelHR/estudio-autopsia/pull/24	MissaelHR	Completado
Analizar Si Es Viable Crear Un Servicio Externo	https://github.com/MissaelHR/estudio-autopsia/issues/12	MissaelHR	Completado
MAESTRIA-2 Actualizar Rutas Incorrectas	https://github.com/MissaelHR/estudio-autopsia/pull/3	MissaelHR	Completado
Identificar Conjuntos De Datos A Utilizar - DB	https://github.com/MissaelHR/estudio-autopsia/issues/13	MissaelHR	Completado
MAESTRIA-7 Soporte De Ventanas Para Reglas Sin Interpretar	https://github.com/MissaelHR/estudio-autopsia/pull/25	MissaelHR	Completado
MAESTRIA-1 Soporte Para Creación De Respaldos De Base De Datos	https://github.com/MissaelHR/estudio-autopsia/pull/2	MissaelHR	Completado
MAESTRIA-0 Configuración De Docker	https://github.com/MissaelHR/estudio-autopsia/pull/1	MissaelHR	Completado
MAESTRIA-8 Funciones Postgresql CSM Datasets Vista_Minable Y Matriz_Binaria	https://github.com/MissaelHR/estudio-autopsia/pull/26	MissaelHR	Completado

Tabla 3.11 Tablero Kanban – Continuación

Título	URL	Asignación	Estatus
Comparar JAR vs Consumo Externo	https://github.com/MissaelHR/estudio-autopsia/issues/10	MissaelHR	Completado
Crear Modelos de Configuración XML	https://github.com/MissaelHR/estudio-autopsia/issues/17	MissaelHR	Completado
Acortar El Numero de Atributos de la Regla	https://github.com/MissaelHR/estudio-autopsia/issues/15	MissaelHR	Completado
Analizar Release RuleKit2-CS	https://github.com/MissaelHR/estudio-autopsia/issues/8	MissaelHR	Completado
Analizar Ultima Versión Sistema Estudio Autopsia	https://github.com/MissaelHR/estudio-autopsia/issues/5	MissaelHR	Completado
Contenerizar Proyecto - Docker	https://github.com/MissaelHR/estudio-autopsia/issues/7	MissaelHR	Completado
MAESTRIA-9 Traducción de Reglas	https://github.com/MissaelHR/estudio-autopsia/pull/27	MissaelHR	Completado
Mejoras UI	https://github.com/MissaelHR/estudio-autopsia/issues/31	MissaelHR	Completado

El uso de la metodología Kanban resultó esencial para mantener una gestión eficiente y ordenada del ciclo de desarrollo del módulo. Su aplicación permitió visualizar el estado real del proyecto en todo momento, identificar rápidamente tareas bloqueadas y reajustar prioridades sin interrumpir el flujo de trabajo.

Además, al integrarse directamente con el repositorio de GitHub que se encuentra en un repositorio privado en el siguiente enlace https://github.com/MissaelHR/estudio-autopsia, cada tarea del tablero quedó vinculada con su respectivo *commit*, *rama* o *pull request*, lo que facilitó un seguimiento transparente y trazable de cada cambio en el código.

Desde una perspectiva de ingeniería de software, Kanban aportó un esquema de entrega continua y mejora incremental, optimizando la planeación de versiones y la gestión del tiempo invertido por actividad.

Su flexibilidad permitió abordar simultáneamente tareas técnicas (contenerización, ajuste de funciones *PL/pgSQL*, integración *NLG*) y operativas (documentación, pruebas,

reportes) sin necesidad de iteraciones rígidas, asegurando un avance sostenido y un producto final estable y verificable.

3.4.2 Contenerización y despliegue

El sistema fue completamente contenerizado para asegurar un entorno de ejecución estable, portable y reproducible en cualquier sistema operativo. Inicialmente, el desarrollo se realizó en entornos Windows, dependientes de rutas absolutas y configuraciones nativas de NetBeans. Sin embargo, durante la fase de integración del módulo de CSM y ante la migración del entorno de desarrollo a *macOs* se detectaron múltiples incompatibilidades de configuración, diferencias en la gestión de rutas y codificación de archivos, así como dependencias acopladas al *IDE*.

Por ello, se implementó una contenerización integral con Docker, acompañada de un proceso de compilación independiente del *IDE* mediante *Apache Ant* y de un sistema de automatización con *Makefile*, lo que permitió consolidar un flujo de despliegue nativo, limpio y replicable.

La contenerización aporta ventajas sustanciales en el contexto del desarrollo científico y académico:

- **Reproducibilidad y consistencia**: cada experimento ejecutado dentro del contenedor produce resultados idénticos independientemente del entorno del sistema operativo o equipo donde se ejecute, lo que elimina dependencias externas y garantiza integridad en las ejecuciones de *RuleKit2-CS* y en las traducciones NLG.
- **Portabilidad y despliegue simplificado**: el sistema puede ejecutarse en cualquier entorno compatible con Docker —local, nube o clúster institucional— sin requerir reconfiguración, gracias a la abstracción de red y almacenamiento gestionada por *docker-compose.yml.*
- **Aislamiento de servicios**: la base de datos y la aplicación web junto con las bibliotecas utilizadas en todos los módulos para la ejecución de minería de datos en específico el descubrimiento supervisado de reglas descriptivas operan en

contenedores separados, comunicándose a través de una red interna segura, evitando conflictos de dependencias o puertos.

- **Escalabilidad y mantenibilidad**: la arquitectura modular facilita la actualización independiente de componentes (por ejemplo, actualizar *RuleKit2-CS* sin reconstruir toda la aplicación), permitiendo incluso escalar horizontalmente la capa de aplicación o incorporar nuevos algoritmos de minería en contenedores paralelo de ser necesario.
- **Integración continua y trazabilidad**: la definición declarativa del entorno en el archivo *Dockerfile* y el *docker-compose.yml* permite la integración en flujos de *CI/CD*, asegurando versiones controladas y documentadas del sistema y su base de datos.

El entorno se compone de dos servicios definidos en *docker-compose.yml*: uno para la aplicación web basada en Apache Tomcat 9, y otro para la base de datos PostgreSQL. Ambos se ejecutan en una red interna Docker (*red_autopsias*), comunicándose por nombre de servicio, eliminando la necesidad de configuraciones locales en el *web.xml*. El código base se observa en el Listado 3.1:

Listado 3.1. Definición de servicios y orquestación de contenedores en Docker Compose

```
services:
  postgres:
    image: postgres:latest
    container_name: postgres
    restart: unless-stopped
    environment:
      POSTGRES_DB: encuesta
      POSTGRES_USER: postgres
      POSTGRES_PASSWORD: password
    ports:
      - "5432:5432"
    volumes:
      - postgres_data:/var/lib/postgresql/data
    networks:
      - red_autopsias

```

Listado 3.1. Definición de servicios y orquestación de contenedores en Docker Compose - Continuación

```
  tomcat:
    build: .
    container_name: autopsias
    depends_on:
      - postgres
    ports:
      - "8080:8080"
    environment:
      - TZ=America/Mexico_City
    networks:
      - red_autopsias

volumes:
  postgres_data:

networks:
  red_autopsias:
    driver: bridge
```

El contenedor de PostgreSQL ejecuta las funciones PL/pgSQL de traducción NLG y materializa las vistas *vista_minable_csm* y *matriz_binaria_csm* así como todas las estructuras de datos mostradas en los modelos de la Figura 3.5.
El contenedor de Tomcat despliega el archivo .war construido con Apache Ant, gestionando toda la capa de aplicación.

El proceso de construcción de la imagen de Tomcat se detalla en el *Dockerfile* (Listado 3.2):

Listado 3.2. Dockerfile para la construcción de la imagen del servidor de aplicaciones

```
FROM tomcat:9.0-jdk8-openjdk

# Borra las apps por defecto (opcional pero recomendable)
RUN rm -rf /usr/local/tomcat/webapps/*

# Copia la carpeta corpus al directorio corpus/ del contenedor
```

Listado 3.2. Dockerfile para la construcción de la imagen del servidor de aplicaciones - Continuación

```
 7. COPY corpus/ corpus/
 8.
 9. # Copia la carpeta epm-framework al directorio epm-framework/ del contenedor
10. COPY epm-framework-master/ epm-framework-master/
11.
12. # Copia la carpeta mineria-conjuntos-contraste al directorio mineria-conjuntos-contraste/ del contenedor
13. COPY mineria-conjuntos-contraste/ mineria-conjuntos-contraste/
14.
15. # Copia tu archivo WAR al directorio de despliegue de Tomcat
16. COPY dist/EstudioAutopsia.war /usr/local/tomcat/webapps/ROOT.war
17.
18. # Exponer el puerto 8080 (Tomcat default)
19. EXPOSE 8080
20.
21. # El CMD viene por defecto en la imagen base de Tomcat
```

3.4.3 Integración con Ant y Makefile

Para independizar la compilación del IDE y resolver los problemas de configuración nativa de NetBeans, el proyecto fue reconstruido sobre Apache Ant con un archivo build.xml que define de forma explícita las tareas de compilación, empaquetado y despliegue.

Como se observa en el Listado 3.3, el WAR se genera por la línea de comandos:

Listado 3.3. Comando de compilación y empaquetado mediante Apache Ant

```
1. ant clean war
```

El archivo Makefile constituye el mecanismo central de automatización del proceso de compilación y despliegue del sistema Estudio Autopsia dentro de su entorno contenerizado.

Su uso permite abstraer los comandos necesarios para construir el proyecto, generar el archivo *.war,* crear las imágenes Docker y administrar los contenedores, evitando dependencias directas del entorno de desarrollo (NetBeans, terminal o *scripts* aislados).

Este archivo sigue la sintaxis estándar de GNU Make, dividiéndose en tres bloques principales:

- Declaración de variables globales,
- Definición de objetivos (*targets*),
- Comandos de automatización para Docker y base de datos

A continuación, se detalla su estructura:

1. Encabezado y declaración de variables (Listado 3.4).

Listado 3.4. Makefile: Declaración de variables globales y configuración del entorno

```
.PHONY: all build prepare image deploy up down logs logs-tomcat db-up db-down clean

# Configuración de variables
WAR_NAME       := EstudioAutopsia.war
WAR_SRC        := dist/$(WAR_NAME)
PG_CONTAINER   := postgres
PG_DB          := encuesta
PG_USER        := postgres
PG_PASSWORD    := password
TOMCAT_SERVICE := tomcat
DB_SERVICE     := postgres

# Parámetros de DB para pruebas locales (puedes sobreescribirlos al invocar make)
DB_URL     ?= jdbc:postgresql://localhost:5432/$(PG_DB)
DB_USER    ?= $(PG_USER)
DB_PASS    ?= $(PG_PASSWORD)
DB_DRIVER  ?= org.postgresql.Driver

# Classpath para ejecutar utilidades Java desde build/classes
CLASSES_DIR := build/classes
WEB_LIB_DIR := web/WEB-INF/lib
CP := $(CLASSES_DIR):$(WEB_LIB_DIR)/*
```

Estas variables permiten ejecutar el *Makefile* en distintos equipos sin modificar manualmente los *scripts* o los archivos de configuración.

También se definen los parámetros de conexión JDBC utilizados en pruebas locales, permitiendo sobrescribirlos en tiempo de ejecución con variables de entorno (DB_URL, DB_USER, DB_PASS).

2. Compilación y construcción del proyecto (Listado 3.5).

Listado 3.5. Makefile: Objetivo de compilación y generación del WAR

```
1. build:
2.         @echo "Compilando WAR con Ant..."
3.         ant clean dist
```

El objetivo *build* invoca Apache Ant para compilar el código fuente y generar el artefacto *EstudioAutopsia.war* dentro del directorio *dist/.*

El uso de Ant garantiza que el proceso sea determinista, libre de configuraciones del IDE, y reproducible en cualquier entorno Linux/macOS.

3. Construcción de la imagen Docker (Listado 3.6).

Listado 3.6. Makefile: Objetivo de construcción de la imagen Docker

```
1. image: prepare
2.     @echo "Construyendo imagen Docker de $(TOMCAT_SERVICE)..."
3.     docker-compose build $(TOMCAT_SERVICE)
```

El *target image* ejecuta *docker-compose build* sobre el servicio de Tomcat, generando la imagen personalizada definida en el *Dockerfile* del proyecto.

4. Despliegue automatizado del sistema.

   El objetivo *deploy* combina los pasos anteriores (*build* + *image*) y levanta el contenedor Tomcat de manera automática, garantizando un flujo *build* → *package* → *run* uniforme como se observa en el Listado 3.7.

Listado 3.7. Makefile: Objetivo de despliegue automatizado de contenedores

```
deploy: build image
    @echo "Desplegando contenedor $(TOMCAT_SERVICE)..."
    docker-compose up -d $(TOMCAT_SERVICE)
```

5. Gestión de servicios y *logs* (Listado 3.8).

Listado 3.8. Makefile: Objetivos para la gestión y monitoreo de registros (*logs*)

```
logs:
    docker-compose logs -f

logs-tomcat:
    docker-compose logs -f $(TOMCAT_SERVICE)
```

El archivo incluye comandos para iniciar, detener o consultar *logs* de manera granular, tanto del conjunto completo de contenedores como de servicios individuales.

6. Operaciones de base de datos.

   Los objetivos *db-import* y *db-dump-bin* automatizan la importación y exportación de copias de seguridad en formato binario *.dump* utilizando los comandos *pg_restore* y *pg_dump*. Esto permite restaurar o respaldar el estado de la base de datos de forma consistente dentro del contenedor mostrado en el Listado 3.9.

Listado 3.9. Makefile: Automatización de respaldo y restauración de base de datos

```
db-import:
    @echo "Importando archivo .dump binario..."
    docker cp $(DUMP_FILE) $(PG_CONTAINER):/tmp/backup.dump
    docker exec -e PGPASSWORD=$(PG_PASSWORD) $(PG_CONTAINER) \
        pg_restore -U $(PG_USER) -d $(PG_DB) --clean /tmp/backup.dump

db-dump-bin:
    @echo "Creando dump binario de la base de datos..."
    docker exec -e PGPASSWORD=$(PG_PASSWORD) $(PG_CONTAINER) \
        pg_dump -U $(PG_USER) -d $(PG_DB) -F c -f /tmp/backup.dump
    docker cp $(PG_CONTAINER):/tmp/backup.dump db/backup.dump
    @echo "Dump binario guardado en db/backup.dump"
```

7. Limpieza de imágenes huérfanas (Listado 3.10).
   El objetivo *delete-images-none* elimina las imágenes *dangling* de Docker (sin etiqueta <*none*>), optimizando el uso de espacio y manteniendo el entorno limpio.

Listado 3.10. Makefile: Rutina de limpieza de imágenes Docker huérfanas

```
1. delete-images-none: docker rmi $(docker images -f "dangling=true" -q)
```

El uso del *Makefile* en este proyecto representó un avance significativo en la automatización del ciclo de vida del módulo CSM, al eliminar las dependencias manuales y garantizar un proceso reproducible, portable y documentado. Además, integró en un solo flujo las herramientas *Ant*, *Docker Compose* y *PostgreSQL*, permitiendo construir, desplegar, importar o respaldar datos con un solo comando.

3.4.4 Construcción de Vistas y Funciones de Soporte en PostgreSQL

Como parte del desarrollo del módulo de minería de conjuntos de contraste, fue necesario crear una nueva estructura de vistas y funciones dentro de PostgreSQL que permitiera homologar completamente la forma en que se generan los conjuntos de datos analíticos. En particular, la creación de las vistas *vista_minable_csm* y *matriz_binaria_csm* respondió a la necesidad de obtener un formato uniforme, ordenado y directamente utilizable por el algoritmo de RuleKit2-CS, el cual requiere un conjunto de datos con un orden de columnas estable y una columna de grupo objetivo claramente definida.

Ambas vistas fueron diseñadas con el propósito de estandarizar la salida de datos, asegurando que cada instancia analítica mantenga la misma estructura, que las columnas sigan un orden lógico consistente, y que la columna *centro_hospitalario* se encuentre ubicada de forma explícita como el atributo que representa el grupo objetivo. Esto resulta indispensable para la ejecución de RuleKit2-CS, ya que este algoritmo utiliza dicha columna para diferenciar los subconjuntos de contraste y derivar patrones discriminantes entre hospitales.

Al concentrar esta lógica dentro de PostgreSQL, se eliminó la dependencia de *scripts* externos y se consiguió un flujo más robusto y reproducible. La base de datos pasó de funcionar únicamente como repositorio a convertirse en un componente activo, responsable de preparar los datos, realizar transformaciones especializadas, garantizar la consistencia estructural y semántica, y generar salidas listas para ser consumidas por RuleKit2-CS.

En este punto se describen estas nuevas vistas, las funciones auxiliares que permiten su construcción y verificación, así como el conjunto de rutinas utilizado para alimentar el módulo de generación de lenguaje natural (NLG). Estos elementos, en conjunto, fortalecen el flujo de procesamiento y consolidan una capa de datos sólida.

Función *vista_minable_csm*

La función *constituye* el componente central del proceso metodológico de preparación de datos empleado para generar el conjunto de análisis utilizado por el algoritmo *RuleKit2-CS*. Su diseño está orientado a transformar una base de datos altamente normalizada en un *dataset* plano, homogéneo y estructuralmente estable, apto para ser consumido directamente por los métodos de minería de conjuntos de contraste.

Desde su inicio, la función adopta un comportamiento idempotente. Antes de reconstruir la tabla final, verifica su existencia y la elimina explícitamente para evitar acumulación de datos residuales (Listado 3.11).

Listado 3.11. Rutina de limpieza e inicialización de la tabla vista_minable_csm

```
1. IF v_table THEN
2.   EXECUTE 'DROP TABLE public.vista_minable_csm';
3. END IF;
```

Esto garantiza que cada ejecución produzca un *dataset* limpio y reproducible, independiente de ejecuciones anteriores.

Una vez eliminada la tabla previa, la función crea una nueva estructura donde se define un orden fijo de atributos. De forma intencional, la primera columna corresponde al

atributo *centro_hospitalario,* requisito indispensable para RuleKit2-CS, ya que este algoritmo interpreta la primera columna como la variable objetivo que determina los grupos a comparar (Listado 3.12).

Listado 3.12. Definición estructural (DDL) de la vista minable

```
CREATE TABLE public.vista_minable_csm(
  centro_hospitalario text NOT NULL,
  id_instancia        integer PRIMARY KEY,
  area                varchar(20),
  categoria           varchar(40),
  ult_grado           varchar(20),
  esc_med_gral        varchar(100),
  esc_esp             varchar(100),
  ...
);
```

Esta estructura fija previene problemas de interpretación posteriores y asegura que las reglas generadas sean trazables a atributos semánticos claros.

Para enfocar el análisis exclusivamente en los centros incluidos en el estudio, la función incorpora un filtrado explícito. Después de identificar el hospital del médico, solo se procesan las instancias correspondientes a H1, H11 y H12 (Listado 3.13).

Listado 3.13. Lógica de filtrado de instancias por centro hospitalario (Vista Minable)

```
IF hospital_code NOT IN ('H1','H11','H12') THEN
  CONTINUE;
END IF;
```

Con ello se garantiza que el *dataset* final represente únicamente las unidades de análisis relevantes para los objetivos de la investigación.

La función realiza un proceso de desnormalización que integra múltiples dimensiones del perfil profesional del médico desde tablas relacionadas (Listado 3.14).

Listado 3.14. Consultas de extracción y desnormalización de atributos médicos

```
SELECT a.sdescc INTO area_medico
FROM area a JOIN medico m ON a.idarea = m.idarea
WHERE m.idmedico = i;

SELECT ce2.snombrec INTO emg
FROM centroestudio ce2
...
WHERE f.idmedico = i AND f.idgrado = 2;
```

Atributos como área, categoría laboral, grado académico, escuelas de formación, años de ejercicio y experiencia en autopsias son extraídos mediante consultas relacionales y transformados en valores textuales que se integran al *dataset* final. Esta información adicional fortalece el valor analítico de RuleKit2-CS al permitir contrastes más ricos entre dimensiones sociodemográficas y percepciones profesionales.

La función también transforma las respuestas de las encuestas, que originalmente se encuentran distribuidas en varias tablas, en atributos textuales consolidados. Las respuestas se obtienen mediante consultas específicas por pregunta, tanto para reactivos de opción única como de opción múltiple. La inserción en la tabla final se realiza mediante SQL parametrizado para mantener la integridad tipada (Listado 3.15).

Listado 3.15. Inserción parametrizada de registros en la vista minable

```
EXECUTE
  $sql$
    INSERT INTO public.vista_minable_csm(centro_hospitalario, id_instancia, area, categoria,
                                         ult_grado, esc_med_gral, esc_esp, anios_prac, casos,
                                         hall_disc, hall_arb, hall_dem, mcc_aut, mcc_no_aut,
                                        rechazo_fam, no_hosp, per_sol_aut, med_aut, fmr_sol_aut,
                                         com_sug_op)
    VALUES ($1,$2,$3,$4,$5,$6,$7,$8,$9,$10,$11,$12,$13,$14,$15,$16,$17,$18,$19,$20)
  $sql$
USING hospital_code, n_inst, area_medico, cat_med, ugrado_med, emg, ee,
      r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8, r_9, r_10, r_11, r_12, r_13;
```

El uso de este mecanismo parametrizado asegura un orden garantizado de columnas, evita inyecciones accidentales y mantiene la consistencia semántica del *dataset*.
Finalmente, después de completar la carga del *dataset*, la columna técnica *id_instancia* utilizada únicamente para identificar internamente cada combinación de respuestas es eliminada para dejar únicamente los atributos semánticos que RuleKit2-CS utilizará durante el análisis (Listado 3.16).

Listado 3.16. Limpieza de columnas técnicas en la vista minable

```
1. ALTER TABLE public.vista_minable_csm DROP COLUMN id_instancia;
```

Este paso simplifica la estructura del *dataset* final y evita la incorporación de identificadores irrelevantes para el proceso de minería.

Función *matriz_binaria_csm*

La función *matriz_binaria_csm()* constituye el segundo mecanismo fundamental dentro del proceso de preparación de datos para RuleKit2-CS. A diferencia de *vista_minable_csm*, que genera un conjunto de atributos textuales, *matriz_binaria_csm* produce una estructura altamente dimensional y completamente binaria, en la cual cada columna representa la selección o ausencia de una respuesta específica por parte del encuestado. Este enfoque resulta especialmente útil para métodos de minería orientados al análisis fino de patrones o combinaciones particulares de respuestas, permitiendo identificar contrastes basados en presencia/ausencia directa de atributos.

Al igual que en la vista minable, la función inicia validando si la tabla ya existe y la elimina para garantizar un proceso limpio y reproducible (Listado 3.17).

Listado 3.17. Rutina de reinicio para la matriz binaria de conjuntos de contraste

```
1. IF v_table THEN
2.   EXECUTE 'DROP TABLE public.matriz_binaria_csm';
3. END IF;
```

Este comportamiento idempotente asegura que cada ejecución genere una matriz consistente y sin residuos de ejecuciones previas.

La característica más importante de este método es la creación dinámica de columnas basadas en todas las posibles combinaciones <*pregunta, respuesta*> presentes en la base de datos. Cada combinación se codifica como una variable del tipo (Listado 3.18).

Listado 3.18. Patrón de nomenclatura para columnas dinámicas

```
1. p<idPregunta>_<idRespuesta>
```

Las columnas se construyen de manera programática (Listado 3.19).

Listado 3.19. Construcción dinámica de columnas dimensionales (*One-Hot Encoding*)

```
1. EXECUTE 'ALTER TABLE public.matriz_binaria_csm ADD COLUMN ' || colname || ' text';
```

donde *colname* corresponde a claves como las mostradas en el Listado 3.20.

Listado 3.20. Ejemplos de identificadores de variables binarias generadas

```
1. p14_-3
2. p18_-1
3. p22_2
```

Cada columna toma valores simples:

- "S" si el médico seleccionó esa opción.
- " (vacío) en caso contrario.

Esta representación permite que RuleKit2-CS detecte patrones muy específicos de comportamiento dentro de las respuestas.

La función mantiene el mismo mecanismo estricto de filtrado temprano que en *vista_minable_csm(),* asegurando que solo se incluyan médicos pertenecientes a los hospitales del estudio (Listado 3.21).

Listado 3.21. Filtrado de instancias en la generación de la matriz binaria

```
1. IF hospital_code NOT IN ('H1', 'H11', 'H12') THEN
2.   CONTINUE;
3. END IF;
```

De esta forma, la primera columna del *dataset* representa directamente la variable *objetivo* usada por RuleKit2-CS: el centro hospitalario.

Cada vez que se analiza un médico, se inserta una nueva fila aún vacía que funcionará como registro base para marcar las respuestas seleccionadas (Listado 3.22).

Listado 3.22. Inicialización de registros base en la matriz binaria

```
1. EXECUTE 'INSERT INTO public.matriz_binaria_csm(centro_hospitalario, id_instancia)
2.          VALUES ($1, $2)'
3. USING hospital_code, n_inst;
```

Aquí:

- *centro_hospitalario* es el grupo objetivo.
- *id_instancia* es la clave técnica temporal que garantiza un registro único por fila.

Una vez creada la fila base, la función recorre todas las respuestas del médico y activa las columnas correspondientes. Para cada respuesta, se ejecuta un bloque como el mostrado en el Listado 3.23.

Listado 3.23. Asignación dinámica de valores de presencia en atributos binarios

```
1. EXECUTE 'UPDATE public.matriz_binaria_csm
2.          SET ' || col_bin || ' = ''S''
3.          WHERE id_instancia = $1'
4. USING n_inst;
```

donde:

- *col_bin* es el nombre de la columna binaria previamente creada (por ejemplo, *p18_-2*).

- *"S"* indica que el encuestado seleccionó dicha opción.

Este patrón asegura que cada registro refleje de manera directa, explícita y binaria todas las elecciones del médico.

Las respuestas no seleccionadas no requieren actualización, ya que la estructura por defecto deja las columnas vacías. El diseño evita valores redundantes como *"N"*, simplificando la interpretación posterior por *RuleKit2-CS*.

Una vez completado el llenado de la matriz, se elimina la columna técnica *id_instancia* ya innecesaria para el análisis, dejando únicamente atributos binarios y la columna objetivo (Listado 3.24).

Listado 3.24. Limpieza final de la estructura de la matriz binaria

```
1. ALTER TABLE public.matriz_binaria_csm DROP COLUMN id_instancia;
```

Con esto, la matriz queda lista para ser exportada, analizada y utilizada como insumo del algoritmo de minería.

Funciones de traducción a lenguaje natural

El módulo de análisis comparativo basado en RuleKit2-CS genera reglas simbólicas que, aunque adecuadas para el procesamiento computacional, no son legibles ni interpretables para un usuario no especializado. Para resolver esta limitación, el sistema incorpora dos funciones complementarias de traducción a lenguaje natural: *csm_translate_rule_text_minable_nlg* y *csm_translate_rule_text_matriz_nlg*. Cada una de ellas está diseñada para interpretar y transformar reglas provenientes de distintos conjuntos de datos vista minable y matriz binaria, respectivamente considerando las particularidades sintácticas y semánticas de cada representación. En ambos casos, el objetivo principal es convertir expresiones como *IF atributo = {código} THEN centro_hospitalario = {H1}* en frases descriptivas en español que inician con el prefijo “Si los médicos…”, integran las condiciones relevantes de manera coherente y concluyen con el hospital correspondiente. De esta forma, ambas funciones constituyen el puente

entre el análisis automatizado y la presentación interpretativa de resultados, permitiendo que las reglas descubiertas por RuleKit-2CS puedan ser comprendidas, evaluadas y utilizadas por especialistas clínicos, administradores hospitalarios y responsables del proceso de toma de decisiones.

La función ***csm_translate_rule_text_minable_nlg*** es la encargada de traducir las reglas generadas a partir del conjunto de datos *vista_minable_csm* a oraciones en lenguaje natural. El encabezado de la función (líneas 377–383) mostrado en el Listado 3.25 define el nombre, el tipo de retorno y las variables básicas.

Listado 3.25. Declaración de la función NLG para traducción de reglas (Vista Minable)

```
377: CREATE OR REPLACE FUNCTION public.csm_translate_rule_text_minable_nlg(in_rule text)
378: RETURNS text
379: LANGUAGE plpgsql STABLE AS $$
380: DECLARE
381:   lhs_text  text;
382:   rhs_text  text;
383:   has_then  boolean := false;
```

- *in_rule* es la regla completa en formato textual (tal como la produce RuleKit2-CS).
- *lhs_text* y *rhs_text* almacenan el lado izquierdo (*antecedente*) y derecho (*consecuente*) de la regla.
- *has_then* indica si la regla contiene realmente un *THEN*.

Además, se declaran las estructuras para ir construyendo el texto (Listado 3.26).

Listado 3.26. Análisis sintáctico (*Parsing*) de la regla generada por RuleKit2-CS.

```
385:   mt        text[];                    -- match tuple
386:   frases    text[] := ARRAY[]::text[];-- lista de antecedentes
387:   akey      text;                      -- atributo (descc_columna)
388:   code      text;                      -- código (interpretacion.atributo)
389:   ant       text;                      -- antecedente resuelto
...
393:   hospital  text;
394:   salida    text;
```

- frases es un arreglo que irá acumulando cada condición ya traducida.
- *akey* y *code* representan, respectivamente, el nombre de la columna y el código de respuesta (tal como están en *interpretacion*).
- *hospital* y *salida* se usan al final para construir la oración completa.

El núcleo del procesamiento del lado izquierdo está en el bucle *FOR mt IN* (líneas 402–426) mostradas en el Listado 3.27.

Listado 3.27. Bucle de interpretación semántica de antecedentes

```
402:   -- LHS: attr = {code} / !{code}  → usar SOLO interpretacion (descc_columna=attr, atributo=code)
403:   FOR mt IN
404:     SELECT regexp_matches(lhs_text, '([A-Za-z0-9_]+)\s*=\s*!?\{([^}]+)\}', 'g')
405:   LOOP
406:     akey := lower(mt[1]);
407:     code := lower(mt[2]);
409:     -- Ignorar variables de matriz si aparecieran accidentalmente
410:     IF akey ~* '^[Pp]\d+_-?\d+$' THEN
411:       CONTINUE;
412:     END IF;
414:     ant := NULL;
415:     SELECT i.antecedente
416:       INTO ant
417:     FROM public.interpretacion i
418:     WHERE lower(i.descc_columna) = akey
419:       AND lower(i.atributo) = code
420:       AND COALESCE(NULLIF(i.antecedente,''),'') <> ''
421:     ORDER BY i.id_interpretacion ASC
422:     LIMIT 1;
424:     -- Si no hay antecedente, deja el código para detectar faltantes
425:     frases := frases || COALESCE(ant, code);
426:   END LOOP;
```

Aquí ocurre lo importante:

- El *regex* en la línea 404 detecta expresiones del tipo atributo = *{codigo}* o *atributo = !{codigo}.*

Ejemplos típicos:

- *esc_med_gral = !{c40}*

- per_sol_aut = {*19b*}
- *akey* toma el nombre del atributo (por ejemplo, *esc_med_gral*) y *code* el código de respuesta (*c40*, *19b*).

Se consulta la tabla *interpretacion* (líneas 415–422) usando:

- *descc_columna* = *akey*
- *atributo* = *code* para recuperar el texto descriptivo (*i.antecedente*) asociado a ese par <*columna, valor*>.

Una vez procesado el antecedente, la función genera el inicio estándar de la oración (Listado 3.28).

Listado 3.28. Construcción del prefijo y concatenación de condiciones

```
428:   -- Prefijo exacto
429:   salida := 'Si los médicos ';
430:   IF array_length(frases,1) IS NOT NULL THEN
431:     salida := salida || public.csm_join_y(frases);
432:   ELSE
433:     salida := salida || '(sin condiciones)';
434:   END IF;
```

- *salida* comienza siempre con el prefijo fijo *"Si los médicos ".*
- Si hay al menos una frase en frases, se llama a *csm_join_y* (*frases*) para concatenarlas con comas y la conjunción final *“y”/“e”.*
- Si no se detectaron condiciones, se coloca explícitamente *“(sin condiciones)”.*

Ejemplo de resultado intermedio: *“Si los médicos se formaron en la Facultad de Medicina y consideran adecuado que el médico tratante solicite autopsias”.*

El bloque final se encarga del lado derecho de la regla (Listado 3.29).

- Se extrae el atributo y el código del consecuente (*rhs_attr, rhs_code).*
- Si el atributo es *centro_hospitalario*, se convierte el código (*H1, H11, H12*) al nombre completo del hospital con *csm_hospital_name*.

La frase final queda de la forma: *“Si los médicos …; entonces son del Hospital Regional Río Blanco.”*

Si no hay *THEN*, se devuelve solo el antecedente con punto final.

Listado 3.29. Generación de la conclusión y ensamblaje final de la oración

```
436:   -- RHS
437:   IF has_then THEN
438:     -- alias 'mm' para evitar ambigüedad
439:     SELECT mm[1], mm[3]
440:       INTO rhs_attr, rhs_code
441:     FROM regexp_matches(COALESCE(rhs_text,''), '([A-Za-z0-9_]+)\s*=\s*(!)?\{([^}]+)\}', 'g') AS
mm
442:     LIMIT 1;
444:     IF rhs_attr IS NOT NULL AND lower(rhs_attr) = 'centro_hospitalario' THEN
445:       hospital := public.csm_hospital_name(rhs_code);
446:       RETURN salida || '; entonces son del ' || hospital || '.';
447:     END IF;
449:     RETURN salida || '; entonces ' || btrim(rhs_text) || '.';
450:   ELSE
451:     RETURN salida || '.';
452:   END IF;
```

La función ***csm_translate_rule_text_matriz_nlg*** cumple el mismo objetivo de *NLG*, pero sobre reglas generadas a partir del conjunto de datos *matriz_binaria_csm*, donde los atributos son variables binarias del tipo *pN_M* que indican si una respuesta fue seleccionada o no.

Antes del traductor de matriz se define un *helper* especializado *csm_interpretar_bin* (Listado 3.30).

- *in_var* es un identificador como *p14_-3*, que representa una combinación *<pregunta, respuesta>.*
- La función busca en *interpretacion.atributo* ese código y devuelve el antecedente asociado.
- Si no encuentra texto, regresa el propio código.

Listado 3.30. Función auxiliar para decodificación de variables binarias

```
460: CREATE OR REPLACE FUNCTION public.csm_interpretar_bin(in_var text)
461: RETURNS text
462: LANGUAGE plpgsql STABLE AS $$
463: DECLARE v text := lower(btrim(in_var));
464:         t text;
465: BEGIN
466:   SELECT i.antecedente
467:     INTO t
468:   FROM public.interpretacion i
469:   WHERE lower(i.atributo) = v
470:     AND COALESCE(NULLIF(i.antecedente,''),'') <> ''
471:   LIMIT 1;
473:   RETURN COALESCE(t, v);
474: END $$;
```

Esta función se usa para “decodificar” las variables binarias del lado izquierdo de la regla de matriz.

El traductor de matriz inicia en la línea 479 (Listado 3.31).

Listado 3.31. Declaración de la función NLG para reglas basadas en patrones binarios

```
479: CREATE OR REPLACE FUNCTION public.csm_translate_rule_text_matriz_nlg(in_rule text)
480: RETURNS text
481: LANGUAGE plpgsql STABLE AS $$
482: DECLARE
483:   lhs_text  text;
484:   rhs_text  text;
485:   has_then  boolean := false;
487:   mt        text[];
488:   frases    text[] := ARRAY[]::text[];
489:   var_key   text;      -- pN_M
490:   neg_flag  text;      -- '!' si aplica
491:   texto     text;
...
495:   hospital  text;
496:   salida    text;
```

De nuevo se separan *lhs_text* y *rhs_text,* y se acumulan frases en *frases*.

var_key es la variable binaria (*pN_M)* y *neg_flag* indica si la condición es negada (!).

La separación *IF … THEN …* es análoga a la versión minable (Listado 3.32).

Listado 3.32. Separación de estructura antecedente-consecuente en reglas de matriz

```
499:   SELECT (mx)[1], (mx)[2]
500:     INTO lhs_text, rhs_text
501:   FROM regexp_matches(in_rule, '(?is)^\s*IF\s*(.*?)\s*THEN\s*(.*)$') AS mx;
503:   has_then := rhs_text IS NOT NULL;
504:   IF lhs_text IS NULL THEN
505:     lhs_text := in_rule;  -- si no trae THEN, todo es LHS
506:   END IF;
```

El bucle responsable de traducir el lado izquierdo de la regla de matriz se encuentra en las líneas 508–522 (Listado 3.33).

Listado 3.33. Lógica de traducción de condiciones binarias (Afirmación/Negación)

```
508:   -- LHS: pN_M = {S} / !{S} / S / !S
509:   FOR mt IN
510:     SELECT regexp_matches(lhs_text, '([Pp]\d+_-?\d+)\s*=\s*(!)?\{?\s*S\s*\}?', 'g')
511:   LOOP
512:     var_key  := lower(mt[1]);   -- 'pN_M'
513:     neg_flag := mt[2];          -- '!' o NULL
515:     texto := public.csm_interpretar_bin(var_key);  -- antecedente, o 'pN_M' si falta
517:     IF neg_flag IS NOT NULL AND neg_flag <> '' THEN
518:       frases := frases || ('no ' || texto);
519:     ELSE
520:       frases := frases || texto;
521:     END IF;
522:   END LOOP;
```

El patrón de la línea 510 detecta expresiones del tipo:

- *p14_-3 = {S}*
- *p14_-3 = !{S}*
- con variantes de espacios y llaves.

var_key identifica la variable (*pN_M*) y se pasa a *csm_interpretar_bin*, que la traduce mediante la tabla interpretación.
Si la condición está negada (*neg_flag = '!'*), la función antepone *"no "* al texto interpretado, construyendo frases como:

- *"no consideran que las autopsias originan demandas"*
- *"no consideran que las autopsias originan discrepancias con el diagnóstico clínico"*

Cada condición traducida se agrega al arreglo *frases*.

El prefijo es idéntico al de la función minable (Listado 3.34).

Listado 3.34. Construcción del prefijo para oraciones basadas en matriz

```
524:   -- Prefijo exacto
525:   salida := 'Si los médicos ';
526:   IF array_length(frases,1) IS NOT NULL THEN
527:     salida := salida || public.csm_join_y(frases);
528:   ELSE
529:     salida := salida || '(sin condiciones)';
530:   END IF;
```

Se reutiliza el mismo patrón de presentación: *"Si los médicos …"* + lista de condiciones unida por *csm_join_y*.
El tratamiento del lado derecho vuelve a ser equivalente al de la función minable (Listado 3.35).

Listado 3.35. Generación de la conclusión para reglas de matriz binaria

```
532:   -- RHS
533:   IF has_then THEN
534:     -- alias 'mm' (evitar ambigüedad)
535:     SELECT mm[1], mm[3]
536:       INTO rhs_attr, rhs_code
537:     FROM regexp_matches(rhs_text, '([A-Za-z0-9_]+)\s*=\s*(!)?\{([^}]+)\}', 'i') AS mm
538:     LIMIT 1;
540:     IF rhs_attr IS NOT NULL AND lower(rhs_attr) = 'centro_hospitalario' THEN
541:       hospital := public.csm_hospital_name(rhs_code);
542:       RETURN salida || '; entonces son del ' || hospital || '.';
543:     END IF;
545:     RETURN salida || '; entonces ' || btrim(rhs_text) || '.';
```

Detecta si el consecuente se refiere a *centro_hospitalario*.
Si es así, usa *csm_hospital_name* para traducir el código del hospital y genera frases como: *Si los médicos …; entonces son del Hospital General de San Juan Bautista Tuxtepec.*

En términos funcionales, ambas funciones cumplen el mismo objetivo general traducir reglas simbólicas a enunciados en lenguaje natural, pero operan sobre representaciones de datos distintas. La función *csm_translate_rule_text_minable_nlg* trabaja sobre reglas construidas a partir de los atributos textuales definidos en *vista_minable_csm*. Para ello, interpreta cada condición utilizando pares (*columna, código*) y los mapea a frases en español mediante la tabla *interpretacion*, a partir de los campos *descc_columna* y *atributo*. Adicionalmente, esta función descarta explícitamente cualquier variable de la forma *pN_M*, ya que dichas claves pertenecen al espacio binario de la matriz y no forman parte del modelo semántico de la vista minable.

Por su parte, la función *csm_translate_rule_text_matriz_nlg* se aplica sobre reglas generadas a partir del conjunto de datos binario *matriz_binaria_csm.* En este caso, el antecedente está formado por variables binarias del tipo *pN_M*, cuya semántica se recupera a través de la función auxiliar *csm_interpretar_bin*, que enlaza cada identificador con su descripción correspondiente en la tabla *interpretacion*. Sobre esta base, la función traduce la presencia (S) o la negación (!S) de cada variable en enunciados afirmativos o negativos, respetando la polaridad de la condición original.

Ambas funciones comparten una estructura de salida homogénea: generan oraciones que inician con el prefijo *“Si los médicos …”* y concluyen con una expresión que identifica el hospital implicado en la regla. No obstante, difieren en la forma de resolver el antecedente: mientras *csm_translate_rule_text_minable_nlg* opera sobre columnas semánticas explícitas de la vista minable, *csm_translate_rule_text_matriz_nlg* trabaja sobre códigos binarios derivados de la matriz, especializados en representar la presencia o ausencia de respuestas específicas.

3.4.5 Implementación de lógica del módulo en Java y ejecución de Rulekit2-CS

En este apartado se documenta la lógica de aplicación implementada en el controlador del módulo de minería de conjuntos de contraste, desarrollado en Java y expuesto como *managed bean* de sesión. Este controlador constituye el componente central que integra la arquitectura de la aplicación con el motor *RuleKit2-CS*, ya que orquesta, de manera secuencial, la selección del conjunto de datos, la preparación del experimento, la generación de archivos de configuración y de datos en formato *ARFF*, la ejecución del algoritmo como proceso externo, la carga y estructuración del archivo *CSV* resultante y, finalmente, la invocación del subsistema de generación de lenguaje natural implementado en PostgreSQL. La descripción se presenta de forma estrictamente técnica y se apoya en fragmentos de código con numeración de líneas, de manera que resulte posible referenciar con precisión la implementación dentro del archivo *conjuntosContrasteControler.java.**

Descripción general del controlador y atributos principales

La clase *conjuntosContrasteControler* se declara como un *managed bean* de sesión y extiende la clase *base Model*, implementando además la interfaz *Serializable*. En las líneas 59 a 88 (Listado 3.36) se definen los atributos que representan tanto los parámetros de configuración del experimento como las estructuras internas necesarias para almacenar los resultados y metadatos de ejecución.

En este fragmento se aprecia que el controlador mantiene explícitamente el estado de los parámetros de minería *corpusSeleccionado*, *nivelAnalisis*, *iteraciones*, *numeroReglas* y *medidaCalidad,* así como colecciones para almacenar la estructura tabular derivada del archivo CSV, las reglas ya interpretadas y diversos modelos de gráficos. Estas estructuras permiten que la lógica de aplicación encapsule todos los elementos necesarios para ejecutar y analizar un experimento de minería de conjuntos de contraste sin depender de componentes externos adicionales.

Listado 3.36. Declaración del controlador y atributos principales

```
59 @ManagedBean(name = "oCCJB")
60 @SessionScoped
61 public class conjuntosContrasteControler extends Model implements Serializable {
62
63     private String corpusSeleccionado; // 'c' o 'd'
64     private int nivelAnalisis; // 1=Exploratorio, 2=Refinado, 3=Específico
65     private int iteraciones;   // MAX_PASSES_COUNT
66     private int numeroReglas;  // MAX_RULE_COUNT
67     private String notas;
68     private String medidaCalidad; // Lift, WRAcc, Correlation, Klosgen, RelativeRisk
69
70     private List<centroHospitalario> lista_hospitales;
71     private List<datosGrafica> reglasPDFA;
72     private List<ReglasSG> reglas2;
73     // Resultados CSV dinámicos
74     private List<String> csvHeaders;
75     private List<java.util.Map<String, String>> csvRows;
76     private List<ReglaTabla> tablaReglas;
77     private List<String> visibleCsvHeaders;
78     private String reglaHeaderName;
79     // Traducciones de reglas (para vista especialista)
80     private List<ReglaInterpretada> reglasInterpretadas;
81
82     private BarChartModel barModel;
83     // Modelos de gráficos para análisis visual
84     private BarChartModel barModelProb;      // series p, P
85     private BarChartModel barModelSP;        // series Soporte, Peso (crudo)
86     private BarChartModel barModelSPNorm;    // series Soporte, Peso normalizados (%)
87     private BarChartModel barModelMedida;    // serie medida de calidad si existe
88     // Modelos para métricas derivadas (supp, negSupp, dSupp, precision, dPrecision, redundancy)
```

De forma complementaria, la clase incorpora atributos específicos relacionados con la configuración de conexión a la base de datos y con las rutas de archivos generados durante la ejecución de RuleKit2-CS, tal como se observa en las líneas 86 a 115 (Listado 3.37).

Estos campos permiten registrar, para cada ejecución, el *dataset* utilizado, el nivel de análisis, las rutas a los archivos de configuración *XML* y *ARFF*, el nombre del archivo

CSV generado, así como la bitácora textual de la ejecución registrada en *logEjecucion*. Ello proporciona un soporte claro para la trazabilidad del proceso analítico desde la perspectiva de la aplicación.

Listado 3.37. Atributos de conexión y metadatos de experimento

```
 86     private BarChartModel barModelSupp;
 87     private BarChartModel barModelNegSupp;
 88     private BarChartModel barModelDSupp;
 89     private BarChartModel barModelPrecision;
 90     private BarChartModel barModelDPrecision;
 91     private BarChartModel barModelRedundancy;
 92     private int chartsBuiltForCount;
 93     private String alturaGrafica;
 94 
 95     // Parámetros de conexión
 96     private String sUrl;
 97     private String sUsr;
 98     private String sPwd;
 99 
100     // Rutas y banderas de estado del modelo y ejecución
101     private String rutaConfigGenerada;
102     private String rutaArffGenerado;
103     private boolean modeloListo;
104     private boolean rulekitEjecutado;
105 
106     // Metadatos del último experimento
107     private String lastDatasetFolder;
108     private String lastNivelStr;
109     private String lastExperimentFolder;
110     private String outDirRelativeWeb;
111     private String resolvedOutDirRelativeWeb;
112     private String lastCsvName;
113     private String lastMdlName;
114     private String lastReportName;
115     private String logEjecucion;
```

Inicialización del módulo y carga de parámetros de conexión

La inicialización del controlador se realiza mediante el método anotado con *@PostConstruct*, en el cual se fijan valores por defecto para los parámetros de minería

y se cargan los parámetros de conexión a *PostgreSQL* desde el contexto de inicialización del contenedor mostrado en el Listado 3.38.

Listado 3.38. Inicialización de parámetros del módulo y configuración de conexión

```
197     @PostConstruct
198     public void init() {
199         // Defaults requeridos por la vista
200         this.corpusSeleccionado = null;
201         this.nivelAnalisis = 1; // Exploratorio
202         this.iteraciones = 1;   // min 1, max 5
203         this.numeroReglas = 1;  // min 1, max 100
204         this.notas = "";
205         this.medidaCalidad = "Lift";
206 
207         this.lista_hospitales = new ArrayList<>();
208         this.reglasPDFA = new ArrayList<>();
209         this.reglas2 = new ArrayList<>();
210         this.barModel = new BarChartModel();
211         this.alturaGrafica = "350px";
212         this.csvHeaders = new ArrayList<>();
213         this.csvRows = new ArrayList<>();
214         this.visibleCsvHeaders = new ArrayList<>();
215         this.reglaHeaderName = null;
216         this.tablaReglas = new ArrayList<>();
217         this.reglasInterpretadas = new ArrayList<>();
218         this.barModelProb = null;
219         this.barModelSP = null;
220         this.barModelSPNorm = null;
221         this.barModelMedida = null;
222         this.chartsBuiltForCount = 0;
223 
224         // Cargar parámetros de conexión
225       Map<String, String> params =
FacesContext.getCurrentInstance().getExternalContext().getInitParameterMap();
226         sUrl = params.get("urldb");
227         sUsr = params.get("usrdb");
228         sPwd = params.get("pwddb");
229         this.modeloListo = false;
230         this.rulekitEjecutado = false;
231     }
```

En este método se inicializa el estado operativo mínimo del módulo, estableciendo el nivel de análisis exploratorio como valor por defecto, una iteración inicial y un número mínimo de reglas, además de instanciar las colecciones internas a estructuras vacías. A su vez, se extraen las credenciales de conexión desde los parámetros de inicialización del contexto *JSF*, evitando codificar directamente los valores dentro del código fuente, lo cual es coherente con buenas prácticas de despliegue y seguridad en aplicaciones empresariales.

Método de orquestación de la minería de conjuntos de contraste

El flujo de ejecución principal del módulo web se concentra en el método *consultarResultados*, el cual encapsula las tres operaciones esenciales de la lógica: preparación del experimento, ejecución de RuleKit2-CS y carga de las traducciones de las reglas (Listado 3.39).

Listado 3.39. Orquestación del flujo de minería en el módulo

```
241      // Consulta con el mismo flujo que "Crear": prepara y ejecuta, pero muestra sólo reglas interpretadas
242     public void consultarResultados() {
243         try {
244             ensureDefaultDatasetC();
245             // 1) Preparar experimento con parámetros actuales
246             prepararExperimento();
247             // 2) Ejecutar algoritmo (carga CSV al terminar)
248             ejecutarRulekit();
249             // 3) Generar traducciones para mostrar en la vista Consultar
250             ensureTraduccionesCargadas();
251         } catch (Exception ex) {
252             // Solo log, sin mensajes visuales
253             System.out.println("[CC] consultarResultados error: " + ex.getMessage());
254         }
255     }
```

Este método se limita a coordinar la secuencia de operaciones, delegando la responsabilidad de cada fase a métodos especializados. La llamada previa a *ensureDefaultDatasetC()* garantiza que el atributo *corpusSeleccionado* tenga un valor

válido (por defecto el conjunto C) antes de construir el experimento, de manera que la lógica no dependa de supuestos implícitos de la interfaz.

Preparación del experimento y generación de archivos de configuración

La preparación del experimento se implementa en el método *prepararExperimento* (Listado 3.40), el cual realiza tres tareas principales: selección del *dataset* y nivel de análisis, construcción dinámica del archivo *XML* de configuración a partir de los modelos base y generación del archivo de datos en formato *ARFF* a partir de las vistas *vista_minable_csm* o *matriz_binaria_csm*. El primer bloque del método determina el *dataset* y el nivel de análisis que se aplicarán en función de los parámetros establecidos en el controlador.

Listado 3.40. Selección del *dataset*, vista y nivel analítico

```
933     public void prepararExperimento() {
934         try {
935             // 1) Determinar dataset (vista vs matriz) y nivel
936              String datasetFolder = ("c".equalsIgnoreCase(corpusSeleccionado)) ? "vista_minable"
: "matriz_binaria";
937             String datasetView = ("c".equalsIgnoreCase(corpusSeleccionado)) ? "vista_minable_csm"
: "matriz_binaria_csm";
938            String nivel = (nivelAnalisis == 1) ? "exploratorio" : (nivelAnalisis == 2 ? "refinado"
: "especifico");
939              System.out.println("[CC] Preparando experimento... + ", iteraciones=" + iteraciones
+ ", reglas=" + numeroReglas);
```

En este fragmento se establece la relación directa entre el valor de *corpusSeleccionado* y las vistas de base de datos diseñadas específicamente para la minería de conjuntos de contraste. Cuando el valor es *c*, el módulo opera sobre *vista_minable_csm*, mientras que, para otros valores, se utiliza *matriz_binaria_csm*. El atributo *nivelAnalisis* se mapea a los modos exploratorio, refinado y específico que disponen cada uno de un modelo *XML* base diferenciado.

Una vez seleccionado el entorno del experimento, el método procede a cargar el archivo *XML* plantilla correspondiente y a realizar los reemplazos sobre los *placeholders* que

parametrizan los modelos base. En los modelos *XML* se utilizan marcadores del tipo *{{MAX_PASSES_COUNT}}, {{MAX_RULE_COUNT}}* y *{{MEASURE}},* lo que permite reutilizar los mismos archivos para múltiples configuraciones de análisis sin duplicar contenido. El siguiente Listado 3.41 muestra esta lógica.

Listado 3.41. Reemplazo de *placeholders* en el modelo XML y configuración de parámetros

```
957             String xml = leerArchivo(template);
958             xml = xml.replace("{{MAX_PASSES_COUNT}}", Integer.toString(iteraciones))
959                      .replace("{{MAX_RULE_COUNT}}", Integer.toString(numeroReglas))
960                      .replace("{{MEASURE}}", medidaCalidad);
961             // Ajustar out_directory e in_file acorde al dataset/nivel y archivos generados
962               String experimentFolder = (medidaCalidad != nul...edidaCalidad : "Lift") + "_" +
iteraciones + "_" + numeroReglas;
963             this.lastDatasetFolder = datasetFolder;
964             this.lastNivelStr = nivel;
965             this.lastExperimentFolder = experimentFolder;
966                  String relativeOut = "./resultados/" + datasetFolder + "/" + nivel + "/" +
experimentFolder + "/";
967             String arffFileName = datasetFolder.equals("vista_minable") ? "vista_minable.arff" :
"matriz_binaria.arff";
968             // Reemplazos tolerantes
```

En estas líneas, la plantilla *XML* leída desde disco se transforma en una configuración concreta mediante la sustitución de los *placeholders* con los valores que el usuario ha definido en la sesión: la métrica de calidad seleccionada, el número máximo de iteraciones del algoritmo y el número máximo de reglas que deben inducirse. Adicionalmente, se construye un identificador de carpeta de experimento basado en la combinación *medidaCalidad_iteraciones_numeroReglas*, lo cual permite organizar los resultados en una jerarquía de directorios que refleja explícitamente la parametrización de cada ejecución. Posteriormente, el código (no mostrado en su totalidad aquí) actualiza en el *XML* los elementos <*out_directory*> y <*in_file*> (Listado 3.42). para ajustarlos a las rutas y nombres de archivo que se generarán durante el proceso.

Finalmente, el método materializa el archivo *XML* resultante y genera el archivo *ARFF* a partir de la vista seleccionada, utilizando *Weka* como herramienta de conversión.

Listado 3.42. Persistencia de la configuración y generación del archivo ARFF

```
969                 xml = xml.replaceAll("<out_directory>.*?</out_d...ectory>", "<out_directory>" +
relativeOut + "</out_directory>");
970                   xml = xml.replaceAll("<in_file>.*?</in_file>", "<in_file>" + arffFileName +
"</in_file>");
971
972             File cfgOut = new File(outDir, nivel + ".xml");
973             guardarArchivo(cfgOut, xml);
974             this.rutaConfigGenerada = cfgOut.getAbsolutePath();
975
976             // 4) Generar ARFF desde la vista correspondiente
977             String sql = "select * from " + datasetView + ";";
978             System.out.println("[CC] Extrayendo instancias con SQL: " + sql);
979             Instances data = obtenerInstancias(sql);
980             File arffOut = new File(outDir, arffFileName);
981             guardarARFF(data, arffOut);
982             this.rutaArffGenerado = arffOut.getAbsolutePath();
```

En la línea 974 se observa que la ruta absoluta del archivo *XML* generado se almacena en *rutaConfigGenerada*, y que la conversión de la vista *SQL* a formato *ARFF* se delega al método auxiliar *obtenerInstancias*, que abstrae el uso de *InstanceQuery* para recuperar instancias directamente desde la base de datos. De manera coherente, el archivo *ARFF* se guarda en el mismo directorio de configuración, y su ruta queda registrada en *rutaArffGenerado*. Ello permite que la fase de ejecución pueda localizar tanto el modelo de datos como el archivo de configuración sin necesidad de recomputar rutas.

El método concluye reiniciando las colecciones de resultados y las banderas de estado asociadas a la ejecución anterior, asegurando que cada preparación de experimento sea independiente de ejecuciones previas.

Ejecución de RuleKit2-CS desde la aplicación

La ejecución del motor de minería se lleva a cabo en el método *ejecutarRulekit*, que resuelve la ubicación del archivo *JAR* de RuleKit2-CS, configura el proceso externo y captura la salida estándar del mismo como bitácora de ejecución. El siguiente Listado 3.43 muestra la estructura fundamental de este método.

Listado 3.43. Ejecución del motor RuleKit2-CS y captura de bitácora

```
1017     public void ejecutarRulekit() {
1018         try {
1019             if (this.rutaConfigGenerada == null) {
1020                 prepararExperimento();
1021             }
1022 
1023             // Resolver jar
1024                                                                 String      baseWeb      =
FacesContext.getCurrentInstance().getExternalContext().getRealPath("/");
1025             File jar1 = new File(baseWeb, "WEB-INF/lib/rulekit-2.1.24-all.jar");
1026             File jar2 = new File("web/WEB-INF/lib/rulekit-2.1.24-all.jar");
1027             File jar = jar1.exists() ? jar1 : jar2;
1028 
1029             System.out.println("[CC] Ejecutando Rulekit con JAR: " + jar.getAbsolutePath());
1030             System.out.println("[CC] Usando configuración: " + this.rutaConfigGenerada);
1031 
1032             // Directorio de trabajo al del config para rutas relativas (ARFF y resultados)
1033             File workDir = new File(this.rutaConfigGenerada).getParentFile();
1034             List<String> cmd = new ArrayList<>();
1035             cmd.add("java");
1036             cmd.add("-jar");
1037             cmd.add(jar.getAbsolutePath());
1038             // Rulekit espera el archivo XML como único argumento
1039             cmd.add(new File(this.rutaConfigGenerada).getName());
1040 
1041             ProcessBuilder pb = new ProcessBuilder(cmd);
1042             pb.directory(workDir);
1043             pb.redirectErrorStream(true);
1044             Process p = pb.start();
1045 
1046             StringBuilder sb = new StringBuilder();
1047             try (InputStream is = p.getInputStream();
```

Listado 3.43. Ejecución del motor RuleKit2-CS y captura de bitácora - Continuación

```
1048                            BufferedReader br = new BufferedReader(new InputStreamReader(is,
StandardCharsets.UTF_8))) {
1049                 String line;
1050                 while ((line = br.readLine()) != null) {
1051                     System.out.println("[Rulekit] " + line);
1052                     sb.append(line).append('\n');
1053                 }
1054             }
1055             int code = p.waitFor();
1056             this.logEjecucion = sb.toString();
```

Este código garantiza que el proceso externo se ejecute en el directorio padre del archivo de configuración *XML*, lo que facilita el uso de rutas relativas dentro de dicho archivo (por ejemplo, hacia el archivo *ARFF* y hacia los directorios de resultados). La salida estándar del proceso se acumula en *logEjecucion*, proporcionando un registro textual de la ejecución que puede utilizarse para diagnóstico o auditoría.

La fase final del método valida el código de terminación del proceso y, en caso de éxito, desencadena la carga del archivo *CSV* de resultados mostrado en el Listado 3.44.

Listado 3.44. Manejo del código de terminación y carga del CSV

```
1057             if (code != 0) {
1058                 System.out.println("[CC] Rulekit terminó con código " + code);
1059                 FacesContext.getCurrentInstance().addMessage(null,
1060                         new FacesMessage(FacesMessage.SEVERITY_ERROR, "Rulekit falló", "Código:
" + code));
1061                 this.rulekitEjecutado = false;
1062             } else {
1063                 System.out.println("[CC] Rulekit finalizado OK");
1064                 // cargar resultados CSV en memoria para mostrar en tabla
1065                 try {
1066                     cargarResultadosCSV();
1067                 } catch (Exception e) {
1068                     System.out.println("[CC] Error cargando CSV: " + e.getMessage());
1069                 }
1070                 this.rulekitEjecutado = true;
```

Listado 3.44. Manejo del código de terminación y carga del CSV - Continuación

```
1071                FacesContext.getCurrentInstance().addMessage(null,
1072                            new FacesMessage(FacesMessage.SEVERITY_INFO, "Rulekit", "Ejecución
finalizada correctamente."));
1073            }
1074        } catch (Exception ex) {
1075            System.out.println("[CC] Error ejecutando Rulekit: " + ex.getMessage());
1076            ex.printStackTrace();
1077            FacesContext.getCurrentInstance().addMessage(null,
1078                    new FacesMessage(FacesMessage.SEVERITY_ERROR, "Error ejecutando Rulekit",
ex.getMessage()));
1079        }
1080    }
```

Este bloque asocia directamente el resultado de la ejecución del motor con la actualización de la estructura interna del controlador, reflejada en la bandera *rulekitEjecutado*, y con la carga de los resultados para su posterior interpretación y análisis.

La obtención de instancias desde la base de datos, utilizada en la preparación del experimento, se encapsula en el método auxiliar *obtenerInstancias*, que abstrae el uso de *InstanceQuery* de *Weka* (Listado 3.45).

Listado 3.45. Recuperación de instancias desde la base de datos mediante Weka

```
1082    private Instances obtenerInstancias(String sql) throws Exception {
1083        InstanceQuery query = new InstanceQuery();
1084        query.setDatabaseURL(sUrl);
1085        query.setUsername(sUsr);
1086        query.setPassword(sPwd);
1087        query.setQuery(sql);
1088        return query.retrieveInstances();
1089    }
```

Este método establece de forma explícita la cadena de conexión, credenciales y consulta a ejecutar, delegando en *Weka* la construcción del objeto *Instances*, lo que permite una transición directa hacia el formato *ARFF* sin pasos intermedios adicionales.

Carga y estructuración del archivo CSV de resultados

Tras la ejecución de *RuleKit2-CS,* los resultados se vuelcan a un archivo *CSV* que contiene, para cada regla inducida, sus medidas de soporte, pesos, probabilidades y otras métricas. El método *cargarResultadosCSV* mostrado en el Listado 4.36 se encarga de localizar este archivo dentro de la jerarquía de resultados, interpretar su contenido y materializarlo en colecciones internas que servirán tanto para la representación tabular como para análisis cuantitativos posteriores.

Listado 3.46. Localización e interpretación del archivo CSV

```
1116     private void cargarResultadosCSV() throws IOException {
1117         ensureResolvedOutDir();
1118        String baseWeb = FacesContext.getCurrentInstance().getExternalContext().getRealPath("/");
1119        String fname = (lastCsvName != null ? lastCsvName ...vista_minable") ? "vista_minable.csv"
: "matriz_binaria.csv"));
1120         File csvFile = new File(baseWeb, resolvedOutDirRelativeWeb + fname);
1121         System.out.println("[CC] Cargando CSV: " + csvFile.getAbsolutePath());
1122         if (!csvFile.exists()) {
1123             System.out.println("[CC] CSV no encontrado");
1124             return;
1125         }
1126         List<String> headers = new ArrayList<>();
1127         List<java.util.Map<String, String>> rows = new ArrayList<>();
1128        try (BufferedReader br = new BufferedReader(new Inp...eader(new FileInputStream(csvFile),
StandardCharsets.UTF_8))) {
1129             String first = br.readLine();
1130             if (first == null) return;
1131             String delimiter = first.contains(";") && !first.contains(",") ? ";" : ",";
1132             for (String h : first.split(delimiter, -1)) {
1133                 headers.add(h.trim());
1134             }
```

Posteriormente, el método recorre las filas de datos (Listado 3.47)., construyendo para cada una un mapa *clave–valor* que asocia cada encabezado con su valor correspondiente. Esta estructura de datos se almacena en *csvRows* y constituye la base para la construcción de objetos de dominio que representan reglas individuales y sus métricas derivadas.

Listado 3.47. Construcción de estructura tabular interna y modelos enriquecidos

```
1184          int num = 1;
1185          for (java.util.Map<String, String> r : rows) {
1186              ReglaTabla t = new ReglaTabla();
1187              t.numero = num++;
1188              t.regla = (keyRegla != null ? r.getOrDefault(keyRegla, "") : "");
1189              t.pLower = (keyp != null ? r.getOrDefault(keyp, "") : "");
1190              t.pUpper = (keyP != null ? r.getOrDefault(keyP, "") : "");
1191              t.soporte = (keySoporte != null ? r.getOrDefault(keySoporte, "") : "");
1192              t.peso = (keyPeso != null ? r.getOrDefault(keyPeso, "") : "");
1193              this.tablaReglas.add(t);
1194          }
1195
1196          // Construir el modelo enriquecido con métricas derivadas
1197          List<ReglaEnriquecida> enr = new ArrayList<>();
1198          int idx = 1;
1199          for (java.util.Map<String, String> r : rows) {
1200              ReglaEnriquecida e = new ReglaEnriquecida();
1201              e.numero = idx++;
1202              e.rule = (keyRegla != null ? r.getOrDefault(keyRegla, "") : "");
1203
1204              Double pv = (keyp != null ? parseDoubleSafe(r.getOrDefault(keyp, null)) : null);
1205              Double Pv = (keyP != null ? parseDoubleSafe(r.getOrDefault(keyP, null)) : null);
...
1210              if (pv != null && Pv != null && Pv != 0.0) {
1211                  e.supp = pv / Pv;
1212              } else {
1213                   e.supp = parseDoubleSafe(keySoporte != null ? r.getOrDefault(keySoporte, null)
: null);
1214              }
1215
1216              // neg_supp = n / N
1217              if (nv != null && Nv != null && Nv != 0.0) {
1218                  e.negSupp = nv / Nv;
1219              } else {
1220                  e.negSupp = null;
1221              }
1222
1223              // d_supp = supp - neg_supp
1224              if (e.supp != null && e.negSupp != null) {
1225                  e.dSupp = e.supp - e.negSupp;
```

Este bloque evidencia que el controlador no solo conserva los valores directamente generados por *RuleKit2-CS,* sino que también calcula métricas derivadas tales como el soporte negativo y la diferencia de soportes entre clases, que son útiles para el análisis cuantitativo y la construcción de gráficos comparativos. De este modo, la lógica de aplicación complementa los resultados crudos del motor de reglas con medidas adicionales de interés analítico.

Invocación del subsistema de generación de lenguaje natural

La última etapa del flujo corresponde a la interpretación en lenguaje natural de las reglas inducidas. Esta funcionalidad se implementa en el método *ensureTraduccionesCargadas* (Listado 3.48), que consume las filas de *csvRows* y delega en funciones almacenadas de *PostgreSQL* la traducción de cada regla en texto estructurado.

Listado 3.48. Integración con el subsistema de generación de lenguaje natural

```
1623     public void ensureTraduccionesCargadas() {
1624         try {
1625             if (this.csvRows == null || this.csvRows.isEmpty()) return;
1626             // Preferir la columna exacta 'Rule' como solicitaste
1627             String keyRule = matchHeaderExact(this.csvHeade...rs != null ? this.csvHeaders : new
ArrayList<String>(), "Rule");
1628             if (keyRule == null) keyRule = this.reglaHeaderName; // fallback si no existe header
exacto
1629             if (keyRule == null) return;
1630 
1631             String dataset = getDatasetFolder();
1632             boolean esMinable = "vista_minable".equalsIgnoreCase(dataset);
1633             String sql = esMinable
1634                     ? "SELECT public.csm_translate_rule_text_minable_nlg(CAST(? AS text))"
1635                     : "SELECT public.csm_translate_rule_text_matriz_nlg(CAST(? AS text))";
1636 
1637             // Cargar driver si está configurado (tolerante)
1638             try {
1639                                                                 Map<String,    String>    params    =
FacesContext.getCurrentInstance().getExternalContext().getInitParameterMap();
1640                 String drv = params != null ? params.get("driverdb") : null;
1641                 if (drv != null && !drv.isEmpty()) { Class.forName(drv); }
1642             } catch (Throwable ignore) {}
```

Listado 3.48. Integración con el subsistema de generación de lenguaje natural - Continuación

```
1643
1644            java.sql.Connection cn = null;
1645            java.sql.PreparedStatement ps = null;
1646            java.sql.ResultSet rs = null;
1647            List<ReglaInterpretada> out = new ArrayList<>();
1648            try {
1649                cn = java.sql.DriverManager.getConnection(sUrl, sUsr, sPwd);
1650                ps = cn.prepareStatement(sql);
1651                int idx = 1;
1652                for (java.util.Map<String, String> r : this.csvRows) {
...
1657                    String sanitized = (raw != null) ? raw.replace("\"", "") : "";
1658                    ps.setString(1, sanitized);
1659                    rs = ps.executeQuery();
1660                    String traducida = raw;
1661                    System.out.println(rs.toString());
1662                    if (rs.next()) {
1663                        String t = rs.getString(1);
1664                            System.out.println("REGLAS INTERPRETADAS");
1665                            System.out.println(t);
1666                        if (t != null) traducida = t;
1667                    }
1668                    try { if (rs != null) { rs.close(); rs = null; } } catch (Exception ignore)
{}
1669                    ReglaInterpretada ri = new ReglaInterpretada();
1670                    ri.numero = idx++;
1671                    ri.texto = traducida;
1672                    out.add(ri);
1673                }
1674                this.reglasInterpretadas = out;
```

El método selecciona de manera dinámica la función de traducción apropiada en función del *dataset* utilizado (*vista_minable_csm o matriz_binaria_csm*), aplicando *csm_translate_rule_text_minable_nlg* o *csm_translate_rule_text_matriz_nlg* respectivamente. Cada regla en crudo se sanea previamente para eliminar comillas superfluas, se envía como argumento a la función *PL/pgSQL* y se recupera el texto interpretado, que se encapsula en instancias de *ReglaInterpretada*. Esta colección

constituye la representación final de las reglas desde la perspectiva semántica, y es la que se pone a disposición de la capa de presentación para su consumo por parte del usuario especialista.

En síntesis, la implementación de la lógica del módulo en *conjuntosContrasteControler.java* integra de manera estructurada todos los pasos necesarios para ejecutar el proceso de minería de conjuntos de contraste utilizando *RuleKit2-CS* desde la propia aplicación. El controlador se responsabiliza de parametrizar los modelos *XML* base mediante *placeholders* que capturan la configuración concreta del experimento, de generar el conjunto de datos en formato *ARFF* a partir de vistas especializadas de la base de datos, de ejecutar el motor analítico como proceso externo y de estructurar el archivo *CSV* resultante en modelos de dominio enriquecidos. Finalmente, la capa de integración con el subsistema de generación de lenguaje natural en *PostgreSQL* completa el flujo, proporcionando descripciones textuales de las reglas que resultan adecuadas para su análisis e interpretación por parte de usuarios expertos. De este modo, el controlador se erige como el núcleo lógico del módulo, articulando la arquitectura de la aplicación con el motor de minería y con las capacidades de interpretación automática de reglas.

3.4.6 Implementación de la vista con JSF y PrimeFaces

En este apartado se documenta la capa de presentación asociada al módulo de minería de conjuntos de contraste, implementada mediante *Facelets*, *JSF* y *PrimeFaces*, y acoplada al controlador descrito previamente. La vista se encuentra distribuida en varias páginas *XHTML* especializadas *modeloCC.xhtml*, *resultadoConjuntosContraste.xhtml*, *analisisVisualCC.xhtml*, *reglasSinInterpretar.xhtml* y *reglasInterpretadas.xhtml* que se apoyan en plantillas comunes y en el *managed bean oCCJB* (instancia de la clase *conjuntosContrasteControler*). La lógica de interfaz se orienta a cubrir tres objetivos principales: (i) permitir la parametrización y ejecución del experimento, (ii) facilitar la consulta de resultados y su interpretación textual, y (iii) habilitar un análisis visual detallado de las métricas de cada regla inducida.

En los apartados siguientes se describen los elementos más relevantes de cada vista, haciendo énfasis en la forma en que se enlazan los componentes de *PrimeFaces* con las propiedades y métodos expuestos por el controlador.

Estructura general y uso de plantillas

Todas las vistas del módulo se construyen sobre plantillas *Facelets* que definen la estructura común menú superior, secciones laterales y área de contenido, garantizando consistencia visual en todo el sistema. Por ejemplo, la página de resultados *resultadoConjuntosContraste.xhtml* (Listado 3.49) declara en sus primeras líneas el uso de la plantilla base mediante *ui:composition*.

Listado 3.49. Estructura base de la vista de consulta

```
1 <?xml version='1.0' encoding='UTF-8' ?>
 2      <!DOCTYPE      html      PUBLIC      "-//W3C//DTD      XHTML      1.0      Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
 3 <html xmlns="http://www.w3.org/1999/xhtml"
 4       xmlns:ui="http://xmlns.jcp.org/jsf/facelets"
 5       xmlns:h="http://xmlns.jcp.org/jsf/html"
 6       xmlns:p="http://primefaces.org/ui"
 7       xmlns:f="http://xmlns.jcp.org/jsf/core">
 ...
15     <h:body>
16         <ui:composition template="./../Template/plantillaBase.xhtml">
17             <ui:define name="top">
18                 <p:menubar style="float: right; width: 40%">
```

La sección *top* define el menú de navegación y, en la sección *content*, se inyecta el contenido específico de la vista del módulo. Este patrón se repite en las vistas *modeloCC.xhtml*, *analisisVisualCC.xhtml*, *reglasSinInterpretar.xhtml* y *reglasInterpretadas.xhtml*, donde se reutilizan las plantillas *plantillaBase.xhtml* y *plantillaBaseFull.xhtml* para ajustar el diseño según la densidad de contenido requerido (por ejemplo, para las tablas y gráficos de gran tamaño en pantalla completa).

Vista de modelado y ejecución del experimento

Listado 3.50. Formulario de configuración y ejecución del modelo

```
101                          <p:panelGrid  id="dgReglas"  columns="2"  style="vertical-align:  top"
styleClass="noBordes">
102                     <h:outputScript name="./js/leer.js"/>
103
104                     <table id="demo"></table>
105
106                     <p:column width="20%">
107                         <p:accordionPanel  id="apAlgoritmos">
108                             <p:tab title="Conjuntos de Contraste" >
109                                 <h:form id="formAlg2">
110                                     <p:growl id="growl" showDetail="true" life="10000"/>
111                                     <h:panelGrid columns="2" cellpadding="10">
112
114                                       <p:outputLabel for="conjunto" value="Dataset" style="font-
weight: bolder"/>
115                                          <p:selectOneRadio id="conjunto" layout="pageDirection"
columns="1"
116                                                                  value="#{oCCJB.corpusSeleccionado}"
required="true">
117                                           <f:selectItem itemLabel="Conjunto C" itemValue="c" />
118                                           <f:selectItem itemLabel="Conjunto D" itemValue="d" />
119                                            <p:ajax listener="#{oCCJB.onParametrosChanged}"
120                                                        update=":apAlgoritmos:formAlg2:btnCrear,
:apAlgoritmos:formAlg2:btnEjecutar2"/>
121                                       </p:selectOneRadio>
122
123                                           <p:outputLabel for="nivel" value="Nivel de Análisis"
style="font-weight: bolder"/>
124                                       <p:selectOneMenu id="nivel" value="#{oCCJB.nivelAnalisis}"
required="true" style="width:180px">
125                                            <f:selectItem itemLabel="Exploratorio" itemValue="1"
/>
126                                           <f:selectItem itemLabel="Refinado" itemValue="2" />
127                                          <f:selectItem itemLabel="Específico" itemValue="3" />
```

La página *modeloCC.xhtml* está orientada al usuario especialista y concentra la parametrización y ejecución del experimento. En ella se ubica un panel principal con un

componente *p:accordionPanel*, donde una de las pestañas contiene el formulario específico para minería de conjuntos de contraste. El Listado 3.50 muestra la estructura central de este formulario.

En este código, los componentes *p:selectOneRadio* y *p:selectOneMenu* se enlazan directamente con las propiedades *corpusSeleccionado* y *nivelAnalisis* del *managed bean oCCJB*, de manera que la selección del conjunto de datos (C o D) y del nivel de análisis (exploratorio, refinado o específico) impacta de inmediato los parámetros que el controlador utilizará al construir el experimento. El componente *p:ajax* asociado a ambos controles invoca el método *onParametrosChanged* del controlador y actualiza dinámicamente el estado de los botones de acción, lo que evita inconsistencias en la interfaz cuando los parámetros se modifican.

Listado 3.51. Botones de creación del modelo y ejecución de RuleKit2-CS

```
156                                                 <p:commandButton id="btnCrear" value="Crear modelo"
157                                                                  action="#{oCCJB.prepararExperimento()}"
158                                                                    onstart="PF('statusDialog20').show()"
159                                                                  onsuccess="PF('statusDialog20').hide()"
160                                                                       update=":frmReglas:resultadosAlgo,
:frmReglas:dgReglas2, :frmReglas:growl"
161                                                                      styleClass="btn-primary btn-lg btn-
wide" style="margin:0 10px;"/>
162                                                   <p:commandButton id="btnEjecutar2" value="Ejecutar
Rulekit" icon="ui-icon-play"
163                                                                    onclick=" PF('statusDialog20').show();"
164                                                                  oncomplete="PF('statusDialog20').hide();"
165                                                                         update=":frmReglas:resultadosAlgo,
:frmReglas:dgReglas2, :frmReglas:growl"
166                                                                      process="@this"
167                                                                      disabled="#{!oCCJB.modeloListo}"
168                                                                        styleClass="#{oCCJB.modeloListo ?
'btn-success btn-lg btn-wide' : 'btn-lg btn-wide'}"
169                                                                      style="margin:0 10px;"/>
```

En la parte inferior del mismo formulario se encuentran los botones que disparan la preparación del experimento y la ejecución del motor RuleKit2-CS. Estos botones se enlazan con los métodos *prepararExperimento()* y *ejecutarRulekit()* definidos en el controlador (Listado 3.51).

Finalmente, en la misma vista se proporciona acceso directo a las páginas de resultados, tanto para reglas sin interpretar como para reglas traducidas, abriendo cada una en una nueva ventana cuando la ejecución de RuleKit2-CS se ha completado (Listado 3.52).

Listado 3.52. Acceso desde la vista de modelado a las vistas de resultados

```
186                             <p:panelGrid id="dgReglas2" columns="2" styleClass="noBordes"
187                                      columnClasses="ui-grid-col-4, ui-grid-col-8" layout="grid">
188                                 <!-- Columna izquierda: botones en vertical -->
189                                 <p:column>
190                                  <p:panelGrid columns="1" styleClass="noBordes" style="width:100%;">
191                                          <p:column>
192                                                  <p:commandButton type="button" value="Reglas Sin
Interpretar" icon="ui-icon-newwin"
193
onclick="window.open('#{request.contextPath}/faces/Views/especialista/reglasSinInterpretar.xhtml','_
blank')"
194                                                                          styleClass="btn-secondary btn-lg"
disabled="#{!oCCJB.rulekitEjecutado}"/>
195                                          </p:column>
196                                          <p:column>
197                                              <p:commandButton type="button" value="Interpretación
De Las Reglas" icon="ui-icon-newwin"
198
onclick="window.open('#{request.contextPath}/faces/Views/especialista/reglasInterpretadas.xhtml','_b
lank')"
199                                                                          styleClass="btn-secondary btn-lg"
disabled="#{!oCCJB.rulekitEjecutado}"/>
```

Este diseño refuerza la separación entre la etapa de construcción/ejecución del modelo (en *modeloCC.xhtml*) y las vistas de análisis de resultados, a la vez que preserva la coherencia de navegación para el usuario especialista.

Vista de consulta integrada de resultados

La vista *resultadoConjuntosContraste.xhtml* está pensada para la consulta directa de los resultados de minería y su interpretación por parte del usuario, sin necesidad de reconfigurar el experimento. Esta página registra tres eventos *preRenderView* (Listado 3.53) que garantizan que siempre haya un *dataset* por defecto, que el archivo *CSV* se encuentre cargado y que las traducciones en lenguaje natural estén disponibles antes de renderizar el contenido.

Listado 3.53. Eventos de inicialización de la vista de consulta

```
65              <ui:define name="content">
66
67                  <f:event type="preRenderView" listener="#{oCCJB.ensureDefaultDatasetC}"/>
68                  <f:event type="preRenderView" listener="#{oCCJB.ensureCSVLoaded}"/>
69                  <f:event type="preRenderView" listener="#{oCCJB.ensureTraduccionesCargadas}"/>
```

Los métodos *ensureDefaultDatasetC*, *ensureCSVLoaded* y *ensureTraduccionesCargadas* del controlador se encargan, respectivamente, de seleccionar el conjunto C como valor inicial en *corpusSeleccionado* cuando no existe sesión previa, de forzar la carga del archivo CSV correspondiente a la última ejecución de *RuleKit2-CS* y de poblar la lista *reglasInterpretadas* mediante invocaciones a las funciones de generación de lenguaje natural en *PostgreSQL*. De esta forma, la vista puede concentrarse en presentar la información, delegando la consistencia de estado al controlador.

El panel principal se organiza en un *p:panelGrid* con dos columnas. La primera contiene un formulario reducido que permite al usuario ajustar el *dataset* y volver a ejecutar el flujo de consulta, mientras que la segunda muestra las reglas interpretadas en un componente *p:dataTable* totalmente interactivo (Listado 3.54).

Listado 3.54. Formulario de consulta de parámetros en la vista de resultados

```
72                          <p:column>
73                              <h:form id="frmConsulta">
```

Listado 3.54. Formulario de consulta de parámetros en la vista de resultados - Continuación

```
74                              <p:fieldset legend="Minería de Conjuntos de Contraste - Consultar"
class="siBordes">
75                               <h:panelGrid columns="2" cellpadding="10">
76
77                                      <p:outputLabel for="conjunto" value="Dataset" style="font-
weight: bolder"/>
78                                          <p:selectOneRadio id="conjunto" layout="pageDirection"
columns="1"
79                                                            value="#{oCCJB.corpusSeleccionado}"
required="true">
```

La segunda columna incorpora el formulario *frmReglas*, que alberga el panel de interpretación de las reglas y la tabla interactiva *dtReglasInterp*, enlazada a la colección *reglasInterpretadas* expuesta por el controlador (Listado 3.55).

Listado 3.55. Tabla de reglas interpretadas en la vista integrada de resultados

```
127                         <p:column>
128                             <h:form id="frmReglas">
129                                 <br/>
130                                    <p:panel header="Interpretación de las reglas" style="margin-
bottom:20px">
131                                 <p:panelGrid columns="1" styleClass="noBordes" style="width:100%;
margin-bottom:6px;">
132                                         <h:panelGroup layout="block" styleClass="inline-filter">
133                                                           <p:inputText id="searchNumInterpC"
style="display:none;width:0"
134                                                                placeholder="N° Regla" styleClass="flt-
num"
135                                                                                onkeypress="return /[0-
9]/.test(String.fromCharCode(event.which||event.keyCode))"/>
136                                          <p:commandButton type="button" value="Buscar" icon="ui-
icon-search"
137                                                                               onclick="(function(){var
t=PF('dtInterpC');var            v=prompt('N°            Regla');var            $f=t.jq.find('.flt-
num');$f.val(v);t.filter();})();"/>
138                                                  <p:commandButton type="button" value="Limpiar"
icon="ui-icon-close"
```

Listado 3.55. Tabla de reglas interpretadas en la vista integrada de resultados - Continuación

```
139                                                                    onclick="(function(){var
t=PF('dtInterpC'); t.jq.find('.flt-num').val(''); t.clearFilters();})();"/>
140                                        </h:panelGroup>
141                                    </p:panelGrid>
142                                                        <p:dataTable  id="dtReglasInterpC"
value="#{oCCJB.reglasInterpretadas}"
143                                                    var="row" rows="20" widgetVar="dtInterpC"
filterEvent="enter" filterDelay="0"
144                                                       style="margin-bottom:20px; width:100%"
rowIndexVar="rowIndex"
145                                                         scrollable="true" scrollHeight="600"
responsiveLayout="scroll" stickyHeader="true"
```

El bloque anterior muestra cómo la tabla permite filtrar por número de regla mediante una columna oculta (*flt-num*), cuya entrada se rellena desde una ventana emergente invocada por *JavaScript* al presionar el botón *"Buscar"*. El atributo *value="#{oCCJB.reglasInterpretadas}"* enlaza cada fila de la tabla con las instancias de *ReglaInterpretada* generadas en el controlador, las cuales contienen el número correlativo de regla y el texto en lenguaje natural construido a partir de la salida de *RuleKit2-CS* y del módulo de *PL/pgSQL*.

Vista de análisis visual de métricas

La vista *analisisVisualCC.xhtml* se centra en ofrecer un análisis visual detallado de las métricas derivadas para cada regla (soporte, soporte negativo, diferencia de soportes, precisión, diferencia de precisión y redundancia). Para ello, se apoya en múltiples componentes *p:chart* de *PrimeFaces*, cada uno vinculado a un modelo de gráfico *BarChartModel* preparado en el controlador (Listado 3.56).

Cada componente *p:chart* consume un modelo de barras construido en el controlador a partir de la colección de *ReglaEnriquecida*, que incluye las métricas derivadas para cada regla. De esta manera, la vista ofrece una representación gráfica de la distribución de las

métricas por número de regla, facilitando la identificación de patrones, contrastes y reglas con comportamientos extremos (por ejemplo, valores altos de *dSupp* o de *dPrecision*).

Listado 3.56. Configuración de gráficos de barras para métricas de reglas

```
21                    <p:panelGrid columns="1" styleClass="noBordes"
22                                 columnClasses="ui-grid-col-12" layout="grid" style="width:100%;">
23                        <p:column>
24                             <p:fieldset legend="supp por regla (X: N° Regla)" toggleable="true"
collapsed="false">
25                                  <p:chart id="chtSupp" type="bar" model="#{oCCJB.barModelSupp}"
style="width:100%; height: 420px" />
26                             </p:fieldset>
27                        </p:column>
28                        <p:column>
29                           <p:fieldset legend="negSupp por regla (X: N° Regla)" toggleable="true"
collapsed="false">
30                             <p:chart id="chtNegSupp" type="bar" model="#{oCCJB.barModelNegSupp}"
style="width:100%; height: 420px" />
31                             </p:fieldset>
32                        </p:column>
33                        <p:column>
34                             <p:fieldset legend="dSupp por regla (X: N° Regla)" toggleable="true"
collapsed="false">
35                                 <p:chart id="chtDSupp" type="bar" model="#{oCCJB.barModelDSupp}"
style="width:100%; height: 420px" />
...
43                        <p:column>
44                         <p:fieldset legend="dPrecision por regla (X: N° Regla)" toggleable="true"
collapsed="false">
45                        <p:chart id="chtDPrecision" type="bar" model="#{oCCJB.barModelDPrecision}"
style="width:100%; height: 420px" />
46                             </p:fieldset>
47                        </p:column>
48                        <p:column>
49                         <p:fieldset legend="redundancy por regla (X: N° Regla)" toggleable="true"
collapsed="false">
50                        <p:chart id="chtRedundancy" type="bar" model="#{oCCJB.barModelRedundancy}"
style="width:100%; height: 420px" />
```

Vistas especializadas de reglas sin interpretar e interpretadas

Las páginas *reglasSinInterpretar.xhtml* y *reglasInterpretadas.xhtml* proporcionan vistas especializadas que permiten revisar en detalle, respectivamente, las reglas en su forma cruda (según el formato de salida de RuleKit2-CS) y las reglas interpretadas en lenguaje natural. Ambas vistas reutilizan el mismo diseño de plantillas y menú superior, pero difieren en el origen de los datos y en las columnas presentadas.

En *reglasSinInterpretar.xhtml*, la tabla *dtReglasSinInterp* se enlaza con una colección de objetos que representan reglas con sus métricas numéricas (soporte, peso, diferencia de soporte, precisión, diferencia de precisión y redundancia). El Listado 3.57 muestra la definición de la tabla.

Listado 3.57. Tabla de reglas sin interpretar

```
31                                    style="margin-bottom:20px; width:100%" rowIndexVar="rowIndex"
32                                   scrollable="true" scrollHeight="600" responsiveLayout="scroll"
stickyHeader="true"
33                                 resizableColumns="true" liveResize="true" columnResizeMode="fit"
34                                       rowStyleClass="#{(rowIndex mod 2) eq 0 ? 'alternativa2' :
'alternativa1'}"
35                                     tableStyle="width:100%; table-layout:fixed;">
36                                <f:facet name="header">Reglas sin interpretar</f:facet>
37
38                                <p:column headerText="N° Regla" width="6%" filterMatchMode="equals"
39                                          style="text-align:center; white-space:nowrap;">
40                                    <f:facet name="filter">
41                                            <p:inputText style="display:none;width:0" placeholder="N°
Regla" styleClass="flt-num"/>
42                                    </f:facet>
43                                    <h:outputText value="#{row.numero}" />
44                                </p:column>
45
46                               <p:column headerText="Regla" width="52%" style="white-space: normal;">
```

La tabla aplica estilos alternos por fila para mejorar la legibilidad y utiliza un campo de filtro oculto similar al de la vista integrada, lo que permite localizar rápidamente una regla por su número. A partir de las mismas estructuras, la vista *reglasInterpretadas.xhtml* enlaza la tabla *dtReglasInterp* con la colección *reglasInterpretadas*, mostrando el texto

completo de cada regla en lenguaje natural, incluyendo el contexto hospitalario y las condiciones de contraste derivadas del algoritmo (Listado 3.58).

Listado 3.58. Tabla de reglas interpretadas

```
21                    <p:panel style="max-width:100%; overflow:hidden;">
22                        <p:panelGrid columns="1" styleClass="noBordes" style="width:100%; margin-
bottom:6px;">
23                            <h:panelGroup layout="block" styleClass="inline-filter">
24                                <p:inputText id="searchNumInterp" style="display:none;width:0"
25                                             placeholder="Nº Regla" styleClass="flt-num"
26                                                                    onkeypress="return  /[0-
9]/.test(String.fromCharCode(event.which||event.keyCode))"/>
27                                <p:commandButton type="button" value="Buscar" icon="ui-icon-search"
...
31                            <p:dataTable id="dtReglasInterp" value="#{oCCJB.reglasInterpretadas}"
var="row" rows="20"
32                                         widgetVar="dtInterp" filterEvent="enter" filterDelay="0"
33                                         style="margin-bottom:20px; width:100%" rowIndexVar="rowIndex"
34                                        scrollable="true" scrollHeight="600" responsiveLayout="scroll"
stickyHeader="true"
35                                resizableColumns="true" liveResize="true" columnResizeMode="expand"
```

En esta vista, cada fila de la tabla corresponde a una instancia de *ReglaInterpretada* y el texto de la regla se presenta en una columna amplia, con estilos CSS específicos para permitir el ajuste del contenido de forma legible incluso cuando las reglas son extensas. Esto complementa la vista de reglas sin interpretar, proporcionando una representación semántica orientada a la interpretación por parte de especialistas clínicos o investigadores.

La implementación de la capa de presentación para el módulo de minería de conjuntos de contraste, distribuida en las vistas *modeloCC.xhtml*, *resultadoConjuntosContraste.xhtml, analisisVisualCC.xhtml, reglasSinInterpretar.xhtml y reglasInterpretadas.xhtml*, materializa de forma coherente los servicios expuestos por el controlador *conjuntosContrasteControler* y por el motor *RuleKit2-CS*. Mediante el uso de componentes de *PrimeFaces* para formularios, tablas y gráficos de barras, las vistas permiten al usuario especialista parametrizar y ejecutar experimentos, consultar

resultados numéricos, explorar métricas derivadas y analizar reglas en lenguaje natural. El empleo sistemático de plantillas *Facelets*, eventos *preRenderView* y *bindings* directos a las propiedades del *managed bean oCCJB* asegura una separación clara de responsabilidades entre la lógica de aplicación y la interfaz, así como una experiencia de usuario consistente y adecuada para la exploración interactiva de conjuntos de contraste.

Capítulo 4 Resultados

El presente capítulo expone los resultados obtenidos en la validación del módulo de minería de conjuntos de contraste desarrollado en esta tesis, utilizando el algoritmo Rulekit2-CS aplicado sobre dos estructuras distintas de datos: el Conjunto C (Vista Minable) y el Conjunto D (Matriz Binaria). Ambos conjuntos forman parte del diseño metodológico del sistema y permiten evaluar el comportamiento del algoritmo bajo representaciones complementarias de la misma información: una perspectiva semántica detallada a nivel atributo (Vista Minable) y una representación codificada, optimizada para el cálculo intensivo (Matriz Binaria). Aunque en esta etapa del documento únicamente se presentan los resultados correspondientes al Conjunto C, la arquitectura del módulo contempla explícitamente la validación cruzada con ambos conjuntos. Este planteamiento garantiza que los patrones obtenidos no dependan de una sola forma de representar los datos, sino que se mantengan estables aun cuando las relaciones entre preguntas y respuestas se convierten en estructuras binarias de mayor abstracción.

Además, antes de exponer los resultados experimentales, este capítulo incorpora la presentación general del módulo de minería de conjuntos de contraste, mostrando su funcionamiento desde la creación del experimento —incluyendo la selección del conjunto de datos, la configuración de parámetros y la definición de la función de calidad— hasta la ejecución del algoritmo Rulekit2-CS, la obtención de reglas sin interpretar y el procesamiento posterior para calcular las métricas derivadas utilizadas en esta investigación. Asimismo, se describe la funcionalidad de visualización y descarga del reporte final, que permite obtener en un solo documento todas las reglas generadas por cada experimento. Esta presentación proporciona al lector el contexto operativo necesario para comprender cómo se producen, estructuran y consolidan los resultados analizados en las siguientes secciones.

Para interpretar adecuadamente los resultados derivados del experimento, fue necesario definir desde el inicio un conjunto de escenarios, premisas metodológicas y casos de estudio que permitieran evaluar las reglas generadas bajo criterios cuantitativos, clínicos e institucionales. Estos elementos funcionan como un marco de referencia para analizar

el comportamiento del algoritmo, comparar los hospitales y validar la coherencia de los patrones identificados.

Los escenarios considerados se construyeron a partir de las diferencias operativas y organizacionales entre los tres hospitales analizados, asumiendo que la disminución de autopsias no es un fenómeno uniforme, sino dependiente de las condiciones propias de cada institución. Desde este enfoque, las reglas se evalúan no solo como patrones estadísticos, sino como reflejos de configuraciones institucionales: trayectorias formativas, percepciones clínicas, experiencia profesional y criterios de solicitud de autopsias.

Las premisas metodológicas definidas para esta etapa del estudio establecen criterios estrictos para aceptar o descartar reglas. Entre ellas destacan la necesidad de consistencia estadística (soporte, diferencia de soporte, precisión y diferencia de precisión), la exclusividad del patrón (soporte negativo cercano a cero), la coherencia semántica una vez traducidas las reglas y la validez clínica evaluada por el experto en patología. Estas premisas permiten garantizar que los patrones identificados por Rulekit2-CS sean simultáneamente robustos, interpretables y aplicables a la práctica médica real.

A partir de estos escenarios y premisas se definió el planteamiento de los casos de estudio, concebidos para evaluar la aplicabilidad del modelo en situaciones que representan configuraciones contrastantes entre hospitales. Aunque inicialmente se planificaron dos casos de estudio principales, el diseño del módulo permite extender esta evaluación a cada una de las funciones de calidad utilizadas. En este sentido, cada función de calidad genera un conjunto independiente de reglas que constituye en sí mismo un caso de estudio adicional. La configuración específica del experimento considera un máximo de veinte reglas por función de calidad para el conjunto de datos C, lo que permite capturar patrones altamente diferenciados incluso en funciones estrictas como *Lift, WRAcc, Correlation y Klosgen.*

La única excepción ocurre en la métrica *RelativeRisk*, donde solo se obtuvieron tres reglas como resultado del nivel de exigencia impuesto por la función. Este comportamiento no constituye un error, sino una consecuencia esperada dada la naturaleza de la métrica, la cual penaliza duramente cualquier relación que no sea extremadamente fuerte y exclusiva. La presencia de únicamente tres reglas bajo *RelativeRisk* ilustra precisamente el nivel de especificidad alcanzado por el modelo.

Finalmente, es importante señalar que, aunque la presentación detallada de este capítulo se centra inicialmente en las reglas generadas para el Conjunto C, estas representan la primera parte del proceso de validación del módulo. En etapas posteriores, las mismas funciones de calidad serán ejecutadas y analizadas sobre el Conjunto D, lo que permitirá evaluar la estabilidad de los patrones cuando los datos son trasformados desde un enfoque semántico hacia una representación numérica abstracta. Esta doble validación es parte fundamental del diseño del módulo de minería y asegura que los contrastes descubiertos sean independientes de la forma de representación de los datos.

4.1 Presentación del módulo de minera de conjuntos de contraste

El módulo de minería de conjuntos de contraste desarrollado en este proyecto constituye la capa funcional donde convergen todos los elementos construidos a lo largo del diseño e implementación del sistema: las vistas estructuradas en PostgreSQL, los mecanismos de preprocesamiento, la lógica ejecutada en Java y la integración final con RuleKit2-CS. Este módulo representa la interfaz operativa mediante la cual el usuario genera, ejecuta y analiza las reglas de contraste derivadas del conjunto de datos utilizado para la validación del algoritmo.

Su propósito es ofrecer un entorno controlado, reproducible y transparente que permita materializar el flujo completo del proceso analítico: seleccionar parámetros, invocar el motor de minería, visualizar la salida generada por RuleKit2-CS, administrar las reglas obtenidas, inspeccionar métricas de calidad (*Lift*, soporte, diferencia de soporte, precisión, diferencia de precisión), y finalmente explorar de forma interactiva la interpretación semántica de los patrones descubiertos. Cada una de estas

funcionalidades se soporta en una arquitectura modular cuidadosamente desacoplada, donde la vista mantiene comunicación directa con los controladores y estos, a su vez, con los servicios especializados para la ejecución del algoritmo y el procesamiento de resultados.

En este apartado se presentan de manera progresiva las pantallas y componentes principales del módulo, explicando su función dentro del flujo de trabajo, su relación con las etapas de la metodología y la forma en que contribuyen a la validación empírica del algoritmo RuleKit2-CS. Para cada imagen se describe su objetivo, la interacción prevista con el usuario, y la manera en que se integra con el *backend* del sistema. Esta presentación no sólo documenta la capa visual, sino que evidencia cómo cada elemento refleja decisiones arquitectónicas y de diseño orientadas a garantizar consistencia, trazabilidad y confiabilidad en la ejecución del proceso de minería de contrastes.

Una vez mostradas las interfaces, se contextualiza su papel dentro del flujo completo del módulo, estableciendo la relación entre la ejecución técnica del algoritmo y la experiencia de usuario proporcionada por la plataforma web.

4.1.1 Pantalla Principal del Sistema

Como se observa en la Figura 4.1, esta vista corresponde a la pantalla principal del sistema, presentada antes de que el usuario acceda a cualquiera de los módulos disponibles. En ella se muestra el menú lateral con las distintas secciones del sistema, junto con información contextual relacionada con el diagnóstico situacional de las autopsias en el H.R.R.B., incluyendo un gráfico histórico que permite visualizar la tendencia registrada a lo largo de los años.

Esta pantalla no forma parte del módulo de minería de conjuntos de contraste; sin embargo, constituye el punto inicial de navegación dentro del sistema. Para ingresar al módulo y a cualquier otra funcionalidad protegida el usuario debe seleccionar la opción “Iniciar sesión” desde el menú lateral, lo que lo redirige a la vista de autenticación correspondiente.



Figura 4.1 Pantalla principal del sistema

4.1.2 Inicio de Sesión

Como se observa en la Figura 4.2, esta pantalla presenta el formulario de autenticación necesario para ingresar al sistema. Su diseño incorpora los campos obligatorios de usuario y contraseña, así como los botones para iniciar sesión o cancelar la operación. La vista permite validar las credenciales contra los registros existentes y asegurar que únicamente usuarios con el permiso de Especialista accedan a los módulos protegidos.

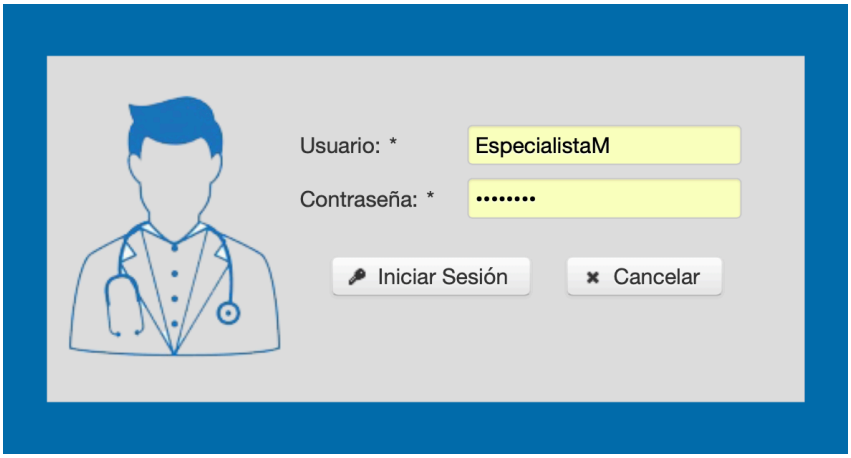


Figura 4.2 Ventana de inicio de sesión

Aunque esta pantalla tampoco pertenece directamente al módulo de minería de conjuntos de contraste, constituye un paso indispensable para garantizar el acceso controlado. Una vez autenticado, el usuario puede navegar hacia el módulo correspondiente y ejecutar las operaciones asociadas al análisis mediante RuleKit2-CS.

4.1.3 Configuración del Modelo para la Ejecución de RuleKit2-CS

Como se observa en la Figura 4.3, esta pantalla constituye la interfaz central de configuración del modelo dentro del módulo de minería de conjuntos de contraste. En esta vista, el usuario define de manera explícita los parámetros necesarios para construir la estructura interna que posteriormente será ejecutada por el algoritmo RuleKit2-CS. Esta fase inicial es fundamental, ya que determina el comportamiento del proceso de minería y las restricciones bajo las cuales se generarán las reglas de contraste.

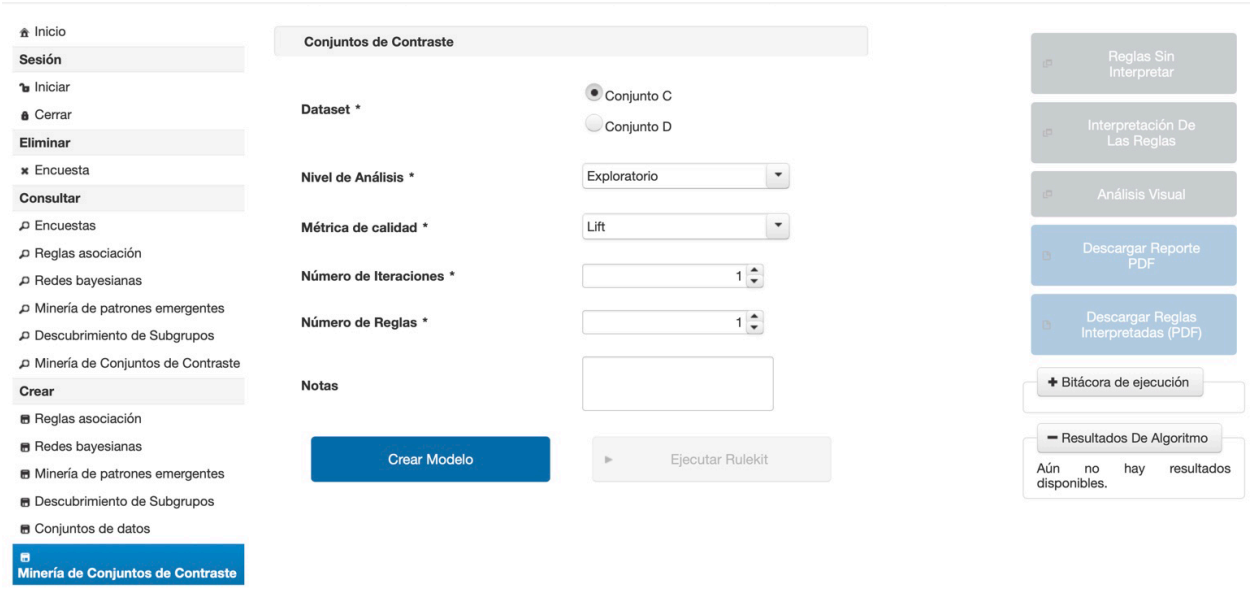


Figura 4.3 Interfaz general del módulo

En primer lugar, el usuario selecciona el *dataset*, pudiendo elegir entre Conjunto C o Conjunto D. Esta elección define directamente la estructura de atributos que se enviará al motor de minería. Conjunto C corresponde a la vista minable enriquecida mientras que Conjunto D corresponde a la matriz binaria. La opción seleccionada cambia el origen del archivo ARFF generado durante la ejecución del módulo.

A continuación, el usuario selecciona el *Nivel de Análisis*, como se observa en la Figura 4.4. Este componente permite elegir entre las tres configuraciones predefinidas implementadas durante el proceso metodológico: Exploratorio, Refinado, y Específico.

Cada uno de estos niveles modifica automáticamente parámetros internos como *min_support, min_support_count, min_difference, min_confidence y min_lift*, alineándose con los valores utilizados durante la validación del módulo.

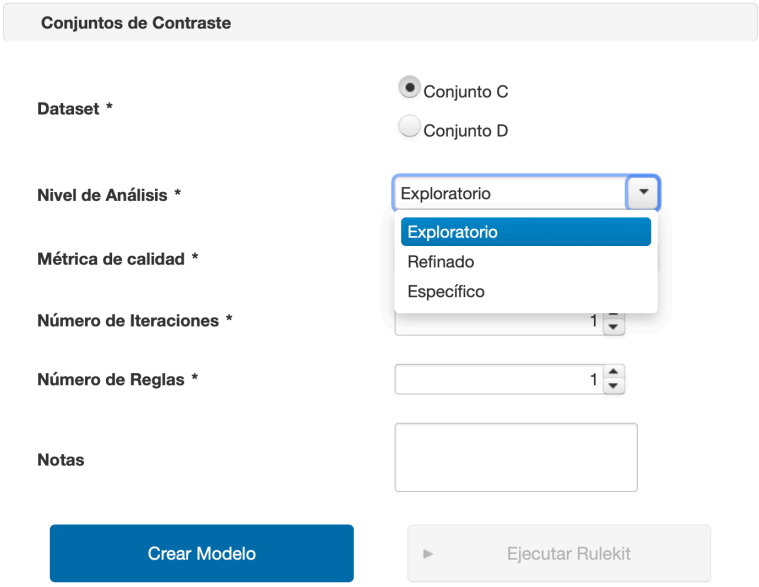


Figura 4.4 Selección del nivel de análisis

Posteriormente, el usuario selecciona la *Métrica de Calidad*, como se aprecia en la Figura 4.5. El sistema integra cinco métricas disponibles en RuleKit2-CS: Lift, WRAcc, Correlation, Klosgen o RelativeRisk.

Para esta tesis, la función de calidad principal utilizada fue *lift*, dado que permite evaluar adecuadamente relaciones discriminativas entre clases mutuamente excluyentes, condición intrínseca a la minería de conjuntos de contraste.

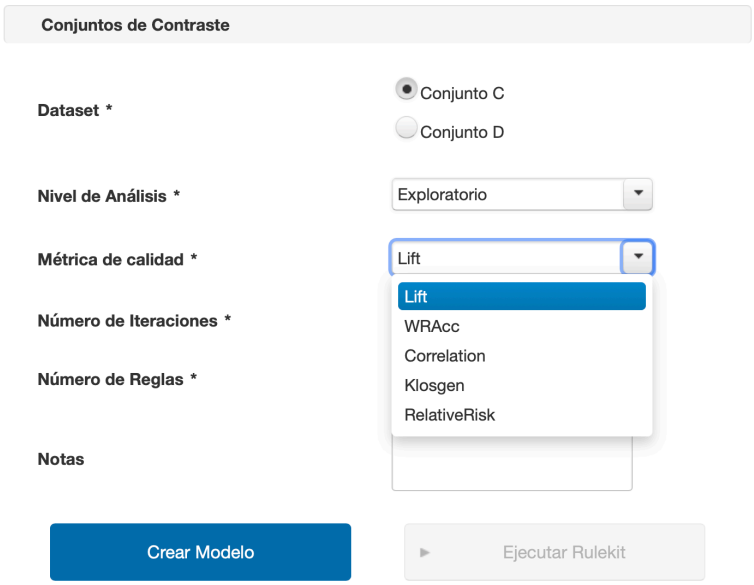


Figura 4.5 Selección de la métrica de calidad

Adicionalmente, los campos *Número de Iteraciones* y *Número de Reglas* permiten establecer los límites del proceso de búsqueda interna del algoritmo. Estos valores repercuten directamente en el espacio explorado y en el número máximo de reglas a generar para cada grupo objetivo.

Finalmente, el usuario registra notas adicionales y generar el modelo mediante el botón “Crear Modelo”. Este botón representa el cierre de la primera fase del proceso, en la cual únicamente se construye el objeto de modelo (DTO) que encapsula todos los parámetros seleccionados por el especialista.

Activación y desactivación dinámica de botones

Como se observa en la Figura 4.3, todos los botones ubicados en la sección derecha *Reglas Sin Interpretar, Interpretación de las Reglas, Análisis Visual, Descargar Reporte PDF, Descargar Reglas Interpretadas* aparecen inicialmente desactivados. Esta condición responde a la lógica de control del módulo, ya que ninguna de estas opciones puede utilizarse antes de ejecutar RuleKit2-CS.

El comportamiento dinámico es el siguiente:

- Antes de crear el modelo:
  a. Todos los botones permanecen deshabilitados. No existe aún un modelo válido.
- Después de crear el modelo (primera fase) como se muestra en la Figura 4.6
  a. Se habilita únicamente el botón "*Ejecutar Rulekit2-CS*", permitiendo iniciar el proceso de minería.
  b. Todos los demás botones permanecen desactivados, ya que no existen todavía reglas ni resultados.
  c. Se muestra mensaje de éxito en la creación del modelo.

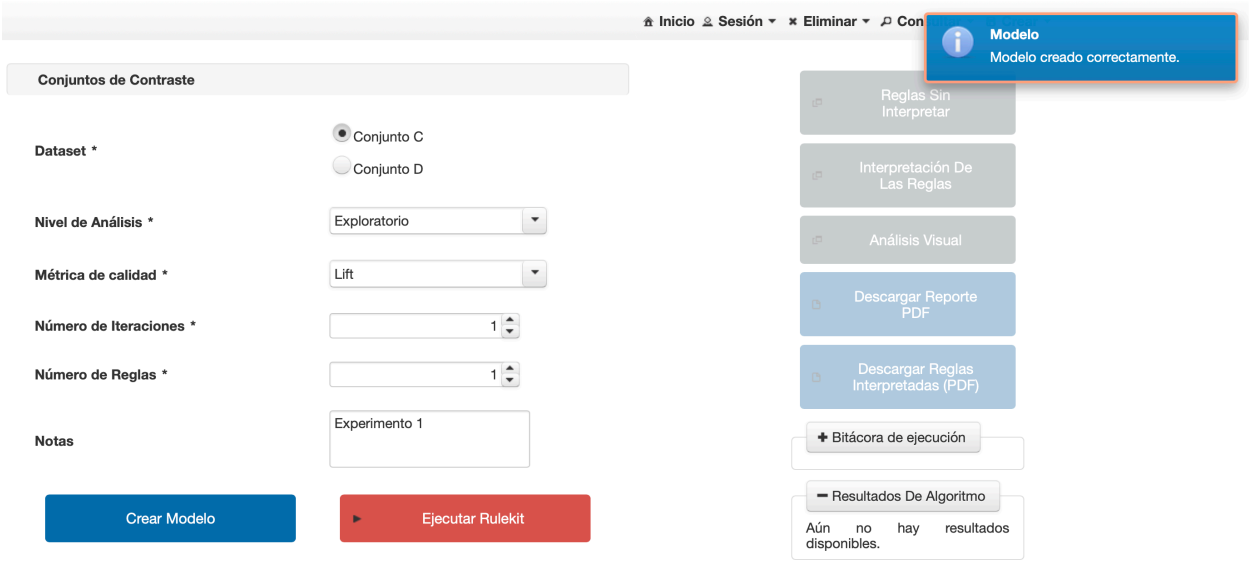


Figura 4.6 Creación del modelo

- Tras ejecutar RuleKit2-CS (segunda fase) mostrado en la Figura 4.7:
  a. Si el proceso termina correctamente:
  b. Se activan todos los botones del lado derecho.
  c. El sistema habilita el acceso a:
    a. Reglas sin interpretar.
    b. Interpretación natural de reglas.
    c. Análisis visual.
    d. Descarga de reporte PDF.

e. Descarga de reglas interpretadas.
f. El bloque *Resultados De Algoritmo* también se actualiza y muestra los datos generados por el motor de minería.

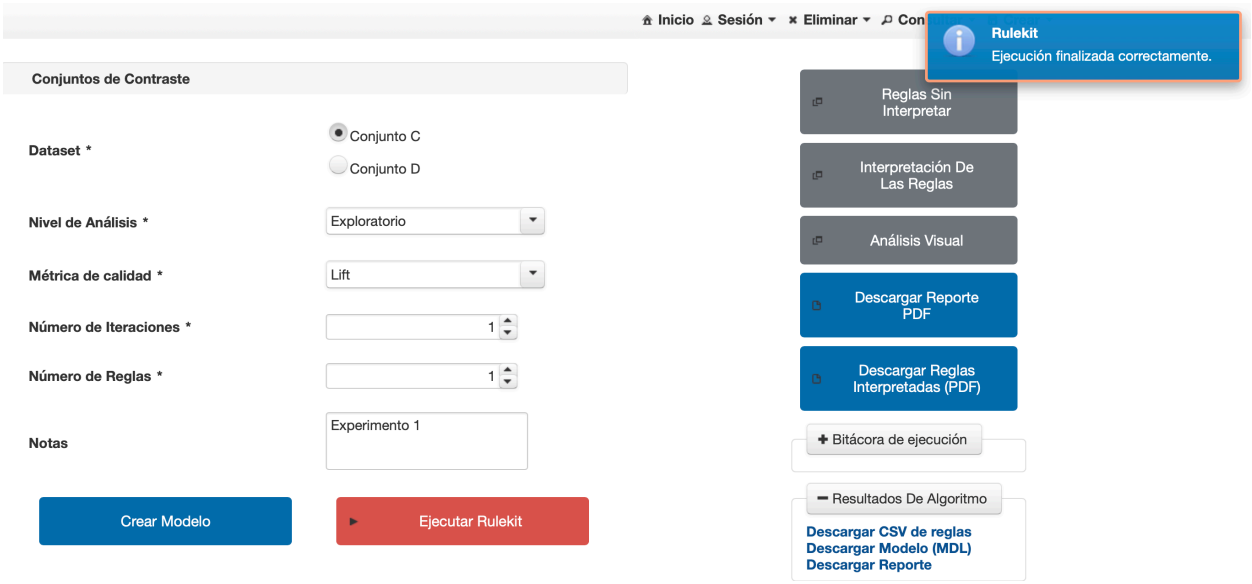


Figura 4.7 Ejecución Rulekit2-CS

- Si el usuario cambia cualquier parámetro del formulario:
  a. El sistema desactiva nuevamente todos los botones,
  b. Invalida el modelo anterior, y
  c. Solicita nuevamente generar el modelo antes de permitir una nueva ejecución de RuleKit2-CS.

Este mecanismo garantiza que toda ejecución se realice bajo un modelo válido y actualizado, evitando inconsistencias.

4.1.4 Reglas Sin Interpretar

Como se observa en la Figura 4.8, esta pantalla presenta la primera salida del proceso de minería de conjuntos de contraste después de ejecutar RuleKit2-CS. En esta etapa, el sistema muestra las reglas sin interpretar, es decir, las reglas en su forma técnica original, acompañadas de las métricas cuantitativas utilizadas para evaluar su relevancia y comportamiento comparativo entre las clases del conjunto de datos.



N° Regl	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF esc_esp = !{c103} AND esc_med_gral = !{c104} THEN centro_hospitalario = {H1}"	1	0.517843	0.482157	0.363673	-0.272653	5
2	"IF categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c125} AND mcc_aut = !{14e} AND mcc_no_aut = !{16m} AND com_sug_op = !{22j} AND no_hosp = !{18h} AND med_aut = !{20h} THEN centro_hospitalario = {H11}"	1	0.321575	0.678425	0.853284	0.706569	10,457
3	"IF esc_med_gral = !{c1} AND mcc_no_aut = !{16p} AND fmr_sol_aut = !{21h} AND esc_esp = !{c6} AND categoria = {c1} AND com_sug_op = !{22p} AND rechazo_fam = !{17z} AND mcc_aut = !{14d} AND no_hosp = !{18h} AND med_aut = !{20f} THEN centro_hospitalario = {H12}"	1	0.444955	0.555045	0.234632	-0.530737	41,101

Figura 4.8 Vista de reglas sin interpretar

Es importante destacar que, aunque el texto de la regla proviene del formato estándar generado por RuleKit2-CS, las métricas presentadas en esta tabla no son devueltas directamente por el algoritmo. En su lugar, estas métricas se calculan dentro de la lógica de la propia aplicación utilizando los datos resultantes de la ejecución, asegurando consistencia con los valores utilizados durante el proceso metodológico. Entre estas métricas se incluyen:

- supp: soporte global de la regla.
- negSupp: soporte del grupo no objetivo.
- dSupp: diferencia de soporte entre clases.
- precision: precisión de la regla.
- dPrecision: diferencia de precisión entre clases.
- redundancy: redundancia de la regla dentro del conjunto generado.

La tabla también contiene un campo de navegación paginada, permitiendo recorrer reglas de manera eficiente cuando el algoritmo genera un conjunto amplio.

Cada regla se muestra en su estructura declarativa, iniciando con la condición “IF … THEN centro_hospitalario = {Hx}”, donde los atributos son los códigos internos correspondientes a los valores categóricos del *dataset*. Esta estructura responde al formato utilizado por RuleKit2-CS, en el cual cada atributo aparece representado por su identificador en las vistas de base de datos (*vista_minable* o *matriz_binaria*). Para su interpretación semántica, se emplean posteriormente las funciones de mapeo incluidas en la fase de NLG del módulo.

Esta vista tiene un propósito estrictamente analítico y corresponde a la segunda fase del flujo del módulo: la ejecución del algoritmo. Solo después de que RuleKit2-CS concluye correctamente, esta tabla se habilita y se presentan las reglas generadas.

Asimismo, cualquier ajuste que el usuario realice a los parámetros iniciales del modelo: *dataset*, nivel de análisis, métrica de calidad, número de reglas o iteraciones provoca que esta sección se desactive nuevamente hasta que se realice una nueva ejecución del algoritmo, garantizando que la información mostrada esté siempre sincronizada con el modelo vigente.

Esta vista constituye el punto de partida para las fases posteriores del módulo, en las cuales las reglas se transforman, interpretan y visualizan mediante análisis comparativos y representaciones gráficas.

4.1.5 Interpretación de las Reglas

Como se observa en la Figura 4.9, esta pantalla corresponde a la sección denominada “Interpretación de las Reglas”, la cual se habilita únicamente después de realizar una ejecución válida del algoritmo RuleKit2-CS. Su propósito es presentar las reglas generadas en un formato legible para el usuario final, transformando la estructura técnica producida por el motor de minería en descripciones en lenguaje natural comprensibles para especialistas clínicos, personal administrativo y tomadores de decisiones.

A diferencia del apartado de *Reglas Sin Interpretar*, donde los atributos aparecen con sus códigos internos provenientes de la vista minable o de la matriz binaria (por ejemplo, *esc_esp = !{c103} o no_hosp = !{18h}),* esta vista muestra la salida ya transformada mediante el conjunto de funciones de interpretación desarrolladas en PostgreSQL. Dichas funciones realizan el mapeo entre: códigos internos de atributos, opciones seleccionadas por los médicos en las encuestas, y valores semánticos legibles correspondientes a preguntas y respuestas.

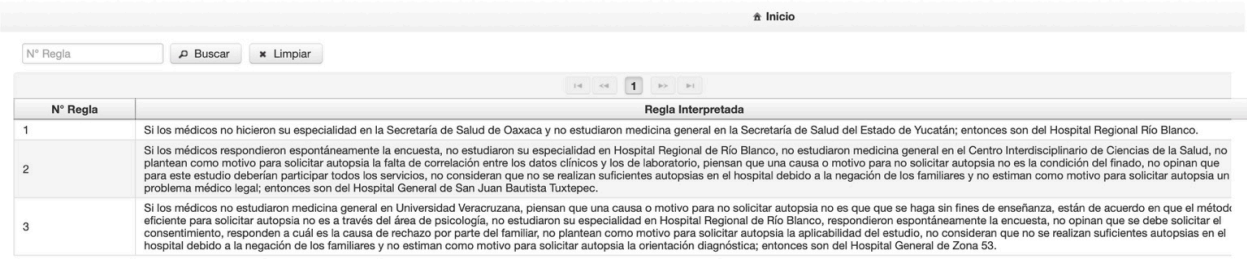


Figura 4.9 Vista de interpretación en lenguaje natural de las reglas generadas

El resultado es un texto fluido, completo y contextualizado, donde cada condición se expresa como un fragmento descriptivo que explica la situación representada por la regla. El objetivo es asegurar que la regla pueda evaluarse por personal especialista sin requerir conocimiento técnico del formato original de RuleKit2-CS.

Cada regla interpretada comienza típicamente con una estructura condicional “Si los médicos…” seguida de las combinaciones de atributos detectados como contrastantes entre los hospitales analizados. La regla finaliza con la identificación del centro hospitalario correspondiente, permitiendo comprender qué combinación de factores distingue al hospital objetivo respecto a los demás.

Esta vista se construye a partir de tres elementos esenciales:

1. La salida base de RuleKit2-CS, la cual indica qué atributos componen cada regla.
2. La capa de interpretación implementada en PostgreSQL, encargada de traducir cada atributo y valor a su forma descriptiva.
3. La estructura de presentación, que compone frases completas siguiendo un patrón lingüístico uniforme para garantizar claridad y coherencia.

La tabla permite paginación, búsqueda por número de regla y limpieza del filtro, facilitando la navegación cuando se generan reglas extensas o cuando su contenido interpretado incluye múltiples condiciones concatenadas.

Esta pantalla constituye un paso clave dentro del módulo, ya que permite convertir conocimiento técnico en conocimiento operable, facilitando la comprensión del comportamiento discriminativo entre hospitales.

4.1.6 Análisis Visual de las Reglas

Como se observa en la Figura 4.10, esta pantalla corresponde a la sección *Análisis Visual*, la cual permite representar gráficamente las métricas cuantitativas asociadas a cada regla generada por RuleKit2-CS. Esta vista constituye un componente clave dentro del módulo, ya que facilita la comparación entre reglas mediante visualizaciones independientes de cada métrica, permitiendo identificar comportamientos contrastantes, tendencias y valores atípicos que no son evidentes únicamente mediante tablas numéricas.

En esta sección se presenta inicialmente el gráfico correspondiente al soporte (*supp*) por regla, mostrado como un histograma donde el eje *X* representa el número de regla y el eje *Y* el valor de soporte. Esta primera gráfica se construye automáticamente al acceder a la vista, sirviendo como referencia inicial del comportamiento global del conjunto de reglas.

Adicionalmente, la pantalla incorpora botones interactivos que permiten visualizar, de manera individual y bajo demanda, cada una de las métricas calculadas por la aplicación. Estos botones se muestran en la parte izquierda y corresponden a:

- negSupp por regla (*X*: Nº Regla)
- dSupp por regla (*X*: Nº Regla)
- precision por regla (*X*: Nº Regla)
- dPrecision por regla (*X*: Nº Regla)
- redundancy por regla (*X*: Nº Regla)

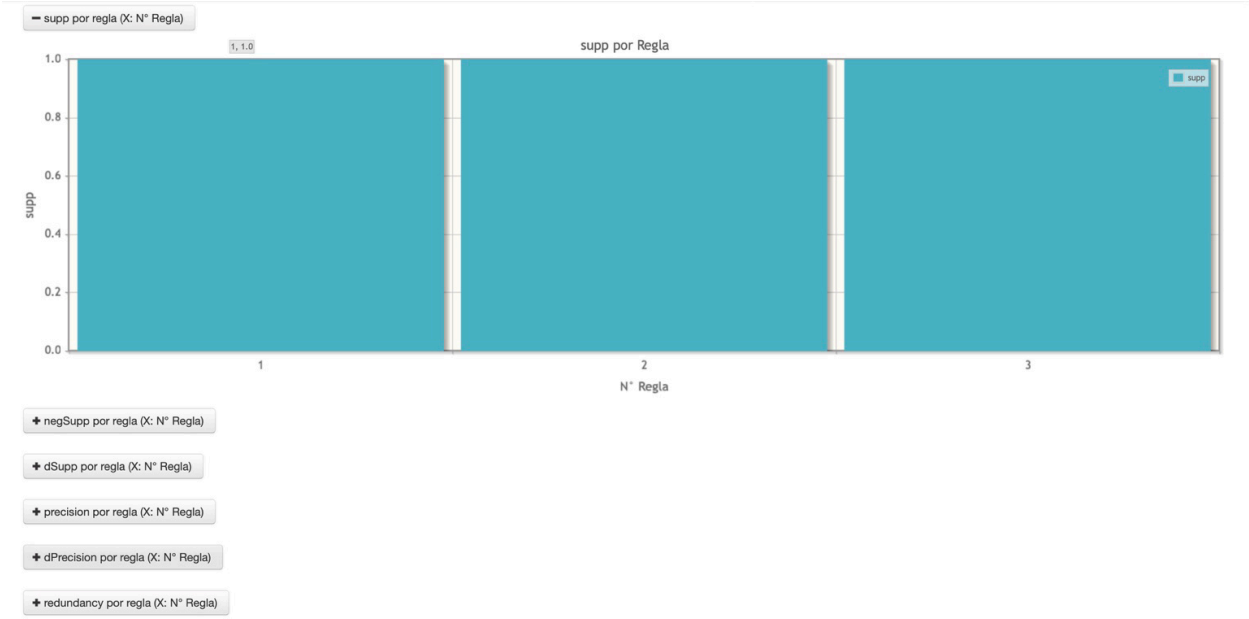


Figura 4.10 Vista del análisis visual de métricas asociadas a las reglas generadas

Cada botón, al seleccionarse, actualiza el área gráfica principal con un nuevo histograma generado dinámicamente mediante el componente gráfico de PrimeFaces. Este diseño modular permite al especialista examinar una métrica a la vez, centrando la atención en la variabilidad particular de cada indicador.

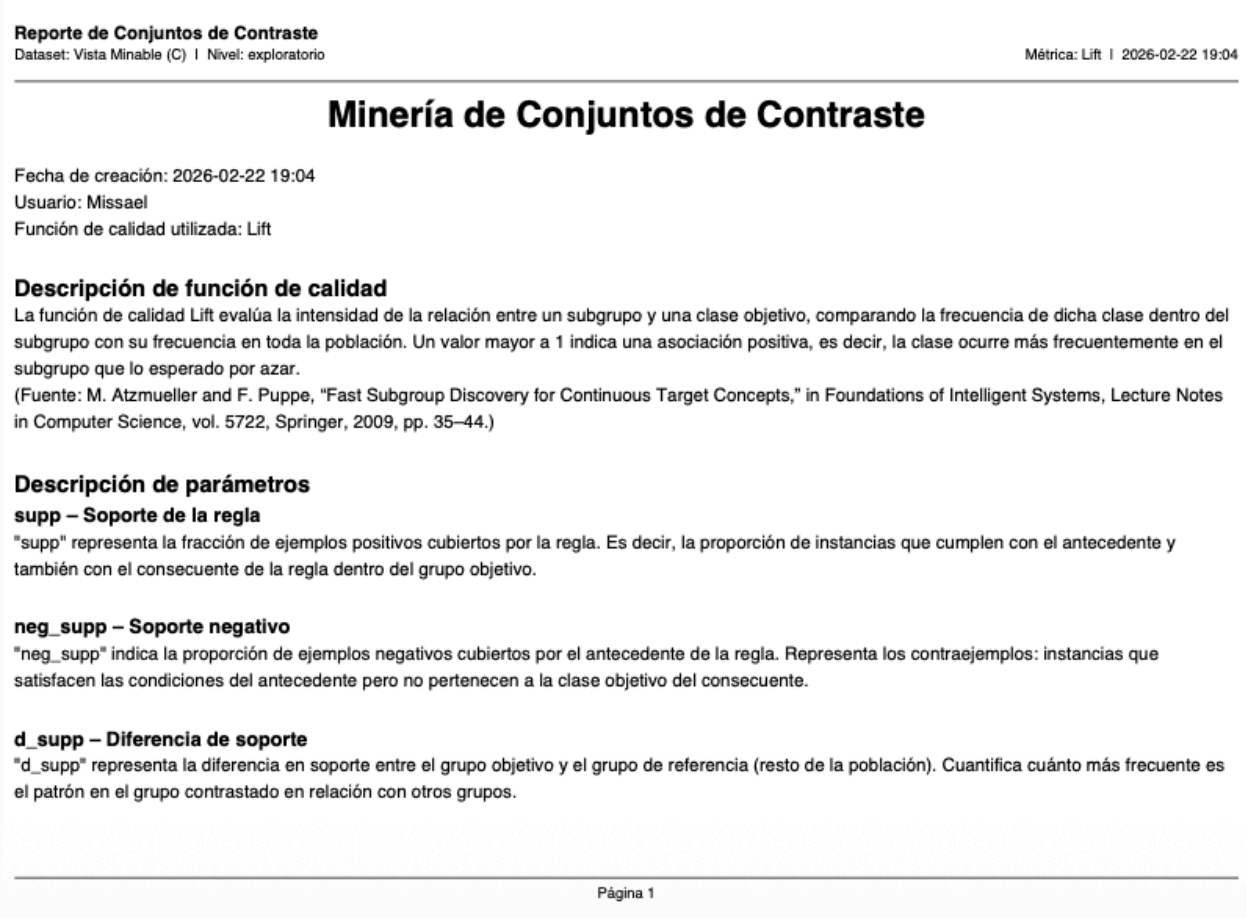
Es importante mencionar que todas estas métricas no son devueltas directamente por RuleKit2-CS, sino que son calculadas por la lógica interna del módulo a partir de los resultados del algoritmo. Esto asegura consistencia con el diseño experimental planteado para la validación del sistema.

Finalmente, esta sección proporciona al usuario una herramienta interpretativa complementaria que permite observar patrones métricos entre reglas, identificar diferencias entre grupos hospitalarios y respaldar la posterior interpretación y discusión de los resultados.

4.1.7 Generación del Reporte PDF

Una vez que RuleKit2-CS concluye satisfactoriamente la ejecución del modelo, el sistema habilita la opción “Descargar Reporte PDF”, mediante la cual se genera un documento completo que consolida todos los elementos esenciales del experimento. Este reporte constituye el producto documental final del módulo y permite preservar, analizar y compartir los resultados obtenidos bajo una estructura formal y reproducible.

En la Figura 4.11, se observa la portada del reporte, en la cual se detallan el *dataset* utilizado, el nivel de análisis, la métrica seleccionada, el usuario ejecutor y la fecha exacta de creación del documento. Esta información se coloca al inicio para fijar desde el principio el contexto experimental bajo el cual fueron obtenidos los resultados.



Reporte de Conjuntos de Contraste
Dataset: Vista Minable (C) | Nivel: exploratorio
Métrica: Lift | 2026-02-22 19:04

Minería de Conjuntos de Contraste

Fecha de creación: 2026-02-22 19:04
Usuario: Missael
Función de calidad utilizada: Lift

Descripción de función de calidad

La función de calidad Lift evalúa la intensidad de la relación entre un subgrupo y una clase objetivo, comparando la frecuencia de dicha clase dentro del subgrupo con su frecuencia en toda la población. Un valor mayor a 1 indica una asociación positiva, es decir, la clase ocurre más frecuentemente en el subgrupo que lo esperado por azar.
(Fuente: M. Atzmueller and F. Puppe, “Fast Subgroup Discovery for Continuous Target Concepts,” in Foundations of Intelligent Systems, Lecture Notes in Computer Science, vol. 5722, Springer, 2009, pp. 35–44.)

Descripción de parámetros

supp – Soporte de la regla
"supp" representa la fracción de ejemplos positivos cubiertos por la regla. Es decir, la proporción de instancias que cumplen con el antecedente y también con el consecuente de la regla dentro del grupo objetivo.

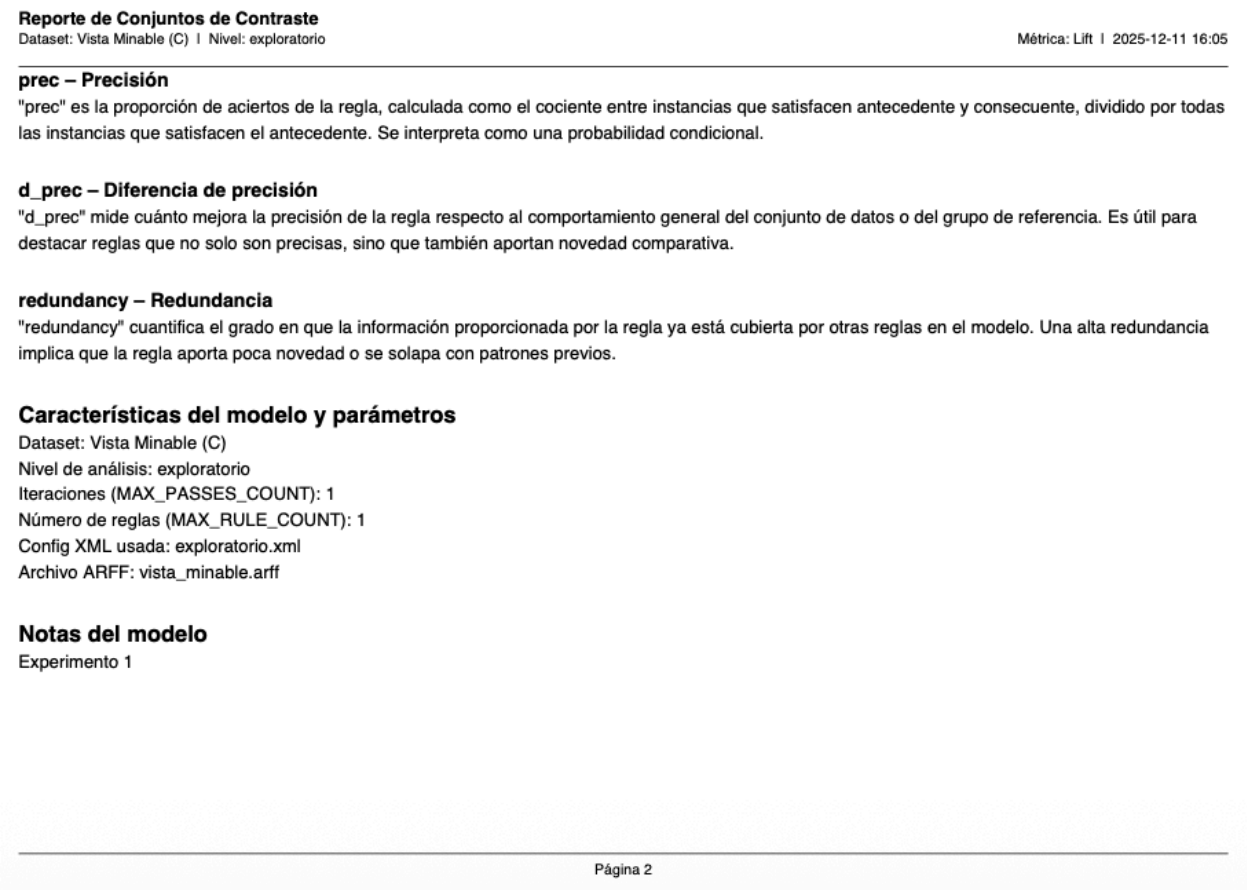
neg_supp – Soporte negativo
"neg_supp" indica la proporción de ejemplos negativos cubiertos por el antecedente de la regla. Representa los contraejemplos: instancias que satisfacen las condiciones del antecedente pero no pertenecen a la clase objetivo del consecuente.

d_supp – Diferencia de soporte
"d_supp" representa la diferencia en soporte entre el grupo objetivo y el grupo de referencia (resto de la población). Cuantifica cuánto más frecuente es el patrón en el grupo contrastado en relación con otros grupos.

Página 1

Figura 4.11 Portada de reporte PDF

En las siguientes secciones, como se aprecia en la Figura 4.11, el sistema incluye una descripción formal de la métrica de calidad utilizada en la ejecución. El texto incluye la definición conceptual de la métrica, su interpretación estadística dentro del análisis de contraste y una referencia bibliográfica precisa que sustenta su aplicación metodológica. Esta parte del reporte tiene el objetivo de ofrecer una base teórica mínima que permita al lector comprender el criterio mediante el cual se evaluaron las reglas.



Reporte de Conjuntos de Contraste
Dataset: Vista Minable (C) | Nivel: exploratorio — Métrica: Lift | 2025-12-11 16:05

prec – Precisión
"prec" es la proporción de aciertos de la regla, calculada como el cociente entre instancias que satisfacen antecedente y consecuente, dividido por todas las instancias que satisfacen el antecedente. Se interpreta como una probabilidad condicional.

d_prec – Diferencia de precisión
"d_prec" mide cuánto mejora la precisión de la regla respecto al comportamiento general del conjunto de datos o del grupo de referencia. Es útil para destacar reglas que no solo son precisas, sino que también aportan novedad comparativa.

redundancy – Redundancia
"redundancy" cuantifica el grado en que la información proporcionada por la regla ya está cubierta por otras reglas en el modelo. Una alta redundancia implica que la regla aporta poca novedad o se solapa con patrones previos.

Características del modelo y parámetros
Dataset: Vista Minable (C)
Nivel de análisis: exploratorio
Iteraciones (MAX_PASSES_COUNT): 1
Número de reglas (MAX_RULE_COUNT): 1
Config XML usada: exploratorio.xml
Archivo ARFF: vista_minable.arff

Notas del modelo
Experimento 1

Página 2

Figura 4.12. Descripción de la métrica y explicación de los parámetros

Posteriormente, el reporte presenta una sección donde se detallan los parámetros de evaluación utilizados para medir el comportamiento y la calidad de cada regla. En esta parte se explica, con un lenguaje técnico pero accesible, el papel de indicadores como el soporte, soporte negativo, diferencia de soporte, precisión, diferencia de precisión y redundancia. Esta información se complementa, como se muestra en la Figura 4.12, con

los parámetros exactos que configuraron el modelo: *dataset*, nivel de análisis, número de iteraciones, número máximo de reglas, archivo ARFF generado y configuración XML utilizada. Esta sección asegura que cualquier usuario pueda reproducir exactamente la misma ejecución sin depender del sistema.

Reporte de Conjuntos de Contraste

Dataset: Vista Minable (C) | Nivel: exploratorio — Métrica: Lift | 2025-11-25 22:55

Reglas sin interpretar

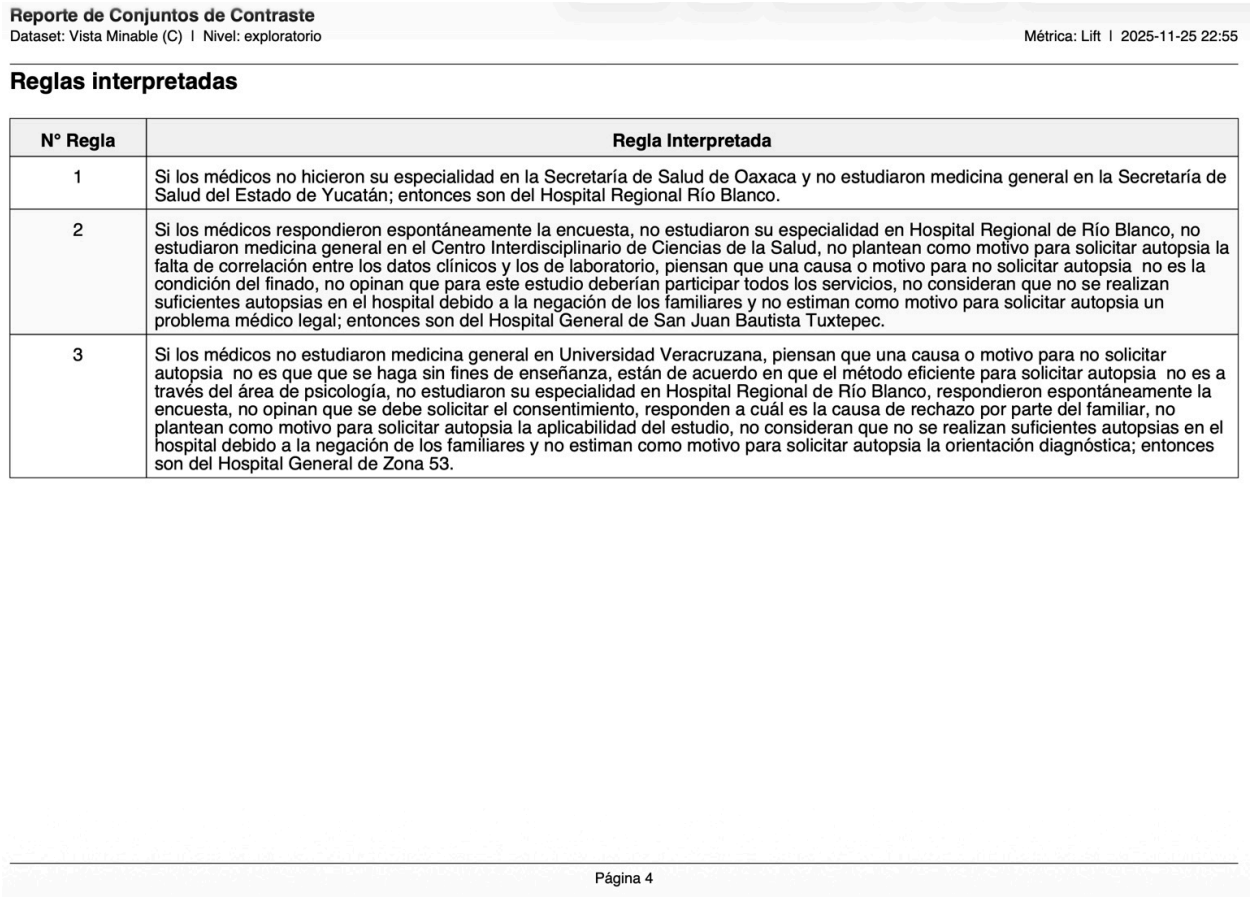
N° Regla	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF esc_esp = !{c103} AND esc_med_gral = !{c104} THEN centro_hospitalario = {H1}"	1.000000	0.517843	0.482157	0.363673	-0.272653	5.000000
2	"IF categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c125} AND mcc_aut = !{14e} AND mcc_no_aut = !{16m} AND com_sug_op = !{22j} AND no_hosp = !{18h} AND med_aut = !{20h} THEN centro_hospitalario = {H11}"	1.000000	0.321575	0.678425	0.853284	0.706569	10457.000000
3	"IF esc_med_gral = !{c1} AND mcc_no_aut = !{16p} AND fmr_sol_aut = !{21h} AND esc_esp = !{c6} AND categoria = {c1} AND com_sug_op = !{22p} AND rechazo_fam = !{17z} AND mcc_aut = !{14d} AND no_hosp = !{18h} AND med_aut = !{20f} THEN centro_hospitalario = {H12}"	1.000000	0.444955	0.555045	0.234632	-0.530737	41101.000000

Página 3

Figura 4.13. Sección del reporte correspondiente a las reglas sin interpretar

En la Figura 4.13, el reporte continúa con la presentación de las reglas sin interpretar, preservando el formato técnico original generado por RuleKit2-CS. La tabla muestra el texto bruto de las reglas en formato IF-THEN, junto con los valores cuantitativos de cada métrica. Esta sección coincide exactamente con lo mostrado en la interfaz del módulo y constituye el registro técnico más directo de los resultados producidos por el motor de minería.

Luego, como se observa en la Figura 4.14, el reporte incluye la versión interpretada de cada regla, la cual se genera mediante las funciones de mapeo y traducción desarrolladas en PostgreSQL. Aquí, cada patrón encontrado se expresa en lenguaje natural, permitiendo que personal médico, administrativo o investigador comprenda sin dificultad las condiciones que caracterizan a cada grupo contrastado y el hospital al que corresponde cada regla. Esta sección es fundamental porque transforma la salida técnica del algoritmo en conocimiento directamente utilizable.



Reporte de Conjuntos de Contraste

Dataset: Vista Minable (C) | Nivel: exploratorio — Métrica: Lift | 2025-11-25 22:55

Reglas interpretadas

N° Regla	Regla Interpretada
1	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca y no estudiaron medicina general en la Secretaría de Salud del Estado de Yucatán; entonces son del Hospital Regional Río Blanco.
2	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud, no plantean como motivo para solicitar autopsia la falta de correlación entre los datos clínicos y los de laboratorio, piensan que una causa o motivo para no solicitar autopsia no es la condición del finado, no opinan que para este estudio deberían participar todos los servicios, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares y no estiman como motivo para solicitar autopsia un problema médico legal; entonces son del Hospital General de San Juan Bautista Tuxtepec.
3	Si los médicos no estudiaron medicina general en Universidad Veracruzana, piensan que una causa o motivo para no solicitar autopsia no es que que se haga sin fines de enseñanza, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología, no estudiaron su especialidad en Hospital Regional de Río Blanco, respondieron espontáneamente la encuesta, no opinan que se debe solicitar el consentimiento, responden a cuál es la causa de rechazo por parte del familiar, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares y no estiman como motivo para solicitar autopsia la orientación diagnóstica; entonces son del Hospital General de Zona 53.

Página 4

Figura 4.14. Reglas interpretadas incluidas dentro del reporte PDF

Finalmente, a partir de la Figura 4.15, el reporte incorpora los gráficos generados durante el análisis visual. El documento dedica una página independiente a cada métrica evaluada, reproduciendo los histogramas que en la interfaz se muestran de manera interactiva. Cada gráfica presenta la variación de la métrica ya sea soporte, soporte negativo, diferencia de soporte, precisión, diferencia de precisión o redundancia asociada

a las reglas generadas. Estas visualizaciones permiten identificar de manera clara comportamientos distintivos entre reglas y facilitan un análisis comparativo que complementa la lectura tabular del reporte.

Debido a que estas páginas son numerosas y cada una corresponde a una métrica distinta, se omite incluirlas de forma individual en esta tesis; sin embargo, el reporte PDF incluye todas las visualizaciones completas generadas a partir del modelo y constituye un registro exhaustivo del comportamiento métrico del conjunto de reglas.

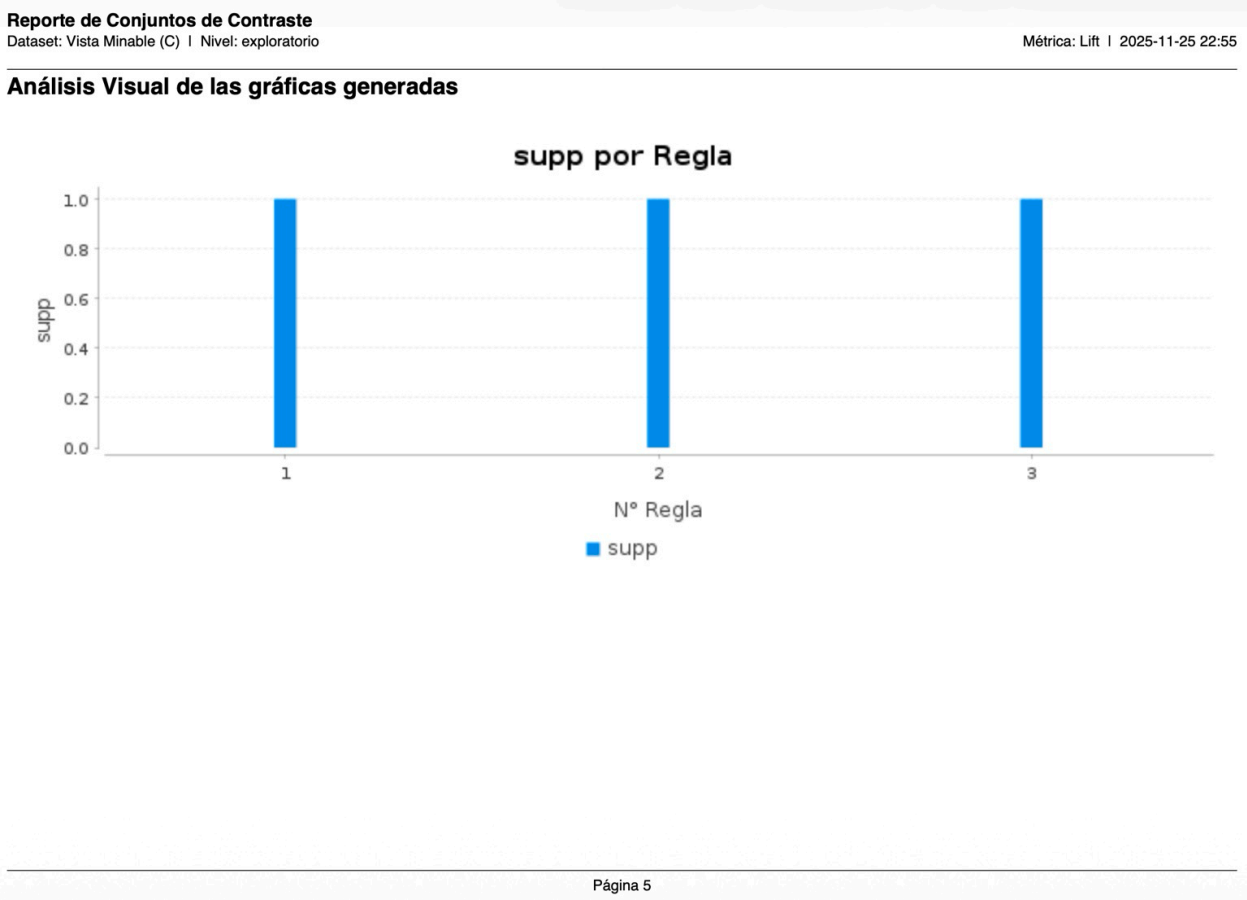


Figura 4.15. Primera de las páginas del análisis visual incluidas en el reporte PDF

Con ello, el reporte PDF se consolida como un documento integral que captura el contexto del modelo, sus parámetros, las reglas generadas y su interpretación tanto textual como visual. Este mecanismo permite preservar la evidencia de cada ejecución,

compartir los resultados con especialistas externos y respaldar el análisis presentado en el capítulo de resultados.

4.2 Vista Minable: Resultados aplicando la función de calidad Lift

La presente sección describe los resultados obtenidos al ejecutar Rulekit2-CS configurado con la función de calidad *Lift* sobre el conjunto de datos Vista Minable (C). La función de calidad *Lift* permitió identificar asociaciones significativas entre los atributos que describen la formación, práctica médica, percepciones sobre la autopsia y la pertenencia a uno de los tres hospitales analizados. A partir del experimento se generaron dieciocho reglas finales, que constituyen patrones estadísticos sólidos para caracterizar diferencialmente a los médicos del Hospital H1, H11 y H12.

La lectura global del experimento revela una estructura altamente contrastada entre hospitales, donde diversas combinaciones de atributos emergen como indicadores casi exclusivos de pertenencia institucional. La robustez de estas reglas se refleja en soportes elevados, precisiones perfectas y diferencias de soporte marcadamente altas, lo que confirma la importancia de la función de calidad *Lift* para identificar subconjuntos clínicos con identidad consistente.

4.2.1 Reglas sin interpretar

La Tabla 4.1 reúne las dieciocho reglas generadas para los tres hospitales. En este caso, los valores de soporte, diferencia de soporte y precisión permiten identificar patrones con asociaciones particularmente fuertes entre los antecedentes y la clase objetivo, coherentes con el comportamiento esperado de la función de calidad *Lift*.

Para el Hospital Regional Río Blanco se obtienen seis reglas con *supp* entre 0.165 y 0.845. Las reglas 1, 2, 3 y 5 muestran precisión perfecta (1.0) y *d_supp* proporcional a su cobertura, destacando la regla 1 (*d_supp* = 0.482) y la regla 2 (*d_supp* = 0.405). La regla 4 presenta un *d_supp* alto (0.428) con precisión de 0.942, mientras que la regla 6, pese a ser la de mayor soporte (*supp* = 0.845), exhibe una precisión reducida (0.427) y

un *dPrecision* negativo, indicando que, aunque frecuente, no incrementa proporcionalmente la probabilidad del hospital objetivo.

Para el Hospital General de San Juan Bautista Tuxtepec se identifican seis reglas con *d_supp* entre 0.164 y 0.661. Las reglas 7, 8, 9 y 10 mantienen precisión igual o cercana a 1.0, con *negSupp* prácticamente nulo. Destaca la regla 12, que alcanza *supp* = 1.0 y *d_supp* = 0.661, convirtiéndose en la regla de mayor cobertura y con un comportamiento estable dentro del subgrupo. Los valores de redundancia, muy elevados en este grupo, indican que múltiples reglas comparten estructuras similares dentro de este hospital.

Para el Hospital General de Zona 53, *Lift* produce seis reglas con *supp* entre 0.126 y 0.982. Las reglas 13, 14, 15 y 16 mantienen precisión igual a 1.0 o muy cercana. La regla 17 destaca por su *d_supp* elevado (0.551) con precisión de 0.938, lo que muestra un patrón altamente diferenciado. La regla 18 presenta el mayor soporte del grupo (*supp* = 0.982), acompañada de *d_supp* = 0.792; sin embargo, su precisión disminuye a 0.414, lo que implica un incremento sustancial en cobertura, pero sin un aumento equivalente en la capacidad de discriminación.

En conjunto, la métrica *Lift* revela patrones con asociaciones fuertes y bien delimitadas para los tres hospitales, apoyados por diferencias de soporte significativas y precisiones elevadas en la mayoría de los casos. Estos valores constituyen la base numérica previa al proceso de interpretación semántica presentado en la siguiente sección.

Tabla 4.1. *Lift* – Vista Minable: Reglas sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	IF categoria = !{c1} THEN centro_hospitalario = {H1}	0.482402	0	0.482402	1	1	5
2	IF esc_esp = {c6} THEN centro_hospitalario = {H1}	0.405214	0	0.405214	1	1	1095
3	IF hall_arb = {8b} AND esc_med_gral = {c1} AND casos = {4a} THEN centro_hospitalario = {H1}	0.165363	0	0.165363	1	1	6
4	IF hall_arb = {8b} AND esc_med_gral = !{c40} AND mcc_no_aut = !{16k} AND anios_prac = !{3c} AND med_aut = !{20z} AND com_sug_op = !{22e} AND fmr_sol_aut = !{21z} AND no_hosp = !{18f} THEN centro_hospitalario = {H1}	0.436406	0.007936	0.42847	0.942111	0.884221	6
5	IF esc_med_gral = {c1} AND area = !{a3} AND ult_grado = !{g1} THEN centro_hospitalario = {H1}	0.214618	0	0.214618	1	1	5
6	IF esc_esp = !{c103} AND med_aut = !{20d} AND per_sol_aut = !{19c} AND hall_dem = !{11e} AND fmr_sol_aut = !{21h} THEN centro_hospitalario = {H1}	0.845158	0.335088	0.510071	0.427414	-0.145171	5
7	IF esc_esp = {c103} AND hall_arb = !{8c} THEN centro_hospitalario = {H11}	0.490439	0	0.490439	1	1	11638
8	IF hall_dem = {11a} AND casos = !{4a} AND per_sol_aut = !{19a} AND categoria = {c1} AND hall_disc = !{7b} THEN centro_hospitalario = {H11}	0.497716	0	0.497716	1	1	10624
9	IF esc_med_gral = {c89} AND mcc_aut = !{14e} AND fmr_sol_aut = !{21b} THEN centro_hospitalario = {H11}	0.320324	0	0.320324	1	1	11641
10	IF hall_dem = {11a} AND ult_grado = {g2} AND area = !{a2} THEN centro_hospitalario = {H11}	0.312198	0.001099	0.3111	0.998122	0.996244	10939

Tabla 4.1. *Lift* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
11	IF hall_arb = {8a} AND anios_prac = !{3a} AND com_sug_op = !{22a} AND mcc_no_aut = {16f} AND med_aut = !{20f} THEN centro_hospitalario = {H11}	0.168581	0.003845	0.164736	0.987952	0.975904	10457
12	IF categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c125} AND no_hosp = !{18h} THEN centro_hospitalario = {H11}	1	0.338541	0.661459	0.84673	0.69346	10457
13	IF esc_med_gral = {c125} THEN centro_hospitalario = {H12}	0.365102	0	0.365102	1	1	42424
14	IF esc_esp = {c111} AND hall_dem = !{11d} THEN centro_hospitalario = {H12}	0.16581	0	0.16581	1	1	41101
15	IF ult_grado = {g8} AND esc_med_gral = {c40} THEN centro_hospitalario = {H12}	0.1938	0.000024	0.193776	0.999087	0.998174	45574
16	IF hall_disc = {7b} AND com_sug_op = !{22k} AND categoria = {c1} AND anios_prac = !{3e} AND casos = {4b} AND per_sol_aut = !{19b} AND rechazo_fam = !{17d} THEN centro_hospitalario = {H12}	0.126661	0	0.126661	1	1	42060
17	IF hall_disc = !{7a} AND esc_med_gral = !{c1} AND casos = !{4c} AND anios_prac = !{3e} AND mcc_no_aut = !{16c} AND mcc_aut = !{14d} AND hall_arb = !{8c} AND fmr_sol_aut = !{21h} AND med_aut = !{20f} THEN centro_hospitalario = {H12}	0.556422	0.005002	0.55142	0.938172	0.876344	41245

Tabla 4.1. *Lift* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
18	IF esc_esp = !{c103} AND esc_med_gral = !{c1} AND per_sol_aut = !{19b} AND fmr_sol_aut = !{21h} AND categoria = {c1} AND mcc_no_aut = !{16e} AND com_sug_op = !{22p} AND hall_arb = !{8d} AND no_hosp = !{18h} THEN centro_hospitalario = {H12}	0.982108	0.1893	0.792808	0.414412	-0.171177	41101

La observación detallada de esta tabla permite apreciar que, incluso antes de incorporar la interpretación semántica, los patrones descubiertos describen estructuras robustas y claramente diferenciadas dentro de los datos. La presencia de soportes elevados, precisiones perfectas y diferencias de soporte muy marcadas confirma que la función de calidad *Lift* capturó relaciones fuertes entre las trayectorias formativas, percepciones clínicas y prácticas institucionales de los médicos. De esta forma, las reglas sin interpretar constituyen el punto de partida para el análisis semántico que se desarrolla en la sección siguiente, donde se examina en profundidad el significado clínico e institucional de cada combinación descubierta por el modelo.

4.2.2 Reglas interpretadas

Una vez obtenidas las reglas sin interpretar y calculadas sus métricas, se procedió a su traducción semántica con el objetivo de identificar el significado clínico, formativo y organizacional de cada patrón. Esta fase es esencial para comprender cómo las combinaciones de atributos encontradas por RuleKit2-CS representan estructuras reales dentro de la práctica médica, permitiendo relacionar los resultados estadísticos con las dinámicas institucionales de cada hospital. Las reglas interpretadas mostradas en la Tabla 4.2 posibilitan reconstruir perfiles distintivos de los médicos, evidenciando diferencias profundas en su formación, percepción de las autopsias y experiencia profesional.

Al analizar las reglas correspondientes al primer hospital se observa que gran parte de los patrones identificados describen a médicos cuya identidad profesional se encuentra estrechamente vinculada a su trayectoria formativa. La presencia reiterada de atributos relacionados con haber estudiado la especialidad en la institución o en universidades cercanas, así como la frecuencia con la que aparecen referencias a estudios de medicina general en facultades específicas, revela que la formación local constituye un rasgo que distingue de manera notable al personal médico de este centro. De igual modo, varios patrones indican percepciones particulares sobre la autopsia, especialmente en torno a su relación con posibles arbitrajes o demandas. En estos casos, la regla traducida muestra un grupo de médicos que, aunque reconoce el valor clínico de la autopsia, también manifiesta cautela respecto a sus implicaciones legales. Estas combinaciones sugieren un entorno institucional donde la experiencia formativa y las consideraciones legales conviven como elementos que influyen directamente en la percepción del procedimiento *post mortem*.

Tabla 4.2. *Lift* – Vista Minable: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos no respondieron espontáneamente la encuesta; entonces son del Hospital Regional Río Blanco.
2	Si los médicos estudiaron su especialidad en Hospital Regional de Río Blanco; entonces son del Hospital Regional Río Blanco.
3	Si los médicos están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, estudiaron medicina general en Universidad Veracruzana y no han participado en ningún estudio de autopsia; entonces son del Hospital Regional Río Blanco.
4	Si los médicos están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en UNAM, piensan que una causa o motivo para no solicitar autopsia no es por cuestiones legales, no tienen entre 11 y 15 años de práctica, especifican motivos para solicitar autopsia, no opinan que son importantes las autopsias, especifican cuál es el método eficiente para solicitar autopsia y no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica; entonces son del Hospital Regional Río Blanco.
5	Si los médicos estudiaron medicina general en Universidad Veracruzana, no son adscritos y no son especialistas; entonces son del Hospital Regional Río Blanco.

Tabla 4.2. *Lift* – Vista Minable: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
6	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estiman como motivo para solicitar autopsia la negligencia médica, alegan que no es el trabajador social el personal adecuado para ordenar autopsia, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda y están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología; entonces son del Hospital Regional Río Blanco.
7	Si los médicos hicieron su especialidad en la Secretaría de Salud de Oaxaca y no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje; entonces son del Hospital General de San Juan Bautista Tuxtepec.
8	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no han participado en ningún estudio de autopsia, alegan que no es el médico el personal adecuado para ordenar autopsia, respondieron espontáneamente la encuesta y no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de San Juan Bautista Tuxtepec.
9	Si los médicos estudiaron medicina general en la Universidad Autónoma de Tlaxcala, no plantean como motivo para solicitar autopsia la falta de correlación entre los datos clínicos y los de laboratorio y están de acuerdo en que el método eficiente para solicitar autopsia no es solicitar el consentimiento de autopsia a los familiares inmediatamente después de la defunción; entonces son del Hospital General de San Juan Bautista Tuxtepec.
10	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, son médicos generales y no son residentes; entonces son del Hospital General de San Juan Bautista Tuxtepec.
11	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no tienen menos de 5 años de práctica, comentan, piensan que una causa o motivo para no solicitar autopsia es por enfermedad de base conocida y no estiman como motivo para solicitar autopsia la orientación diagnóstica; entonces son del Hospital General de San Juan Bautista Tuxtepec.
12	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud y no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares; entonces son del Hospital General de San Juan Bautista Tuxtepec.
13	Si los médicos estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud; entonces son del Hospital General de Zona 53.

Tabla 4.2. *Lift* – Vista Minable: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
14	Si los médicos hicieron su especialidad en la Secretaría de Salubridad y Asistencia y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.
15	Si los médicos son estudiantes de medicina y estudiaron medicina general en UNAM; entonces son del Hospital General de Zona 53.
16	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no agradecen por la encuesta, respondieron espontáneamente la encuesta, no tienen más de 20 años de práctica, han participado en menos de 5 casos de autopsias, alegan que no es la enfermera el personal adecuado para ordenar autopsia y creen que la causa de rechazo por parte del familiar no es por no haber solicitado la autorización en forma adecuada; entonces son del Hospital General de Zona 53.
17	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Veracruzana, no han participado entre 6 y 10 casos de autopsia, no tienen más de 20 años de práctica, piensan que una causa o motivo para no solicitar autopsia no es la negativa de los familiares, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología y no estiman como motivo para solicitar autopsia la orientación diagnóstica; entonces son del Hospital General de Zona 53.
18	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Veracruzana, alegan que no es la enfermera el personal adecuado para ordenar autopsia, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología, respondieron espontáneamente la encuesta, piensan que una causa o motivo para no solicitar autopsia no es el desinterés, no opinan que se debe solicitar el consentimiento, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje y no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares; entonces son del Hospital General de Zona 53.

Las reglas interpretadas del segundo hospital presentan un perfil marcadamente distinto. En este caso, las combinaciones encontradas por el algoritmo destacan reiteradamente el convencimiento de los médicos sobre la capacidad de la autopsia para originar demandas o conflictos legales. Esta percepción no aparece de manera aislada, sino acompañada de condiciones relacionadas con el rol de los médicos en la solicitud de

autopsias, la suficiencia de los diagnósticos clínicos y la naturaleza administrativa del proceso. En su conjunto, las reglas muestran un entorno institucional donde la interpretación del riesgo legal constituye un eje central de la práctica médica. Este perfil se manifiesta de forma especialmente homogénea, lo que explica por qué las reglas de este hospital presentan precisiones perfectas y soportes elevados: los médicos comparten una visión común sobre la autopsia y su potencial impacto en el ámbito jurídico.

El tercer hospital exhibe un patrón más heterogéneo. Las reglas traducidas revelan una diversidad formativa significativa: los médicos provienen de instituciones educativas variadas y de prestigio nacional, lo que genera una estructura interna más compleja. En estas reglas también aparecen atributos asociados a discrepancias diagnósticas, participación limitada en casos de autopsia y percepciones matizadas respecto a la responsabilidad de ordenar el procedimiento. En este centro se identifican tanto reglas que describen grupos amplios de médicos con características comunes, como reglas destinadas a subgrupos más específicos que exhiben condiciones relacionadas con la interacción con familiares o con la forma en que se justifica la solicitud de un estudio *post mortem*. Esta variedad indica que los patrones institucionales no se articulan en torno a un solo eje como ocurre en el segundo hospital, sino que surgen de la combinación de múltiples dimensiones de la práctica médica.

La interpretación conjunta de estas reglas permite observar tres estructuras claramente diferenciadas. El primer hospital se caracteriza por una identidad fuertemente arraigada en la formación local y en percepciones moderadas sobre la autopsia. El segundo presenta una visión institucional compacta, centrada en la dimensión jurídico-legal del procedimiento. El tercero, en contraste, muestra un sistema más amplio y diverso, donde convergen trayectorias formativas distintas y percepciones técnicas sobre la utilidad diagnóstica y administrativa del estudio *post mortem*.

En síntesis, las reglas interpretadas no solo traducen combinaciones lógicas en descripciones comprensibles, sino que revelan las bases culturales, formativas y

operativas que estructuran la práctica médica en cada hospital. Estas reglas constituyen la evidencia semántica que sustenta las diferencias estadísticas observadas en el experimento y preparan el terreno para el análisis visual, la validación y la comparación detallada que se presentan en las secciones siguientes.

4.2.3 Análisis Visual

El análisis visual constituye una fase esencial para comprender la forma en que las reglas identificadas por RuleKit2-CS se distribuyen a lo largo del conjunto de datos y cómo se comportan en términos de cobertura, exclusividad, contraste y estabilidad. Las seis gráficas generadas a partir de las métricas calculadas por el módulo permiten observar patrones que complementan y fortalecen la lectura numérica presentada en la sección previa. Cada gráfica revela una dimensión distinta del comportamiento de las reglas, y en conjunto forman una representación integral del desempeño de la función de calidad *Lift* en este experimento.

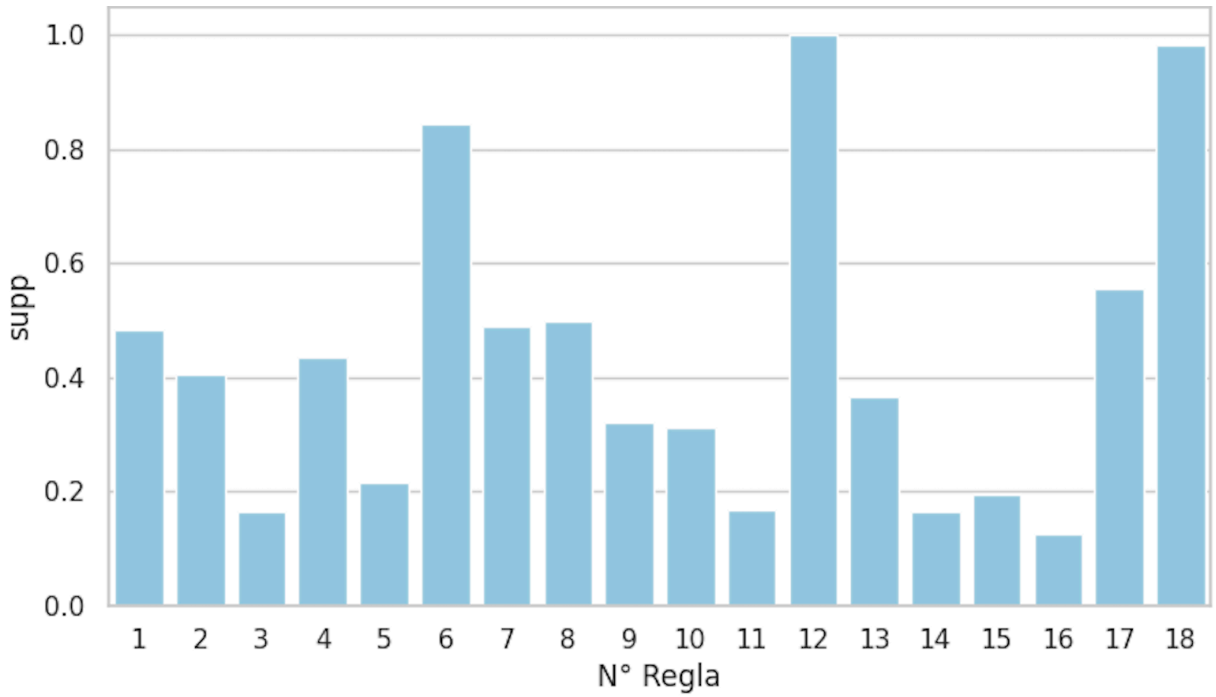


Figura 4.16 *Lift* – Vista Minable: Soporte por Regla

La primera gráfica representada en la Figura 4.16, correspondiente al soporte por regla, muestra la proporción real de médicos en el hospital objetivo que cumplen las

condiciones establecidas por cada antecedente. En esta visualización se observa un comportamiento heterogéneo, con reglas que alcanzan soportes modestos como las reglas 3, 5, 14 y 16 y otras que describen grupos muy amplios, destacando especialmente la regla 6 con un soporte superior al 0.84, así como la regla 12 y 18, cuyo valor se aproxima al 1.0. Las oscilaciones en esta gráfica reflejan la diversidad estructural del conjunto de datos: algunos patrones representan segmentos específicos con identidad marcada, mientras otros abarcan grandes proporciones del personal médico. Esta variabilidad confirma que la función *Lift* no solo identifica reglas exclusivas, sino también patrones generales que describen tendencias amplias dentro de un hospital.

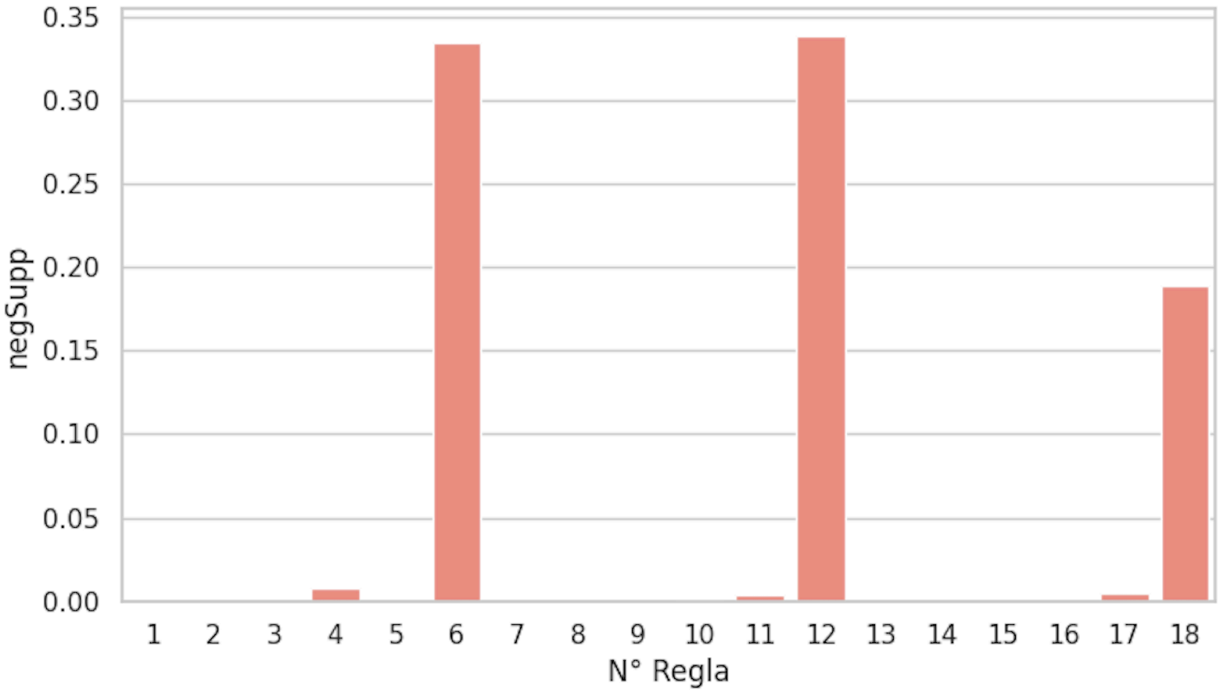


Figura 4.17 *Lift* – Vista Minable: Soporte Negativo por Regla

La segunda gráfica representada en la Figura 4.17, correspondiente al soporte negativo, evidencia la proporción de contraejemplos, es decir, casos en los que el antecedente se cumple, pero el médico no pertenece al hospital sugerido por la regla. La mayor parte de las reglas presenta valores cercanos a cero, lo que indica que los atributos descritos por RuleKit2-CS son altamente exclusivos y rara vez aparecen en médicos de otros hospitales. Solo algunas reglas como la regla 6, 12 y 18 muestran valores superiores, lo

que sugiere la presencia de patrones más difusos en estos casos. La interpretación de esta gráfica resulta crucial, ya que la ausencia de contraejemplos constituye una de las características más deseables en la minería de conjuntos de contraste: garantiza que la regla representa una estructura realmente distintiva.

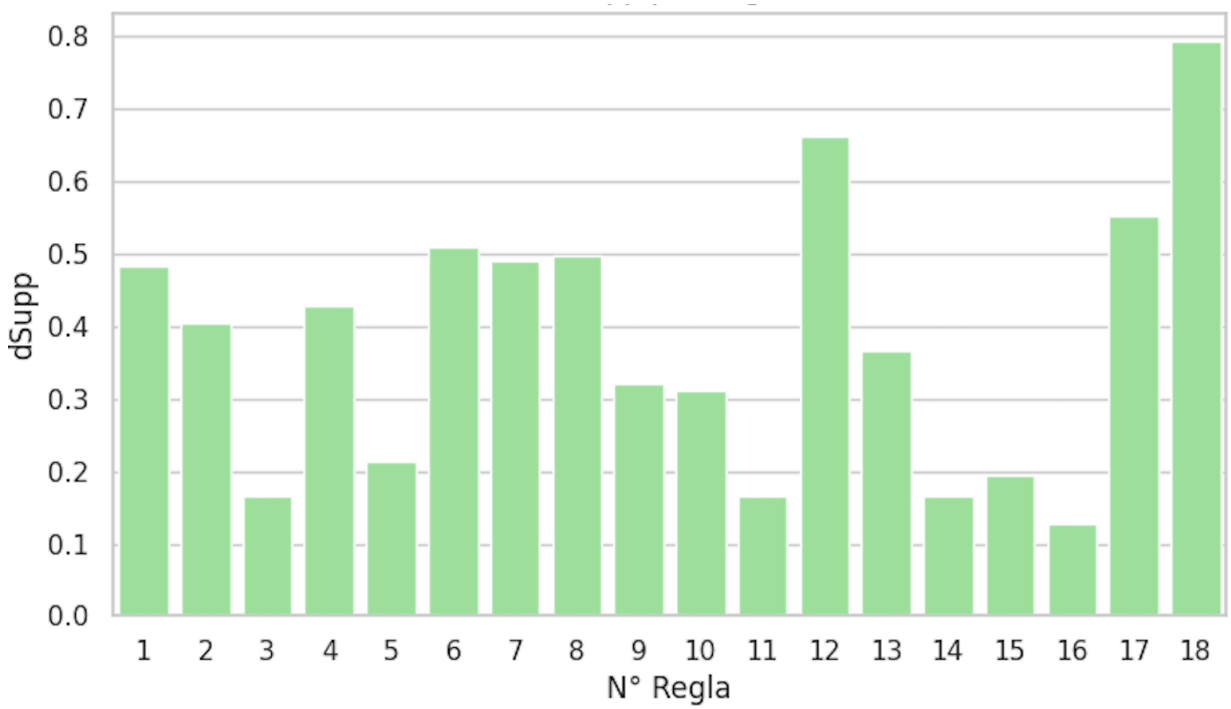


Figura 4.18 *Lift* – Vista Minable: Diferencia de Soporte por Regla

La tercera gráfica representada en la Figura 4.18, que muestra la diferencia de soporte, permite visualizar el grado de contraste entre el hospital objetivo y el resto. Aquí se observa un conjunto de reglas con valores consistentemente altos, especialmente la regla 18, cuyo valor supera el 0.79. Este comportamiento indica que las reglas generadas con *Lift* describen patrones cuya presencia es significativamente mayor en el hospital objetivo que en los demás. La distribución de esta gráfica confirma la existencia de perfiles institucionales bien definidos, donde ciertos atributos se concentran de manera notable en un hospital específico. La estabilidad de valores elevados en esta métrica demuestra que la función *Lift* operó con una capacidad consistente para identificar diferencias profundas entre los grupos del estudio.

En la gráfica correspondiente a la precisión por regla representada en la Figura 4.19, se observa que la mayoría de los patrones alcanzan valores prácticamente perfectos. Catorce de las dieciocho reglas muestran precisiones iguales a uno, lo que significa que cada médico que cumple el antecedente pertenece efectivamente al hospital objetivo, sin excepción. La gráfica destaca la regla 6 y la regla 18 como los únicos patrones con precisiones inferiores, alcanzando valores alrededor de 0.42. La homogeneidad visual de esta gráfica sugiere que *Lift* tiende a privilegiar reglas con una relación altamente determinística entre antecedente y consecuente, lo que refuerza su utilidad como métrica para identificar asociaciones fuertes en contextos clínicos e institucionales.

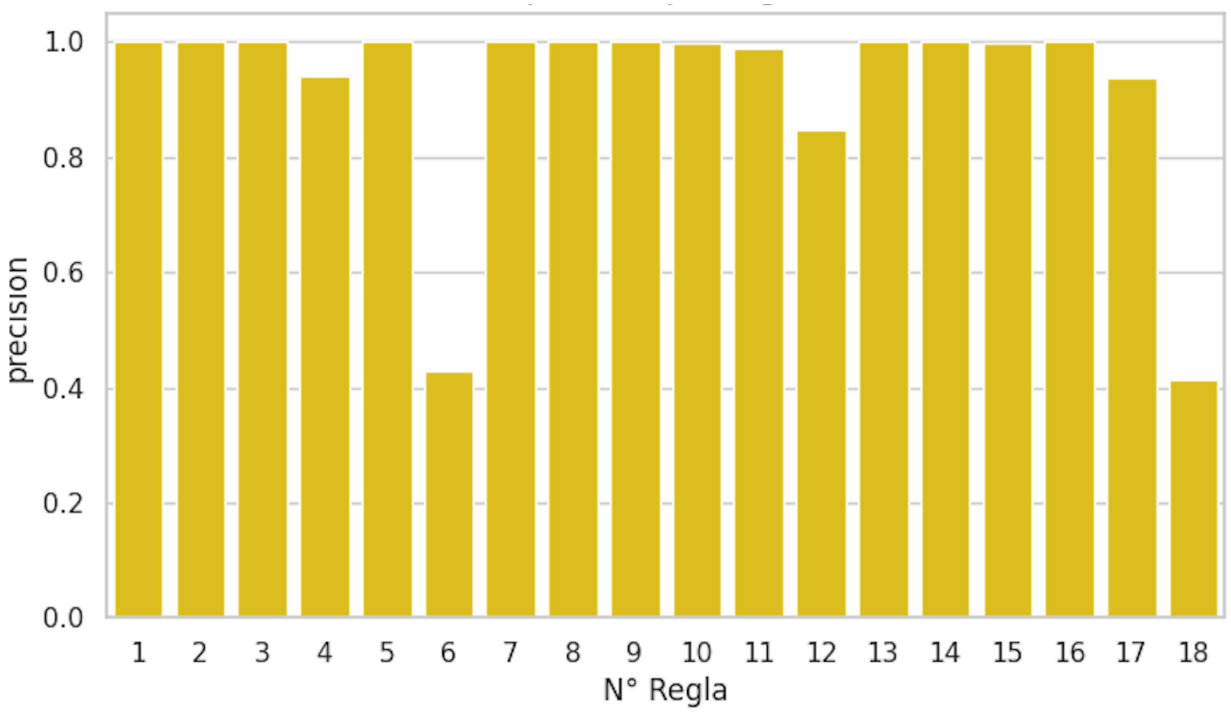


Figura 4.19 *Lift* – Vista Minable: Precisión por Regla

La quinta gráfica, que muestra la diferencia de precisión representada en la Figura 4.20, profundiza la lectura anterior, ya que indica cuánto mejora la precisión de la regla con respecto a la distribución global del conjunto de datos. Una vez más, la mayoría de las reglas exhibe valores elevados, cercanos a uno, señalando una mejora significativa respecto al comportamiento general. Las reglas 6 y 18 presentan valores negativos o

reducidos, lo que indica que, aunque cubren segmentos amplios, especialmente la regla 18 su especificidad relativa es menor. Esta gráfica permite identificar la estabilidad de los patrones: valores altos sugieren reglas que no solo son exactas, sino que superan claramente la distribución natural del conjunto.

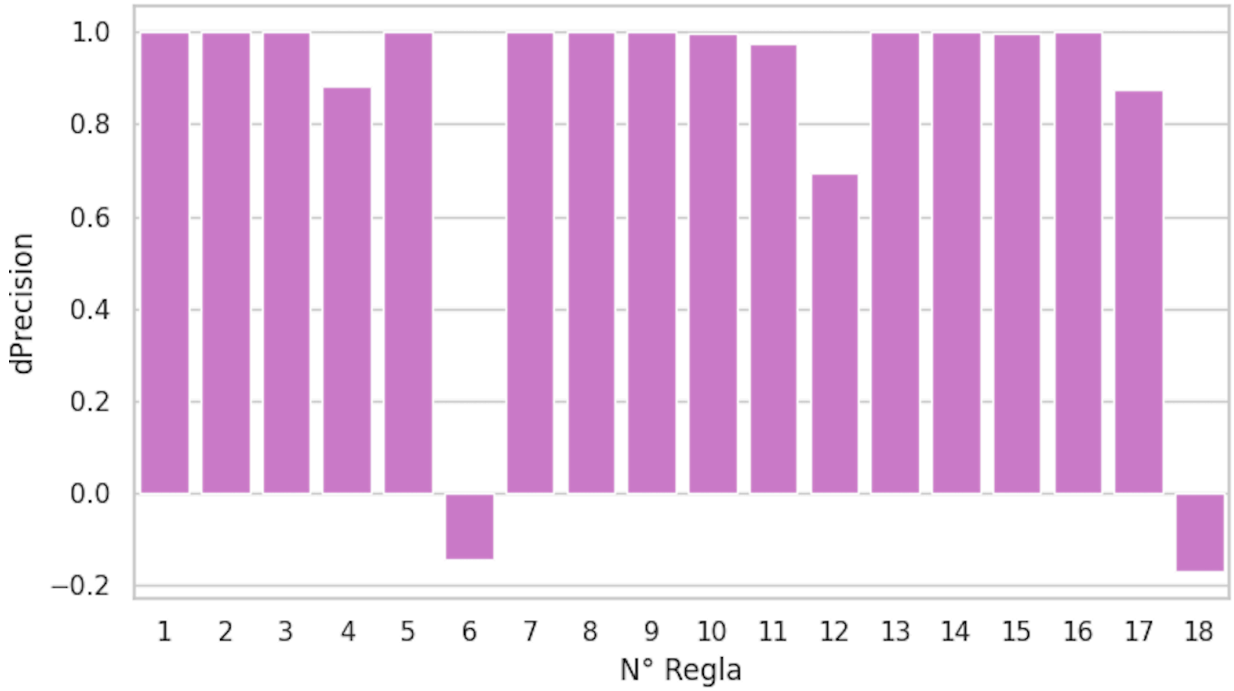


Figura 4.20 *Lift* – Vista Minable: Diferencia de Precisión por Regla

Finalmente, la gráfica de redundancia por regla representada en la Figura 4.21 ofrece un análisis sobre la complejidad combinatoria asociada a cada patrón. Se observa un comportamiento bifurcado: mientras las reglas correspondientes al primer y segundo hospital muestran valores moderados, las reglas asociadas al tercero presentan redundancias extremadamente elevadas, con algunos valores que superan los cuarenta mil. Esto significa que existen múltiples configuraciones diferentes de atributos que llevan al mismo patrón institucional, especialmente en el caso del tercer hospital. Esta alta redundancia indica un sistema interno más variado, donde diversas trayectorias formativas y percepciones clínicas convergen hacia un mismo resultado estadístico. La gráfica de redundancia es, por lo tanto, una evidencia visual de que la función *Lift* no solo identifica patrones exclusivos, sino también patrones complejos cuya estructura interna requiere una interpretación más matizada.

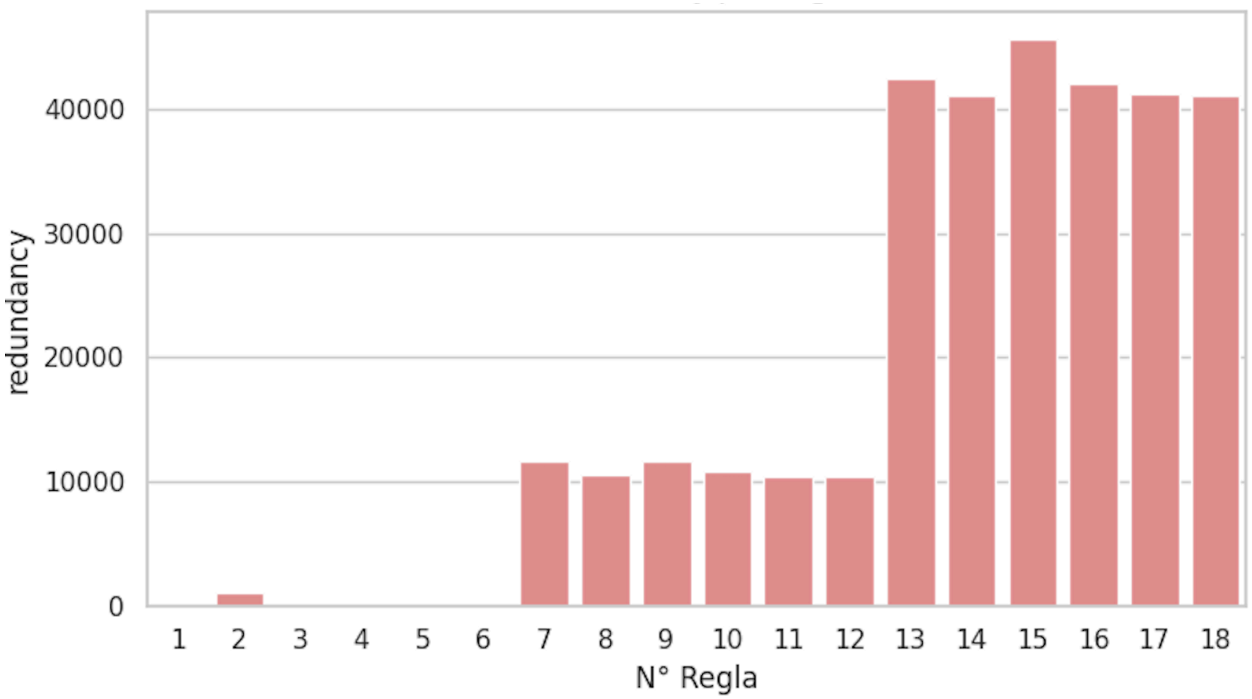


Figura 4.21 *Lift* – Vista Minable: Redundancia por Regla

En conjunto, las seis gráficas permiten apreciar que la función de calidad *Lift* mantuvo un desempeño notablemente robusto en términos de cobertura, exclusividad y contraste. La subordinación visual entre soporte, diferencia de soporte y soporte negativo ilustra la capacidad del algoritmo para capturar patrones claramente discriminativos. Las gráficas de precisión y diferencia de precisión revelan asociaciones altamente determinísticas, mientras que la gráfica de redundancia proporciona una lectura adicional sobre la diversidad combinatoria de los atributos que estructuran dichos patrones. Este análisis visual confirma la solidez de la función *Lift* en este experimento y constituye un fundamento claro para la evaluación cuantitativa y cualitativa que se presenta en las secciones siguientes.

4.2.4 Evaluación Objetiva

La evaluación objetiva de los resultados obtenidos mediante la función de calidad *Lift* se fundamenta en la magnitud y consistencia de las métricas calculadas para las dieciocho

reglas generadas en el experimento. Los valores numéricos asociados a estas reglas muestran un comportamiento estable y altamente discriminativo entre los hospitales analizados, lo que confirma que los patrones identificados tienen una base estadística sólida.

Un primer indicador de validez se observa en los valores de soporte, donde doce reglas presentan coberturas superiores al 0.30, destacando casos como el de la regla 6, cuyo soporte alcanza 0.845, describiendo a más del 84% de los médicos del hospital objetivo. Del mismo modo, la regla 18 presenta un soporte de 0.982, lo que significa que prácticamente todo el personal médico del hospital considerado cumple el patrón descrito por la regla. Estas cifras evidencian que los antecedentes identificados por RuleKit2-CS no se limitan a subgrupos marginales, sino que capturan comportamientos centrales dentro de la población estudiada.

En contraste, el soporte negativo exhibe una tendencia opuesta. Diez de las dieciocho reglas presentan valores exactamente iguales a 0, y otras cinco muestran valores menores a 0.01, como ocurre en la regla 15 (0.000024) y en la regla 10 (0.001099). Esta ausencia casi total de contraejemplos confirma que la mayoría de los patrones son exclusivos del hospital objetivo. La única excepción notable corresponde a las reglas 6, 12 y 18, con soportes negativos de 0.335, 0.338 y 0.189,, respectivamente; sin embargo, incluso en estos casos el soporte positivo supera ampliamente al negativo, manteniendo la validez contrastiva.

La diferencia de soporte, métrica esencial para la minería de conjuntos de contraste, presenta valores altamente diferenciados entre hospitales. La regla 18, con un *dSupp* de 0.793, y la regla 12, con 0.661, representan contrastes institucionales particularmente marcados. En total, doce reglas muestran valores superiores a 0.30, lo que refleja una concentración sustancial de los patrones dentro del hospital correspondiente y una ausencia significativa de estas condiciones en los demás centros. La predominancia de diferencias de soporte elevadas descarta la posibilidad de que los patrones surgieran por azar o por distribuciones uniformes entre los hospitales.

La precisión de las reglas constituye uno de los aspectos más categóricos de esta validación. Diez reglas presentan una precisión igual a 1.000, lo que significa que cada médico que cumple el antecedente pertenece de manera inequívoca al hospital objetivo. Esta proporción, equivalente al 55.6% de las reglas totales, es excepcional en análisis aplicados a dominios clínicos reales, donde es común encontrar solapamientos en percepciones y prácticas profesionales. Incluso las reglas con precisiones inferiores 0.9991 en la regla 15, 0.988 en la regla 11 y 0.943 en la regla 4 mantienen niveles de exactitud elevados. Las únicas reglas con precisiones considerablemente bajas, la 6 y la 18, aún conservan soportes altos y diferencias de soporte significativas, lo que preserva su relevancia estadística.

La diferencia de precisión refuerza estas observaciones. Trece reglas presentan valores iguales a 1.000 o muy cercanos a él, como 0.998 en la regla 15 y 0.996 en la regla 10. Estas cifras indican que la precisión de las reglas supera ampliamente el comportamiento general del conjunto de datos, incrementando significativamente la probabilidad de pertenecer al hospital objetivo al cumplirse el antecedente. Solo dos reglas exhiben diferencias de precisión negativas (–0.145 y –0.171), lo que refleja una menor exclusividad relativa, pero aun así mantienen valores de dSupp y soportes suficientes para ser consideradas patrones válidos.

Finalmente, el análisis de redundancia aporta información sobre la estabilidad interna de los patrones. Las reglas asociadas al tercer hospital alcanzan valores de redundancia extremadamente altos, como los 45 574 de la regla 15 o los 42 424 de la regla 13. Estos valores implican que existen decenas de miles de combinaciones alternativas de atributos que podrían generar el mismo patrón, lo que indica una estructura institucional donde convergen múltiples trayectorias formativas y perceptuales hacia un mismo comportamiento global. La redundancia moderada de las reglas asociadas a los otros hospitales generalmente por debajo de 12,000 sugiere estructuras más homogéneas, pero igualmente estables.

En conjunto, los valores cuantitativos obtenidos muestran que las reglas generadas con la función *Lift* satisfacen plenamente los criterios estadísticos de consistencia, exclusividad y diferencia contrastiva. La magnitud de sus métricas permite afirmar que los patrones identificados representan asociaciones robustas y reproducibles dentro del conjunto C, proporcionando la base numérica necesaria para avanzar hacia la interpretación clínica y la comparación institucional.

4.2.5 Evaluación Subjetiva

La evaluación subjetiva fue realizada por el experto clínico Dr. José Antonio Palet Guzmán, médico patólogo con amplia experiencia en práctica hospitalaria y conocimiento profundo de los procesos asociados a la autopsia, la discrepancia diagnóstica y la dinámica institucional. Durante el proceso de evaluación, el especialista revisó las 18 reglas interpretadas en lenguaje natural, verificando la coherencia de cada una respecto al comportamiento clínico real y a la estructura organizacional de los hospitales analizados.

El Dr. Palet Guzmán confirmó que todas las reglas traducidas fueron entendibles, consistentes y clínicamente plausibles, destacando que la traducción realizada desde los códigos del conjunto de datos hacia expresiones comprensibles mantuvo la fidelidad conceptual de cada atributo. En su dictamen, señaló que las combinaciones identificadas mediante la función *Lift* reflejan patrones reales asociados a diferencias en la formación académica, la percepción del riesgo legal, la experiencia en autopsias y la interacción con procesos administrativos internos. Estas características corresponden, según su evaluación, a escenarios observables y habituales dentro del contexto hospitalario.

El especialista concluyó que las 18 reglas generadas fueron aceptadas en su totalidad, tanto en su forma estadística como en su traducción semántica, lo que respalda la pertinencia del modelo para capturar patrones diferenciales entre instituciones. La coincidencia entre los resultados del módulo y la experiencia profesional del experto constituye evidencia sólida de que las reglas no solo son válidas desde el punto de vista matemático, sino también desde la práctica clínica y la realidad institucional.

4.3 Vista Minable: Resultados aplicando la función de calidad WRAcc

La presente sección expone los resultados obtenidos al aplicar la medida *Weighted Relative Accuracy* (*WRAcc*) sobre el Conjunto C dentro del módulo de minería de conjuntos de contraste desarrollado para esta tesis. WRAcc constituye una de las métricas más empleadas en CSM debido a su capacidad para equilibrar simultáneamente dos aspectos fundamentales: la frecuencia con la que un patrón aparece en el subgrupo objetivo (*coverage*), y la mejora real que dicho patrón aporta respecto al comportamiento promedio del conjunto de datos (*accuracy lift*).
A diferencia de otras funciones de calidad como *Lift*, *Correlation* o *Relative Risk*, cuyo enfoque se centra en la fuerza de asociación entre el antecedente y la clase objetivo, *WRAcc* prioriza la utilidad práctica de las reglas dentro del análisis contrastivo. Esta utilidad surge de combinar tamaño del subgrupo y ganancia informativa, lo que permite identificar patrones que no sólo son correctos, sino también representativos y estadísticamente relevantes. Este enfoque resulta especialmente adecuado en el análisis de encuestas sobre autopsias, donde ciertas combinaciones de atributos pueden ser poco frecuentes pero altamente discriminativas, mientras que otras pueden ser amplias pero carecer de valor analítico.

El análisis con *WRAcc* fue ejecutado utilizando RuleKit2-CS sobre la estructura Vista Minable, manteniendo la misma configuración operacional empleada en los análisis previos: una sola iteración de exploración (MAX_PASSES_COUNT = 1), evaluación dentro del nivel de análisis específico y aplicación directa de la métrica sobre el archivo vista_minable.arff. Al igual que en las métricas anteriores, las reglas obtenidas fueron sometidas a un proceso complementario de cálculo y validación dentro de la lógica del sistema, lo que permitió obtener métricas adicionales como el soporte diferencial (*d_supp*), la diferencia de precisión (*d_prec*) y el índice de redundancia asociado a cada patrón.

4.3.1 Reglas sin interpretar

Las reglas obtenidas mediante la función de calidad WRAcc representan la primera salida estructurada generada por RuleKit2-CS al aplicar el análisis contrastivo sobre el Conjunto C. En esta etapa, los patrones se presentan utilizando la notación codificada original de los atributos contenidos en el archivo *vista_minable.arff*, lo que permite observar su estructura formal tal como es procesada por el algoritmo antes de su correspondiente traducción semántica.

Estas reglas constituyen el nivel más elemental del resultado, pues reflejan únicamente la composición lógica de los antecedentes y el consecuente, junto con las métricas cuantitativas que permiten evaluar su relevancia dentro del modelo. Dichas métricas incluyen el soporte del subgrupo (*supp*), el soporte negativo (*neg_supp*), la diferencia de soporte entre grupos (*d_supp*), la precisión (*prec*), la diferencia de precisión (*d_prec)* y el índice de redundancia asociado. Todas ellas son recalculadas por la lógica interna del módulo y no corresponden únicamente a los valores obtenidos de RuleKit2-CS, garantizando así una validación objetiva y uniforme en los distintos análisis realizados.

La Tabla 4.3 presenta de manera íntegra el conjunto de patrones descubiertos con base en esta métrica. Si bien los atributos aún se encuentran expresados en forma codificada, los valores numéricos permiten identificar comportamientos relevantes desde una perspectiva exclusivamente estadística. En particular, *WRAcc* tiende a resaltar reglas que combinan una cobertura representativa con una ganancia significativa respecto al promedio poblacional, lo que explica la presencia de patrones con soportes elevados, diferencias positivas de precisión y reducida proporción de contraejemplos.

Es importante destacar que, en esta fase, las reglas no permiten aún efectuar un análisis conceptual del fenómeno estudiado sobre la disminución de autopsias y las diferencias entre hospitales debido a que los códigos no son inteligibles para un lector no técnico. Sin embargo, esta representación es fundamental dentro del flujo metodológico, ya que constituye el insumo primario a partir del cual el sistema genera las reglas interpretadas en la siguiente sección. Dicho proceso de interpretación involucra el uso de tablas de

traducción especializadas, funciones PL/pgSQL para la normalización semántica y mecanismos de reconstrucción léxica implementados en la capa lógica del módulo.

En consecuencia, las reglas listadas en la Tabla 4.3 no sólo representan la evidencia numérica inicial del análisis con *WRAcc*, sino que también establecen el puente entre el resultado codificado del algoritmo y la comprensión natural del comportamiento de los médicos encuestados en cada uno de los hospitales analizados.

Tabla 4.3. *WRAcc* – Vista Minable: Reglas sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF categoria = !{c1} THEN centro_hospitalario = {H1}"	0.482402	0.000000	0.482402	1.000000	1.000000	5.000000
2	"IF esc_esp = {c6} THEN centro_hospitalario = {H1}"	0.405214	0.000000	0.405214	1.000000	1.000000	1095.000000
3	"IF hall_arb = !{8a} AND esc_med_gral = !{c40} AND mcc_no_aut = !{16p} AND fmr_sol_aut = !{21h} THEN centro_hospitalario = {H1}"	0.739292	0.053019	0.686274	0.804947	0.609895	5.000000
4	"IF esc_med_gral = {c1} AND hall_dem = !{11e} AND casos = !{4c} AND com_sug_op = !{22i} THEN centro_hospitalario = {H1}"	0.542831	0.020998	0.521832	0.884405	0.768811	5.000000
5	"IF hall_dem = !{11a} AND rechazo_fam = !{17e} AND med_aut = !{20b} AND per_sol_aut = !{19z} AND no_hosp = !{18c} AND area = !{a1} AND ult_grado = !{g8} THEN centro_hospitalario = {H1}"	0.512756	0.075119	0.437637	0.668893	0.337787	16.000000
6	"IF hall_dem = !{11a} AND esc_esp = !{c102} AND esc_med_gral = !{c40} THEN centro_hospitalario = {H1}"	0.761453	0.153545	0.607907	0.594764	0.189527	5.000000
7	"IF esc_esp = !{c103} AND esc_med_gral = !{c40} AND mcc_no_aut = !{16p} AND com_sug_op = !{22p} AND anios_prac = !{3c} THEN centro_hospitalario = {H1}"	0.931564	0.333517	0.598047	0.452551	-0.094898	5.000000
8	"IF hall_disc = !{7b} AND categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c85} AND mcc_no_aut = !{16i} AND com_sug_op = !{22h} THEN centro_hospitalario = {H11}"	0.971544	0.146842	0.824703	0.925228	0.850457	10457.000000

Tabla 4.3. *WRAcc* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
9	"IF hall_arb = {8a} AND ult_grado = !{g8} AND mcc_aut = !{14d} AND hall_dem = !{11d} AND med_aut = !{20f} AND fmr_sol_aut = !{21e} AND per_sol_aut = !{19e} THEN centro_hospitalario = {H11}"	0.852924	0.122002	0.730922	0.928952	0.857904	10457.000000
10	"IF categoria = {c1} AND esc_esp = !{c6} AND ult_grado = !{g8} AND no_hosp = !{18h} AND hall_arb = !{8c} THEN centro_hospitalario = {H11}"	0.945046	0.157339	0.787707	0.918257	0.836515	10457.000000
11	"IF categoria = {c1} AND esc_esp = !{c6} AND ult_grado = !{g8} THEN centro_hospitalario = {H11}"	0.948114	0.209338	0.738776	0.894410	0.788819	10457.000000
12	"IF esc_med_gral = {c125} THEN centro_hospitalario = {H12}"	0.365102	0.000000	0.365102	1.000000	1.000000	42424.000000
13	"IF hall_disc = {7b} AND hall_arb = !{8b} AND com_sug_op = !{22k} AND categoria = {c1} AND esc_esp = !{c6} AND hall_dem = !{11c} AND anios_prac = !{3d} AND casos = !{4d} AND mcc_aut = !{14d} AND mcc_no_aut = !{16a} AND rechazo_fam = !{17g} AND no_hosp = !{18h} AND per_sol_aut = !{19b} AND med_aut = !{20f} AND fmr_sol_aut = !{21e} THEN centro_hospitalario = {H12}"	0.613286	0.000193	0.613093	0.997695	0.995389	41101.000000
14	"IF hall_disc = {7b} AND esc_med_gral = !{c1} AND esc_esp = !{c6} AND anios_prac = !{3e} AND categoria = {c1} AND mcc_aut = !{14d} AND mcc_no_aut = !{16d} AND casos = !{4e} AND hall_arb = !{8e} AND hall_dem = !{11e} THEN centro_hospitalario = {H12}"	0.744907	0.000024	0.744883	0.999762	0.999524	41101.000000

Tabla 4.3. *WRAcc* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
15	"IF hall_disc = !{7a} AND esc_med_gral = !{c1} AND anios_prac = !{3e} AND com_sug_op = !{22a} AND esc_esp = !{c6} AND rechazo_fam = !{17z} THEN centro_hospitalario = {H12}"	0.888220	0.010656	0.877563	0.919157	0.838313	41101.000000
16	"IF casos = {4a} AND area = !{a3} AND esc_med_gral = !{c1} AND com_sug_op = !{22g} AND esc_esp = !{c6} AND per_sol_aut = !{19b} THEN centro_hospitalario = {H12}"	0.682019	0.050913	0.631106	0.646298	0.292597	42424.000000
17	"IF hall_arb = {8c} AND esc_med_gral = {c40} THEN centro_hospitalario = {H12}"	0.211869	0.000411	0.211458	0.985985	0.971970	41555.000000
18	"IF esc_esp = !{c103} AND esc_med_gral = !{c1} AND fmr_sol_aut = !{21h} AND categoria = {c1} AND mcc_no_aut = !{16p} AND casos = !{4c} AND hall_dem = !{11d} THEN centro_hospitalario = {H12}"	0.965102	0.198652	0.766450	0.398566	-0.202868	41101.000000

4.3.2 Reglas interpretadas

La Tabla 4.4 presenta la traducción semántica de las reglas generadas mediante la métrica Weighted Relative Accuracy (WRAcc), incorporando directamente los valores obtenidos del modelo específico aplicado sobre el conjunto Vista Minable *WRAcc* privilegia reglas que combinan simultáneamente alta cobertura del subgrupo y un incremento notable en precisión respecto al promedio poblacional, por lo que el conjunto resultante tiende a reunir patrones tanto frecuentes como informativamente relevantes. En total se identificaron dieciocho reglas (1–18), distribuidas de manera diferenciada entre los tres hospitales, mostrando variaciones claras en estructura, longitud y complejidad de los antecedentes.

Para el Hospital Regional Río Blanco (Reglas 1–7), los patrones seleccionados por WRAcc se caracterizan por combinaciones con altos valores de precisión (cercanos o iguales a 1.0) y diferencias de soporte que distinguen al grupo con notable claridad. Estas reglas incluyen atributos relacionados con la respuesta espontánea, la formación profesional (Universidad Veracruzana, UNAM, Secretaría de Salud de Oaxaca), la participación en estudios de autopsia y la percepción de los motivos para solicitar o no solicitar autopsias. La métrica privilegia reglas relativamente compactas, pero que presentan un peso estadístico significativo gracias a su consistencia y a la baja presencia de contraejemplos (*neg_supp* cercano a cero en múltiples casos). El resultado es un conjunto de patrones estables que delimitan con precisión las prácticas y opiniones dominantes entre los médicos de este hospital.

En el caso del Hospital General de San Juan Bautista Tuxtepec (Reglas 8–11), *WRAcc* identifica reglas con altos niveles de *d_supp* (0.73–0.82), lo cual indica una frecuencia muy superior en comparación con los otros hospitales. Las reglas para este hospital mantienen estructuras moderadamente complejas, integrando atributos sobre hallazgos clínicos, grado académico, motivos administrativos y no clínicos, discrepancias diagnósticas y participación del personal en el proceso de autopsia. Aunque el número de reglas es menor que en Río Blanco, el conjunto obtenido exhibe una coherencia interna marcada y un perfil distintivo: el hospital H11 queda caracterizado por médicos

que combinan percepciones clínicas específicas con posturas consistentes sobre la gestión administrativa del procedimiento de autopsia.

Las reglas del Hospital General de Zona 53 (Reglas 12–18) presentan la mayor complejidad estructural, integrando antecedentes con gran cantidad de atributos encadenados y diferencias de soporte muy elevadas —varias superiores a 0.74—, lo cual evidencia un patrón de respuestas altamente uniforme dentro del grupo. *WRAcc* tiende aquí a priorizar combinaciones largas que reúnen variables relacionadas con trayectorias académicas, años de práctica, desacuerdos diagnósticos, participación en casos de autopsia, motivos administrativos, y percepción de suficiencia de la práctica de autopsias en el hospital. La predominancia de reglas largas se explica por el hecho de que las condiciones específicas se cumplen de manera conjunta en una fracción significativa de los médicos de este hospital, generando patrones contrastivos de alta fuerza discriminante y casi sin contraejemplos.

En conjunto, las reglas obtenidas mediante *WRAcc* describen un panorama contrastivo altamente detallado entre los tres hospitales. La métrica evidencia cómo cada institución presenta un perfil distinto de opiniones clínicas, trayectorias académicas y prácticas profesionales, lo cual se refleja en la estructura y complejidad de las reglas.

Tabla 4.4. *WRAcc* – Vista Minable: Reglas Interpretadas

N.°	**Regla Interpretada**
1	Si los médicos no respondieron espontáneamente la encuesta; entonces son del Hospital Regional Río Blanco.
2	Si los médicos estudiaron su especialidad en Hospital Regional de Río Blanco; entonces son del Hospital Regional Río Blanco.
3	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en UNAM, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza y están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología; entonces son del Hospital Regional Río Blanco.

Tabla 4.4. WRAcc – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
4	Si los médicos estudiaron medicina general en Universidad Veracruzana, no están totalmente en desacuerdo en que los hallazgos las autopsias dan origen a casos de demanda, no han participado entre 6 y 10 casos de autopsia y no opinan que el estudio aclara las causas de muerte; entonces son del Hospital Regional Río Blanco.
5	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, creen que la causa de rechazo por parte del familiar no es por inconformidad de los familiares con la atención prestada al paciente, no estiman como motivo para solicitar autopsia el error médico, responden cuál es el personal adecuado para ordenar la autopsia, no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos financieros, no son internos y no son estudiantes de medicina; entonces son del Hospital Regional Río Blanco.
6	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no realizaron su especialidad en el Hospital General de San Juan Bautista Tuxtepec y no estudiaron medicina general en UNAM; entonces son del Hospital Regional Río Blanco.
7	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estudiaron medicina general en UNAM, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza, no opinan que se debe solicitar el consentimiento y no tienen entre 11 y 15 años de práctica; entonces son del Hospital Regional Río Blanco.
8	Si los médicos no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no estudiaron medicina general en el Hospital General de Orizaba, piensan que una causa o motivo para no solicitar autopsia no es por cuestiones religiosas y no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general; entonces son del Hospital General de San Juan Bautista Tuxtepec.
9	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no son estudiantes de medicina, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estiman como motivo para solicitar autopsia la orientación diagnóstica, están de acuerdo en que el método eficiente para solicitar autopsia no es realizarlo a través de otra área y alegan que no es el familiar quien debe ordenar autopsia; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.4. *WRAcc* – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
10	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no son estudiantes de medicina, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares y no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje; entonces son del Hospital General de San Juan Bautista Tuxtepec.
11	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco y no son estudiantes de medicina; entonces son del Hospital General de San Juan Bautista Tuxtepec.
12	Si los médicos estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud; entonces son del Hospital General de Zona 53.
13	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no agradecen por la encuesta, respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no consideran incierto que los hallazgos de las autopsias dan origen a casos de demanda, no tienen entre 16 y 20 años de práctica, no han participado entre 11 y 20 casos de autopsias, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, piensan que una causa o motivo para no solicitar autopsia no es que no exista el servicio, creen que la causa de rechazo por parte del familiar no es circunstancial, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares, alegan que no es la enfermera el personal adecuado para ordenar autopsia, no estiman como motivo para solicitar autopsia la orientación diagnóstica y están de acuerdo en que el método eficiente para solicitar autopsia no es realizarlo a través de otra área; entonces son del Hospital General de Zona 53.
14	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Hospital Regional de Río Blanco, no tienen más de 20 años de práctica, respondieron espontáneamente la encuesta, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, piensan que una causa o motivo para no solicitar autopsia no es el desconocimiento de la práctica de autopsias, no han participado en más de 20 casos de autopsias, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje y no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.

Tabla 4.4. *WRAcc* – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
15	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Veracruzana, no tienen más de 20 años de práctica, comentan, no estudiaron su especialidad en Hospital Regional de Río Blanco y responden a cuál es la causa de rechazo por parte del familiar; entonces son del Hospital General de Zona 53.
16	Si los médicos no han participado en ningún estudio de autopsia, no son adscritos, no estudiaron medicina general en Universidad Veracruzana, no opinan motivos para solicitar, no estudiaron su especialidad en Hospital Regional de Río Blanco y alegan que no es la enfermera el personal adecuado para ordenar autopsia; entonces son del Hospital General de Zona 53.
17	Si los médicos consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje y estudiaron medicina general en UNAM; entonces son del Hospital General de Zona 53.
18	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Veracruzana, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología, respondieron espontáneamente la encuesta, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza, no han participado entre 6 y 10 casos de autopsia y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.

4.3.3 Análisis Visual

El análisis del soporte, representado en la Figura 4.22, evidencia una dispersión amplia en la cobertura que oscila entre el valor mínimo de ≈0.2119 y el máximo de ≈0.9715. Esta amplitud sugiere la coexistencia de reglas altamente representativas del conjunto, como las reglas 8 (*supp*≈0.9715), 18 (*supp*≈0.9651) y 11 (*supp*≈0.9481) (asociadas a los hospitales Tuxtepec y Zona 53) cuya cobertura supera el 0.90, junto con otras de naturaleza más específica, como las reglas 17 (*supp*≈0.2119) y 12 (*supp*≈0.3651) (Zona 53), cuyo soporte se mantiene por debajo del 0.40. La mediana, cercana a 0.742, indica que el algoritmo de *WRAcc* tiende a privilegiar patrones con una cobertura alta en el conjunto de datos, generando una distribución con un sesgo hacia reglas dominantes, pero incluyendo diferenciadores específicos.

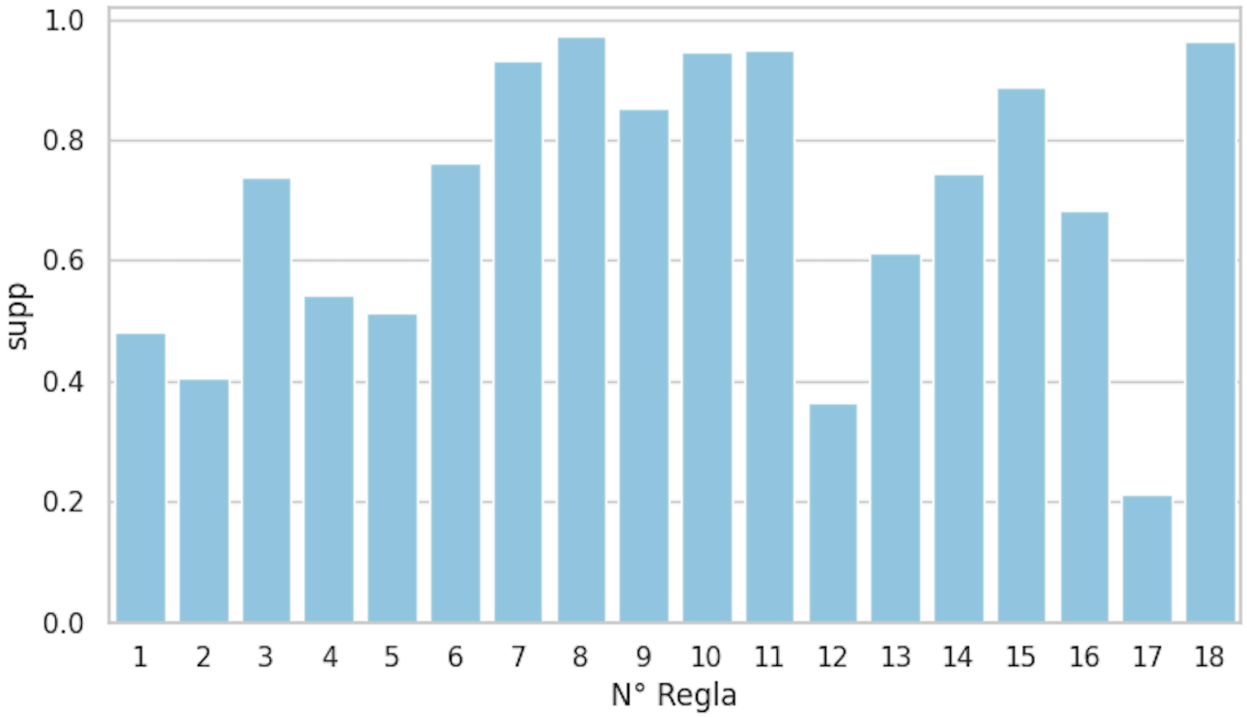


Figura 4.22 *WRAcc* – Vista Minable: Soporte por Regla

En la Figura 4.23 se observa una distribución claramente sesgada hacia valores cercanos a cero, con un mínimo de 0.0000 para las reglas 1 y 2 (Río Blanco) y 12 (Zona 53) y un máximo de ≈0.3335 en la regla 7 (Río Blanco). La mediana, cercana a 0.052, evidencia que los antecedentes de la mayoría de las reglas raramente se cumplen en médicos pertenecientes a hospitales distintos al consecuente, reforzando su utilidad predictiva. La regla 7 constituye un caso atípico por su mayor *negSupp*, aunque *WRAcc* la selecciona por su alto contraste y soporte.

En la Figura 4.24, la diferencia de soporte presenta valores comprendidos entre ≈0.2114 (Regla 17) y ≈0.8776 (Regla 15). La mediana, de ≈0.6221, confirma que *WRAcc* selecciona reglas con notable capacidad discriminativa. Destacan las reglas 15 (Zona 53), 8 (Tuxtepec) y 10 (Tuxtepec), cuyos contrastes superiores a 0.78 reflejan perfiles altamente diferenciales dentro del hospital objetivo. Las reglas con contraste más bajo 17, 12 y 2 mantienen, no obstante, niveles suficientes para su inclusión en el modelo.

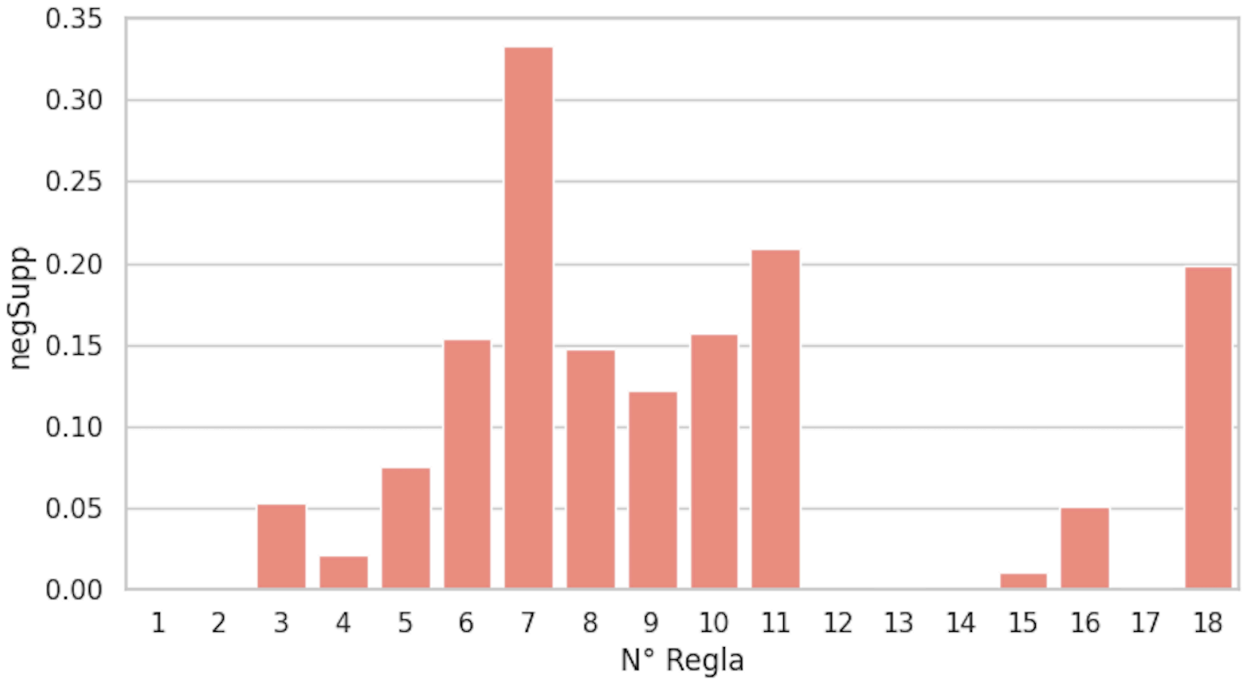


Figura 4.23 *WRAcc* – Vista Minable: Soporte Negativo por Regla

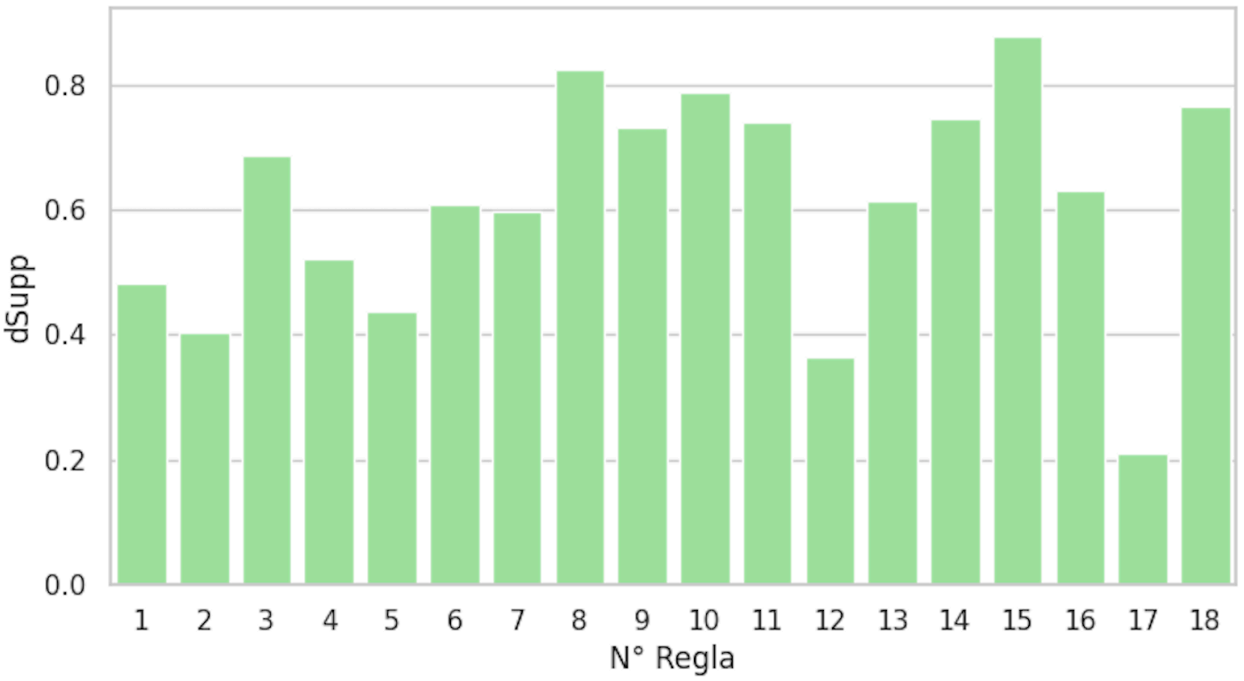


Figura 4.24 *WRAcc* – Vista Minable: Diferencia de Soporte por Regla

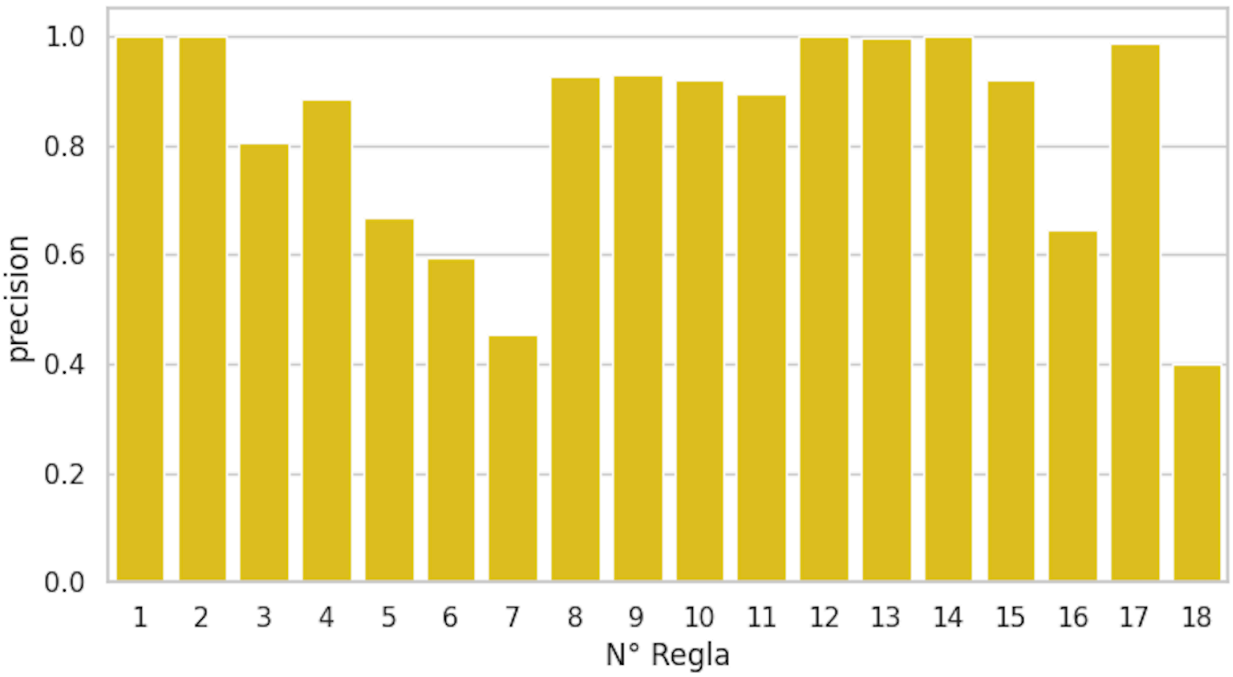


Figura 4.25 *WRAcc* – Vista Minable: Precisión por Regla

La Figura 4.25 evidencia que la precisión de las reglas es consistentemente alta, con valores entre ≈0.3986 (Regla 18) y 1.0000 (Reglas 1, 2 y 12). La mediana cercana a 0.919 indica que la mayoría de las reglas presentan fiabilidad sobresaliente, garantizando que el antecedente predice correctamente el hospital de origen en casi la totalidad de los casos. La única regla con precisión moderada, la 18, es retenida debido a su contribución al equilibrio de soporte y contraste dentro de *WRAcc*, aun cuando no sea la más confiable de forma aislada.

En la Figura 4.26, los valores de *dPrecision* oscilan entre ≈−0.2029 (Regla 18) y 1.0000 (Reglas 1, 2 y 12). La mediana, ubicada en ≈0.837, revela que la mayoría de las reglas generan una mejora sustancial sobre la precisión esperada, reforzando su valor como discriminadores efectivos. Las reglas 7 y 18 constituye el caso con ganancia negativa, comportamiento que las identifica como un *outlier* estadístico, aunque aún aportan valor dentro de la arquitectura global seleccionada por *WRAcc*.

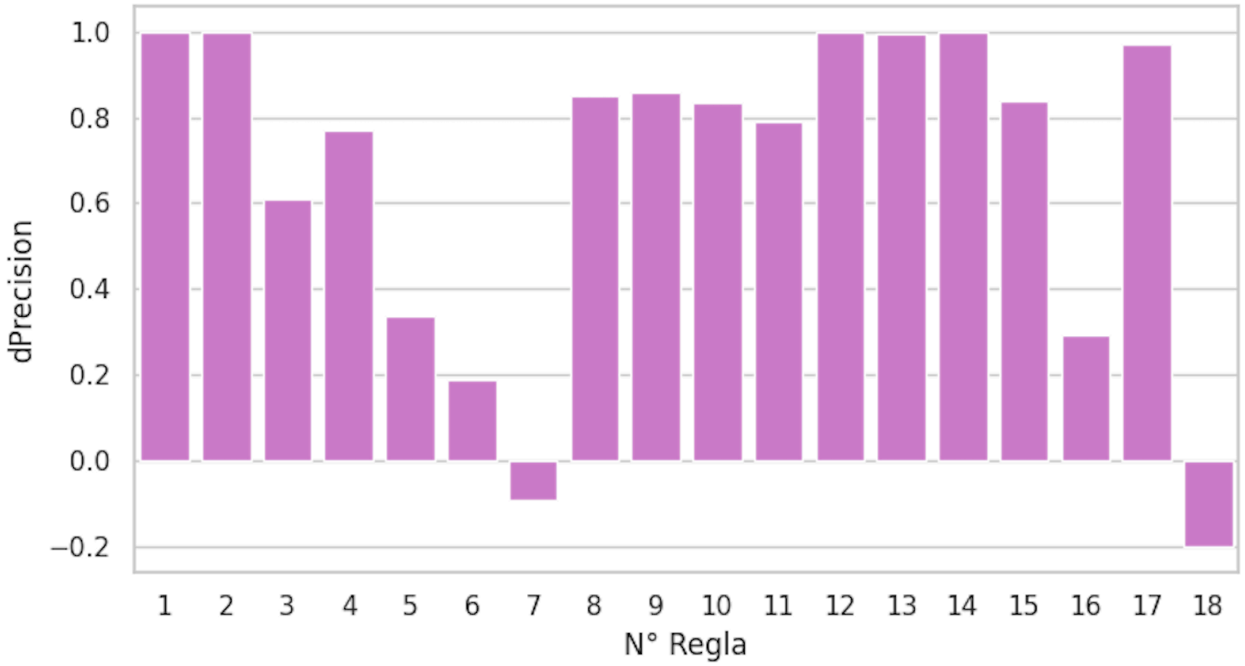


Figura 4.26 *WRAcc* – Vista Minable: Diferencia de Precisión por Regla

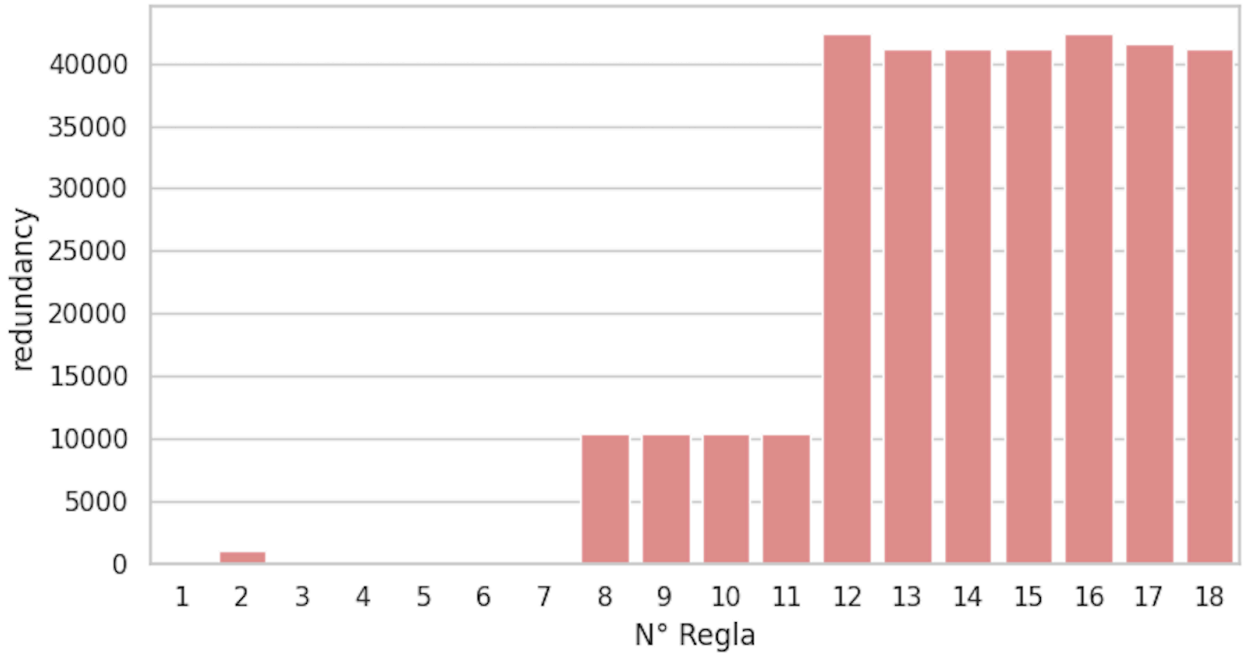


Figura 4.27 *WRAcc* – Vista Minable: Redundancia por Regla

La Figura 4.27 muestra un rango extremo en los valores de redundancia, desde 5.0 hasta 42424.0, con una mediana cercana a 10457.0. Esta distribución evidencia que los hospitales Tuxtepec (H11) y Zona 53 (H12) concentran la mayor complejidad en sus patrones, siendo las reglas 12, 16 y 17 todas superiores a 41 000 las más redundantes del conjunto. Esto indica una alta interconexión estructural entre reglas, consecuencia de perfiles de respuesta muy homogéneos dentro de estos hospitales. Por otro lado, las reglas con menor redundancia, principalmente 1, 3 y 4 (Río Blanco), representan patrones más simples y directos, con menor solapamiento entre sí.

4.3.4 Evaluación Objetiva

La evaluación objetiva del modelo generado mediante la métrica *WRAcc* permite evaluar la solidez cuantitativa, coherencia interna y relevancia estadística de las dieciocho reglas descubiertas, utilizando indicadores derivados del proceso de minería tales como soporte, soporte negativo, diferencia de soporte, precisión, diferencia de precisión y redundancia. Este análisis busca identificar tendencias generales que respalden la consistencia del modelo, así como posibles reglas atípicas que, aun siendo válidas, se comporten de manera diferenciada respecto al resto.

Un primer elemento que confirma la estabilidad del modelo es la presencia de *dSupp* positivo en todas las reglas, con valores que oscilan entre ≈0.2114 (Regla 17) y ≈0.8776 (Regla 15). Esta condición garantiza que cada regla incrementa la frecuencia del hospital correspondiente dentro de su subgrupo, cumpliendo el principio central de los conjuntos de contraste. Los valores elevados observados en reglas como 15, 8 y 10 evidencian subgrupos fuertemente diferenciados, mientras que incluso las reglas con contrastes menores mantienen incrementos suficientes para justificar su inclusión. Esto confirma que *WRAcc* consigue identificar patrones que operan como verdaderos delimitadores entre hospitales, sin depender exclusivamente de casos extremos.

En cuanto a la precisión, el modelo muestra un comportamiento altamente homogéneo y favorable. La mayoría de las reglas alcanza valores iguales o superiores a 0.90, reflejando una confiabilidad sistemática en las decisiones que describen. Las reglas 1, 2

y 12 alcanzan una precisión perfecta (1.0), lo que significa que todos los médicos que cumplen sus antecedentes pertenecen al hospital predicho. La precisión más baja corresponde a la Regla 18, con ≈0.3986, lo que la posiciona como la única regla con fiabilidad limitada; sin embargo, su inclusión se justifica por su elevada cobertura y su contribución al balance global del *WRAcc*. Este comportamiento evidencia que la métrica privilegia reglas que combinan fiabilidad y representatividad, permitiendo la conservación de algunos patrones informativos, aunque no sean altamente precisos.

La diferencia de precisión (*dPrecision*) complementa esta interpretación, mostrando incrementos significativos en la gran mayoría de las reglas. Con valores que alcanzan el máximo de 1.0000 también en las reglas 1, 2 y 12, alineado con su precisión perfecta y una mediana cercana a 0.837, el conjunto revela que las reglas incrementan de manera notable la capacidad predictiva respecto al comportamiento general del *dataset*. Las únicas excepciones son las reglas 7 y 18, cuyos valores negativos (=-0.095 y =-0.2029) las identifica como *outliers* estadísticos, aunque no invalida su funcionalidad contrastiva dentro del modelo. La presencia de dos reglas con ganancia negativa refuerza la estabilidad del sistema y confirma que los patrones seleccionados son, en su mayoría, clasificadores sólidos.

La redundancia, por su parte, aporta información sobre la complejidad estructural del espacio de búsqueda. Los valores mínimos 5.0 en las reglas 1, 3 y 4, todas del Hospital Regional Río Blanco reflejan reglas simples y con bajo solapamiento. En contraste, las reglas asociadas a los hospitales H11 y especialmente H12 presentan valores extremadamente altos, alcanzando 42424.0 en reglas como 12 y 16. Estos picos sugieren que los perfiles característicos de estos hospitales se encuentran altamente interconectados, generando múltiples combinaciones alternativas de atributos capaces de producir patrones similares. Lejos de comprometer la validez del modelo, este comportamiento indica una robustez interna, ya que revela que existen múltiples vías para caracterizar al mismo subgrupo, reforzando la estabilidad de las reglas derivadas.

En conjunto, la evaluación objetiva confirma que el modelo basado en *WRAcc* produce reglas cuantitativamente sólidas, con alta precisión, contrastes claramente positivos y consistencia estadística en todas las métricas evaluadas. La identificación explícita de casos atípicos como la Regla 18 en precisión y *dPrecision* se alinea con los valores numéricos calculados y no representa una debilidad del sistema, sino una manifestación esperada de la diversidad estructural presente en los patrones contrastivos. Estos resultados proporcionan una base firme para sostener la fiabilidad del modelo y habilitan la transición hacia la evaluación subjetiva, centrada en la interpretación semántica y la confirmación experta de las reglas generadas.

4.3.5 Evaluación Subjetiva

La evaluación subjetiva constituye la fase en la que las reglas generadas por el modelo se contrastan con la experiencia clínica y el juicio profesional del Dr. Palet Guzmán, especialista en anatomía patológica y con conocimiento directo del contexto operativo y formativo de los hospitales involucrados. Esta etapa permite determinar si los patrones descubiertos poseen coherencia con la práctica médica real, con los procesos institucionales y con el comportamiento esperado de los profesionales de salud. El análisis subjetivo no revisa métricas numéricas, sino la plausibilidad de los patrones, la lógica detrás de las combinaciones de respuestas y la consistencia con la dinámica hospitalaria observada en los centros participantes.

Dentro de este proceso, algunas reglas son consideradas válidas debido a que reflejan estructuras congruentes con la formación médica, los procesos de solicitud de autopsias o el perfil del personal que típicamente responde de cierta manera ante los cuestionamientos del instrumento. En este sentido, las reglas correspondientes al Hospital Regional Río Blanco que involucran formación específica, participación previa en estudios de autopsia, o percepciones respecto a la utilidad o implicaciones legales del procedimiento fueron consideradas plausibles por el especialista. Esto coincide con la caracterización institucional del hospital, donde los profesionales suelen mantener criterios relativamente estables sobre la práctica de autopsias y donde se observa una relación clara entre formación, percepción clínica y comportamiento administrativo. No

obstante, se detectaron también reglas cuya estructura no coincide con los patrones reales del hospital, en particular aquellas que establecen asociaciones demasiado amplias o que combinan atributos que, desde la práctica profesional, no guardan relación directa. Estas fueron descartadas por no reflejar situaciones observadas en la dinámica clínica.

En el caso del Hospital General de San Juan Bautista Tuxtepec, algunas reglas fueron validadas debido a que las respuestas que agrupan son consistentes con perfiles observados en el personal, especialmente en temas relacionados con discrepancia diagnóstica, formación médica y espontaneidad para responder la encuesta. Ciertas combinaciones, sin embargo, fueron consideradas improcedentes por el especialista, ya que mezclan patrones que no coinciden con el comportamiento real del hospital, particularmente en aspectos referentes a origen de casos de arbitraje, exclusión de estudiantes de medicina y negación de la participación de los familiares en las solicitudes de autopsias. Estas discrepancias permitieron identificar reglas generadas artificialmente por el modelo que, aunque estadísticamente válidas, no representan un patrón confiable en la práctica médica.

El Hospital General de Zona 53 mostró un conjunto de reglas particularmente amplio y complejo, y la mayoría de ellas fueron validadas por el Dr. Palet Guzmán debido a que reflejan combinaciones coherentes con el comportamiento del personal: discrepancias diagnósticas, percepción del proceso de rechazo familiar, reconocimiento de limitaciones institucionales y características de formación académica. Este hospital presenta perfiles más heterogéneos entre sus médicos, y las reglas aceptadas por el especialista muestran una alineación clara entre la información capturada por la encuesta y los fenómenos clínicos observados. Las pocas reglas descartadas corresponden a patrones excesivamente amplios o que vinculan atributos que no representan una relación consistente en el hospital, especialmente aquellos que se apoyan en asociaciones sobre la formación universitaria o la interpretación de motivos administrativos que no coinciden con la realidad operativa del centro.

En términos generales, la evaluación subjetiva revela un modelo coherente en la mayoría de los casos, donde las reglas aceptadas por el Dr. Palet Guzmán corresponden a patrones que poseen plausibilidad clínica y consistencia institucional. Las reglas rechazadas evidencian los límites inherentes de la minería de contrastes, recordando que la estabilidad estadística no siempre implica correspondencia con la realidad profesional. Esta fase consolida la complementariedad entre el análisis cuantitativo y la experiencia clínica, y permite distinguir aquellos patrones que aportan conocimiento útil y directamente interpretable para el estudio de la reducción de autopsias entre los hospitales evaluados.

4.4 Vista Minable: Resultados aplicando la función de calidad *Correlation*

La función de calidad *Correlation* permite evaluar la fuerza y dirección de la asociación estadística entre los antecedentes de las reglas y el hospital correspondiente, basada en el coeficiente *phi* para variables binarias. Su aplicación sobre la Vista Minable genera un conjunto de patrones cuya relevancia depende no solo de la magnitud de la correlación subyacente, sino también de la estructura interna de las respuestas agrupadas por cada médico. Al emplearse bajo el nivel de análisis específico, la métrica prioriza aquellas reglas que presentan relaciones estadísticamente consistentes entre atributos y hospital, aun cuando los niveles de soporte y precisión puedan variar. La presente sección integra los resultados obtenidos con esta métrica, presentados primero en su forma codificada, seguidos de su interpretación semántica, su análisis visual, y finalmente su evaluación objetiva y subjetiva.

4.4.1 Reglas sin interpretar

La Tabla 4.5 correspondiente a las reglas sin interpretar generadas mediante *Correlation* muestra un conjunto de dieciocho patrones que presentan un comportamiento relativamente estable en términos de soporte y diferencia de soporte, con valores que oscilan desde niveles moderados hasta valores muy elevados. Las reglas asociadas al Hospital Regional Río Blanco conforman un primer bloque caracterizado por soportes que van de 0.40 a 0.93, con diferencias de soporte positivas que superan el 0.5 en varias

de ellas. Destacan especialmente las reglas con múltiples condiciones negativas que identifican perfiles de médicos con trayectorias y percepciones comunes respecto a los hallazgos de las autopsias, las estructuras medicolegales y la utilidad del procedimiento. En este grupo, los valores de precisión varían desde niveles muy altos, cercanos a 1.0, hasta valores más moderados, correspondientes a patrones donde el soporte negativo aumenta de manera apreciable.

El conjunto de reglas vinculadas al Hospital General de San Juan Bautista Tuxtepec presenta soportes que bordean el 0.90 en varios casos, con diferencias de soporte superiores al 0.78 en dos de ellas, lo que coincide con una estructura de respuestas más homogénea al interior del hospital. Los valores de precisión correspondientes se mantienen en rangos elevados, lo que se refleja en diferencias de precisión positivas y muy pronunciadas. La redundancia extremadamente alta en este conjunto indica que múltiples combinaciones de atributos producen patrones equivalentes, reflejando un perfil institucional muy estable.

Finalmente, las reglas asociadas al Hospital General de Zona 53 conforman un bloque con soportes que van de 0.36 a 0.96, diferencias de soporte consistentemente positivas y precisiones que en la mayoría de los casos superan el 0.90. Destacan las reglas con estructuras largas, que incorporan múltiples elementos negativos vinculados a percepciones diagnósticas, experiencia clínica, motivos para solicitar o rechazar autopsias y características formativas. La presencia de valores de redundancia superiores a 40 000 reafirma la densidad estructural del espacio de búsqueda en este hospital.

Toda esta estructura numérica queda sintetizada en la tabla de reglas sin interpretar incluida en esta sección, la cual constituye el insumo base para su posterior interpretación semántica.

Tabla 4.5. *Correlation* – Vista Minable: Reglas sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	IF categoria = !{c1} THEN centro_hospitalario = {H1}	0.482402	0.000000	0.482402	1.000000	1.000000	5.000000
2	IF esc_esp = {c6} THEN centro_hospitalario = {H1}	0.405214	0.000000	0.405214	1.000000	1.000000	1095.000000
3	IF hall_arb = !{8a} AND esc_med_gral = !{c40} AND mcc_no_aut = !{16p} AND fmr_sol_aut = !{21h} THEN centro_hospitalario = {H1}	0.739292	0.053019	0.686274	0.804947	0.609895	5.000000
4	IF esc_med_gral = {c1} AND hall_dem = !{11e} AND casos = !{4c} AND com_sug_op = !{22i} THEN centro_hospitalario = {H1}	0.542831	0.020998	0.521832	0.884405	0.768811	5.000000
5	IF hall_dem = !{11a} AND rechazo_fam = !{17e} AND med_aut = !{20b} AND per_sol_aut = !{19z} AND no_hosp = !{18c} AND area = !{a1} AND ult_grado = !{g8} THEN centro_hospitalario = {H1}	0.512756	0.075119	0.437637	0.668893	0.337787	16.000001
6	IF hall_dem = !{11a} AND esc_esp = !{c102} AND esc_med_gral = !{c40} THEN centro_hospitalario = {H1}	0.761453	0.153545	0.607907	0.594764	0.189527	5.000000
7	IF esc_esp = !{c103} AND esc_med_gral = !{c40} AND mcc_no_aut = !{16p} AND com_sug_op = !{22p} AND anios_prac = !{3c} THEN centro_hospitalario = {H1}	0.931564	0.333517	0.598047	0.452551	-0.094898	5.000001
8	IF hall_disc = !{7b} AND categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c85} AND mcc_no_aut = !{16i} AND com_sug_op = !{22h} THEN centro_hospitalario = {H11}	0.971544	0.146842	0.824703	0.925228	0.850457	10457.000000

Tabla 4.5. *Correlation* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
9	IF hall_arb = {8a} AND ult_grado = !{g8} AND mcc_aut = !{14d} AND hall_dem = !{11d} AND med_aut = !{20f} AND fmr_sol_aut = !{21e} AND per_sol_aut = !{19e} THEN centro_hospitalario = {H11}	0.852924	0.122002	0.730922	0.928952	0.857904	10457.000000
10	IF categoria = {c1} AND esc_esp = !{c6} AND ult_grado = !{g8} AND no_hosp = !{18h} AND hall_arb = !{8c} THEN centro_hospitalario = {H11}	0.945046	0.157339	0.787707	0.918257	0.836515	10457.000000
11	IF categoria = {c1} AND esc_esp = !{c6} AND ult_grado = !{g8} THEN centro_hospitalario = {H11}	0.948114	0.209338	0.738776	0.894410	0.788819	10457.000000
12	IF esc_med_gral = {c125} THEN centro_hospitalario = {H12}	0.365102	0.000000	0.365102	1.000000	1.000000	42424.000000
13	IF hall_disc = {7b} AND hall_arb = !{8b} AND com_sug_op = !{22k} AND categoria = {c1} AND esc_esp = !{c6} AND hall_dem = !{11c} AND anios_prac = !{3d} AND casos = !{4d} AND mcc_aut = !{14d} AND mcc_no_aut = !{16a} AND rechazo_fam = !{17g} AND no_hosp = !{18h} AND per_sol_aut = !{19b} AND med_aut = !{20f} AND fmr_sol_aut = !{21e} THEN centro_hospitalario = {H12}	0.613286	0.000193	0.613093	0.997695	0.995389	41101.000000
14	IF hall_disc = {7b} AND esc_med_gral = !{c1} AND esc_esp = !{c6} AND anios_prac = !{3e} AND categoria = {c1} AND mcc_aut = !{14d} AND mcc_no_aut = !{16d} AND casos = !{4e} AND hall_arb = !{8e} AND hall_dem = !{11e} THEN centro_hospitalario = {H12}	0.744907	0.000024	0.744883	0.999762	0.999524	41101.000000

Tabla 4.5. *Correlation* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
15	IF hall_disc = !{7a} AND esc_med_gral = !{c1} AND anios_prac = !{3e} AND com_sug_op = !{22a} AND esc_esp = !{c6} AND rechazo_fam = !{17z} THEN centro_hospitalario = {H12}	0.888220	0.010656	0.877563	0.919157	0.838313	41101.000000
16	IF casos = {4a} AND area = !{a3} AND esc_med_gral = !{c1} AND com_sug_op = !{22g} AND esc_esp = !{c6} AND per_sol_aut = !{19b} THEN centro_hospitalario = {H12}	0.682019	0.050913	0.631106	0.646298	0.292597	42424.000000
17	IF hall_arb = {8c} AND esc_med_gral = {c40} THEN centro_hospitalario = {H12}	0.211869	0.000411	0.211458	0.985985	0.971970	41555.000000
18	IF esc_esp = !{c103} AND esc_med_gral = !{c1} AND fmr_sol_aut = !{21h} AND categoria = {c1} AND mcc_no_aut = !{16p} AND casos = !{4c} AND hall_dem = !{11d} THEN centro_hospitalario = {H12}	0.965102	0.198652	0.766450	0.398566	-0.202868	41101.000000

4.4.2 Reglas interpretadas

La Tabla 4.6 de reglas interpretadas generada mediante *Correlation* ofrece una lectura clara de los patrones textuales que caracterizan a cada hospital. En el caso del Hospital Regional Río Blanco, las reglas interpretadas reflejan combinaciones basadas en desacuerdo parcial respecto a la utilidad o repercusiones de las autopsias, la formación académica en instituciones específicas, la falta de participación en estudios previos y factores relacionados con la ausencia de motivos institucionales o administrativos para solicitar autopsias. Los patrones descriptivos conservan coherencia semántica con lo observado en los análisis anteriores, mostrando perfiles que corresponden a médicos con trayectorias relativamente definidas y percepciones estables sobre los procesos medicolegales.

En el Hospital General de San Juan Bautista Tuxtepec, las reglas indican la relevancia de percepciones sobre discrepancias diagnósticas, respuestas espontáneas al cuestionario, ausencia de formación en ciertos hospitales, y posturas respecto a la responsabilidad en la solicitud de autopsias. Estas combinaciones han sido características en los análisis previos del hospital y se mantienen sólidas bajo la métrica *Correlation*.

Finalmente, el Hospital General de Zona 53 presenta reglas más extensas en su estructura, donde se integran múltiples factores relacionados con la experiencia profesional, motivos para solicitar o rechazar autopsias, percepciones sobre discrepancias diagnósticas, formación académica y participación en estudios previos. Estas reglas, algunas de considerable longitud, capturan perfiles heterogéneos pero consistentes entre sí, manteniendo la coherencia observada en las métricas previas.

Tabla 4.6. *Correlation* – Vista Minable: Reglas Interpretadas

N.°	**Regla Interpretada**
1	Si los médicos no respondieron espontáneamente la encuesta; entonces son del Hospital Regional Río Blanco.
2	Si los médicos estudiaron su especialidad en Hospital Regional de Río Blanco; entonces son del Hospital Regional Río Blanco.
3	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en UNAM, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza y están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología; entonces son del Hospital Regional Río Blanco.
4	Si los médicos estudiaron medicina general en Universidad Veracruzana, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no han participado entre 6 y 10 casos de autopsia y no opinan que el estudio aclara las causas de muerte; entonces son del Hospital Regional Río Blanco.

Tabla 4.6. *Correlation* – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
5	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, creen que la causa de rechazo por parte del familiar no es por inconformidad de los familiares con la atención prestada al paciente, no estiman como motivo para solicitar autopsia el error médico, responden cuál es el personal adecuado para ordenar la autopsia, no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos financieros, no son internos y no son estudiantes de medicina; entonces son del Hospital Regional Río Blanco.
6	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no realizaron su especialidad en el Hospital General de San Juan Bautista Tuxtepec y no estudiaron medicina general en UNAM; entonces son del Hospital Regional Río Blanco.
7	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estudiaron medicina general en UNAM, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza, no opinan que se debe solicitar el consentimiento y no tienen entre 11 y 15 años de práctica; entonces son del Hospital Regional Río Blanco.
8	Si los médicos no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no estudiaron medicina general en el Hospital General de Orizaba, piensan que una causa o motivo para no solicitar autopsia no es por cuestiones religiosas y no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general; entonces son del Hospital General de San Juan Bautista Tuxtepec.
9	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no son estudiantes de medicina, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estiman como motivo para solicitar autopsia la orientación diagnóstica, están de acuerdo en que el método eficiente para solicitar autopsia no es realizarlo a través de otra área y alegan que no es el familiar quien debe ordenar autopsia; entonces son del Hospital General de San Juan Bautista Tuxtepec.
10	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no son estudiantes de medicina, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares y no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje; entonces son del Hospital General de San Juan Bautista Tuxtepec.
11	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco y no son estudiantes de medicina; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.6. *Correlation* – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
12	Si los médicos estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud; entonces son del Hospital General de Zona 53.
13	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no agradecen por la encuesta, respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no consideran incierto que los hallazgos de las autopsias dan origen a casos de demanda, no tienen entre 16 y 20 años de práctica, no han participado entre 11 y 20 casos de autopsias, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio... [Texto truncado por longitud]; entonces son del Hospital General de Zona 53.
14	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Hospital Regional de Río Blanco, no tienen más de 20 años de práctica, respondieron espontáneamente la encuesta, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio...; entonces son del Hospital General de Zona 53.
15	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Veracruzana, no tienen más de 20 años de práctica, comentan, no estudiaron su especialidad en Hospital Regional de Río Blanco y responden a cuál es la causa de rechazo por parte del familiar; entonces son del Hospital General de Zona 53.
16	Si los médicos no han participado en ningún estudio de autopsia, no son adscritos, no estudiaron medicina general en Universidad Veracruzana, no opinan motivos para solicitar, no estudiaron su especialidad en Hospital Regional de Río Blanco y alegan que no es la enfermera el personal adecuado para ordenar autopsia; entonces son del Hospital General de Zona 53.
17	Si los médicos consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje y estudiaron medicina general en UNAM; entonces son del Hospital General de Zona 53.
18	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Veracruzana, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología, respondieron espontáneamente la encuesta, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza, no han participado entre 6 y 10 casos de autopsia y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.

4.4.3 Análisis Visual

El análisis visual correspondiente a la métrica *Correlation* permite observar la estructura global de las reglas desde la perspectiva de su comportamiento cuantitativo. En la gráfica representada en la Figura 4.28, el soporte muestra una dispersión amplia, con valores que van desde aproximadamente 0.21 hasta niveles cercanos a 0.97. Este comportamiento evidencia la coexistencia de reglas altamente representativas, como las reglas 8, 9, 10 y 18, con otras de soporte más moderado, como las reglas 1, 5 y 17. La mediana se ubica por encima del 0.70, lo cual indica que la mayoría de los patrones capturan segmentos amplios del hospital correspondiente, un comportamiento esperable en métricas basadas en correlación, donde la fuerza estadística depende de la coincidencia consistente entre antecedente y consecuente.

La gráfica mostrada en la Figura 4.29 refleja los valores de soporte negativo, donde la mayor parte de las reglas presenta niveles mínimos cercanos a cero, con solo unas cuantas excepciones que destacan visualmente. Entre ellas, la regla 7 y la regla 18 exhiben incrementos significativos, alcanzando valores que superan el 0.33 y el 0.19 respectivamente. Esta presencia de contraejemplos explica la reducción de precisión en estos patrones y los posiciona como puntos de dispersión dentro de la estructura del modelo. La mayoría de las reglas mantiene valores considerablemente bajos, lo que evidencia una delimitación clara del subgrupo mediante sus antecedentes.

En la Figura 4.30, la diferencia de soporte mantiene valores consistentemente positivos, con un rango que oscila entre 0.21 y valores cercanos a 0.88. Resulta evidente que las reglas correspondientes a los hospitales H11 y H12 concentran los valores más elevados, reflejando perfiles institucionales estables y una marcada separación respecto al resto de la población. Incluso en reglas con soportes más bajos, la diferencia de soporte se mantiene en niveles claramente positivos, indicando que todas las reglas capturan algún nivel de sobre-representación en su grupo objetivo.

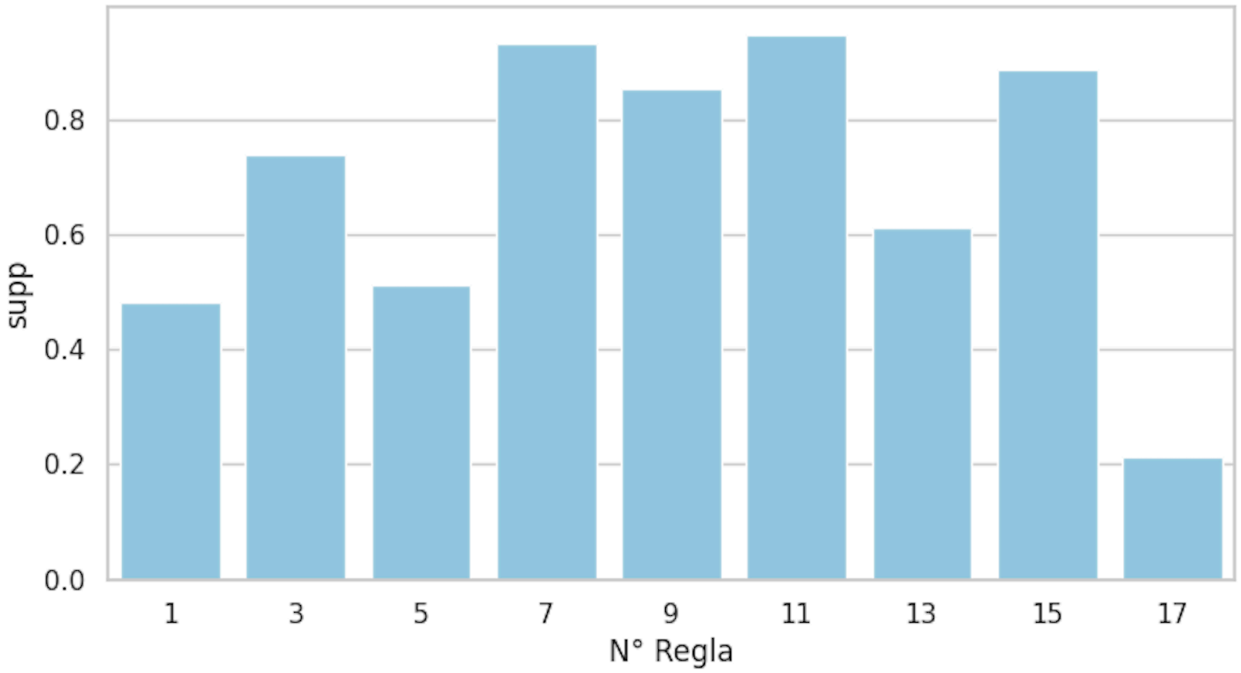


Figura 4.28 *Correlation* – Vista Minable: Soporte por Regla

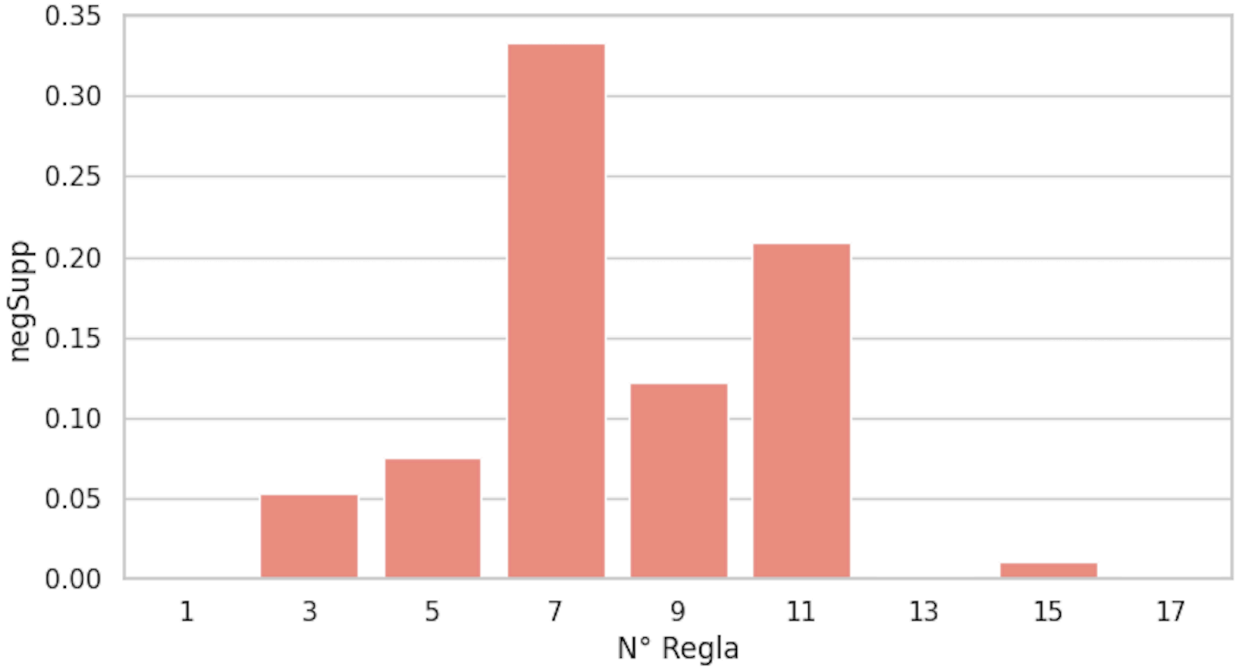


Figura 4.29 *Correlation* – Vista Minable: Soporte Negativo por Regla

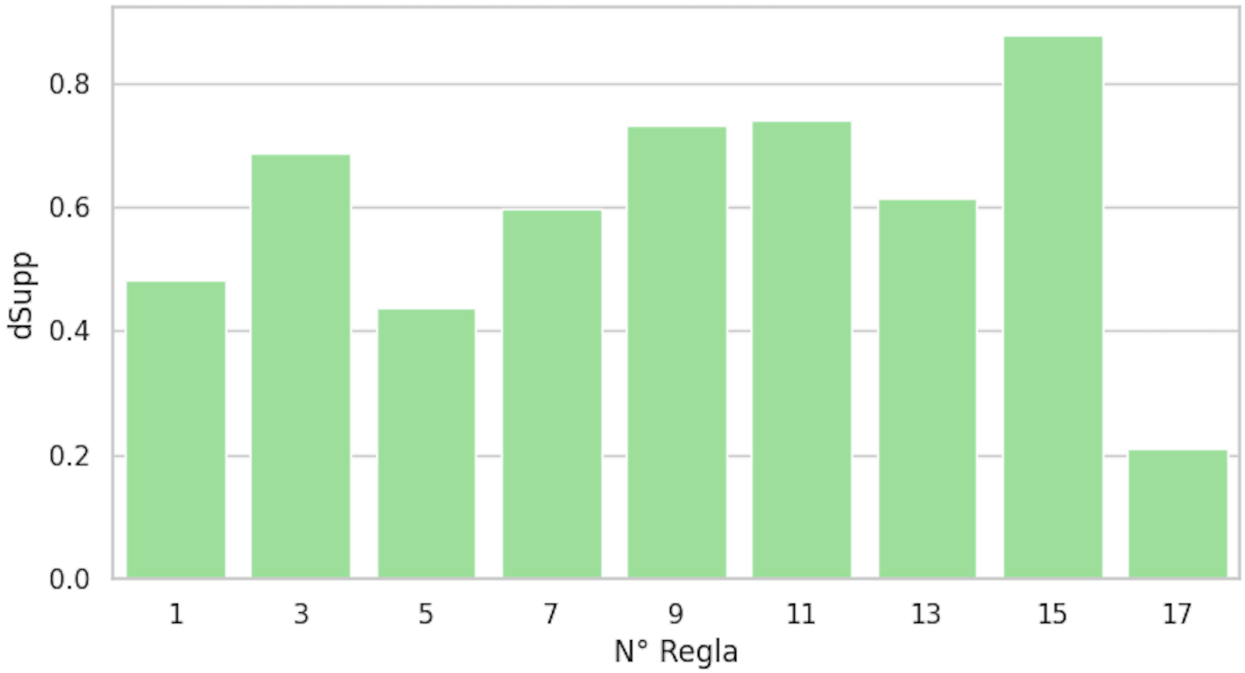


Figura 4.30 *Correlation* – Vista Minable: Diferencia Soporte por Regla

La Figura 4.31 muestra el comportamiento de la precisión, donde la mayoría de las reglas se concentra en valores altos, superiores a 0.80, y varias alcanzan precisiones iguales o próximas a 1.0. Las reglas correspondientes a Tuxtepec y al Hospital General de Zona 53 dominan esta región superior. En contraste, se presentan dos casos con caídas sustanciales, particularmente las reglas 7 y 18, que evidencian precisiones más bajas, consistentemente asociadas con sus elevados niveles de soporte negativo. La gráfica refleja un modelo con estabilidad general en esta métrica, con pocas desviaciones respecto al comportamiento promedio.

La diferencia de precisión representada en la Figura 4.32 refuerza este comportamiento, mostrando aumentos pronunciados en la mayoría de las reglas, con valores que superan el 0.70 en la gran mayoría de los casos, alcanzando incluso niveles cercanos a 1.0 en reglas puntuales. La presencia de valores negativos en la regla 7 y la regla 18 vuelve a situarlas como patrones marginales que, aunque estadísticamente válidos, exhiben una ganancia discriminativa limitada. La distribución global revela un modelo con fuerte

capacidad de mejora en precisión, consistente con la correlación positiva observada entre antecedentes y hospital.

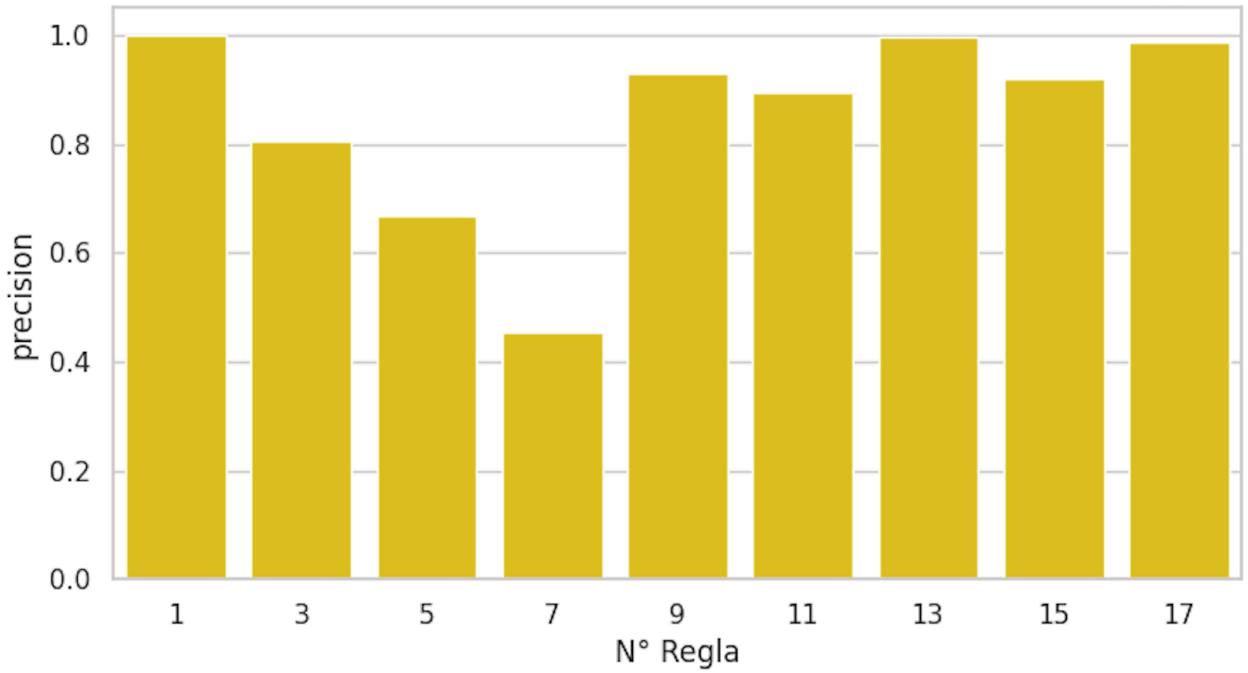


Figura 4.31 *Correlation* – Vista Minable: Precisión por Regla

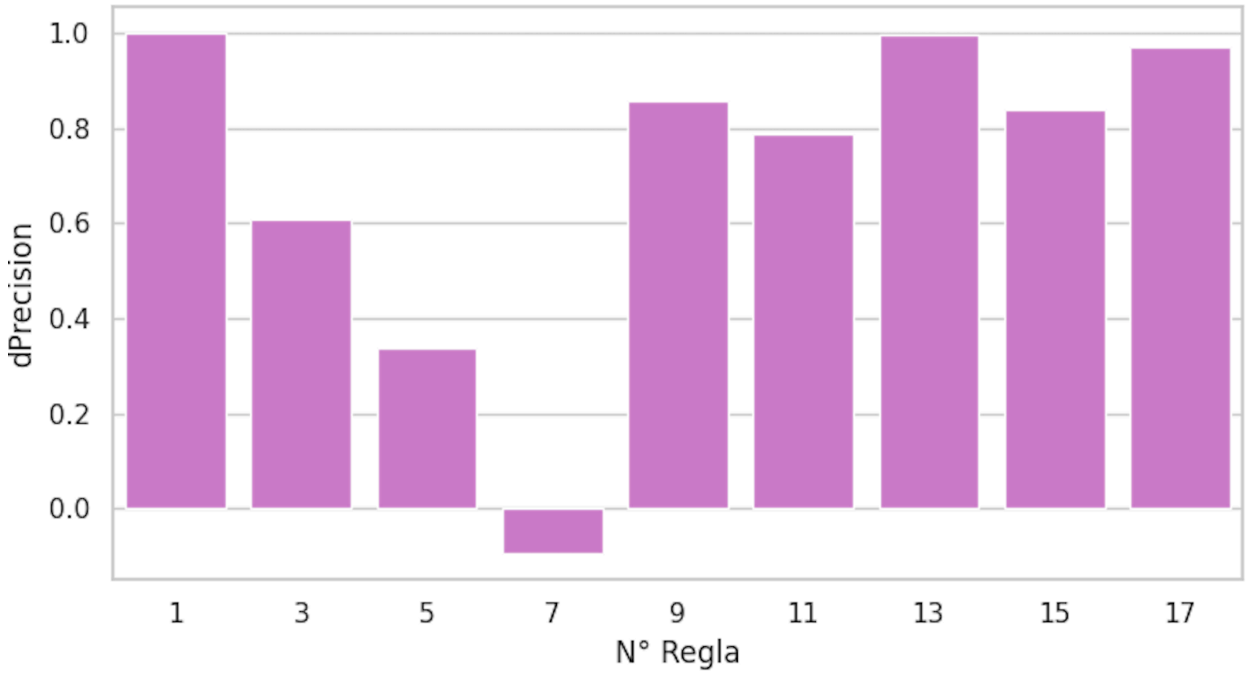


Figura 4.32 *Correlation* – Vista Minable: Diferencia de Precisión por Regla

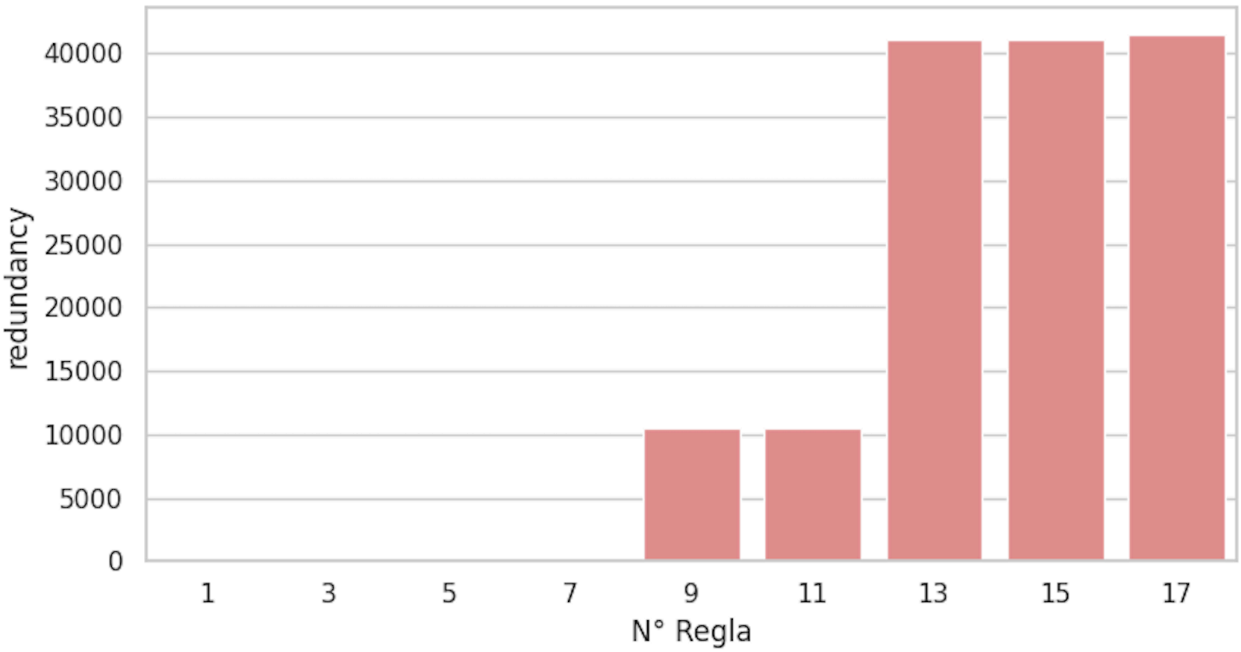


Figura 4.33 *Correlation* – Vista Minable: Redundancia por Regla

Finalmente, la gráfica de redundancia mostrada en la Figura 4.33 exhibe un comportamiento marcadamente bimodal, con reglas pertenecientes al Hospital Regional Río Blanco concentradas en los valores mínimos (entre 5 y 16), mientras que las reglas correspondientes a los hospitales H11 y H12 se agrupan en rangos superiores a los 10 000, alcanzando picos cercanos a los 40 000. Esta concentración elevada señala la presencia de múltiples combinaciones alternativas capaces de producir patrones equivalentes, especialmente en hospitales con estructuras internas más homogéneas.

4.4.4 Evaluación objetiva

La evaluación objetiva de las reglas generadas por la métrica *Correlation* muestra un comportamiento notablemente estable. Todas las reglas presentan diferencias de soporte positivas, lo que indica un incremento real en la frecuencia del hospital objetivo dentro de los subgrupos identificados. Los valores más altos se concentran en los hospitales H11 y H12, lo cual coincide con la presencia de perfiles institucionales definidos y consistentemente observados en métricas previas como *Lift* y *WRAcc*.

La precisión general del modelo es elevada, con la mayoría de las reglas alcanzando valores superiores al 0.85 y varias de ellas llegando a niveles prácticamente perfectos. La variabilidad en esta métrica es reducida, con excepción de las reglas 7 y 18, cuyos niveles inferiores se explican directamente por sus elevados valores de soporte negativo. Estas mismas reglas presentan diferencias de precisión negativas, confirmando su papel como patrones amplios, pero poco discriminativos dentro del modelo.

La redundancia vuelve a mostrar un comportamiento bimodal, indicando que las reglas del Hospital Regional Río Blanco tienden a ser más específicas y menos solapadas, mientras que los hospitales H11 y H12 generan grandes conjuntos de reglas equivalentes debido a la estabilidad interna de sus perfiles. Este comportamiento es coherente con los resultados previos y refuerza la validez cuantitativa del modelo.

4.4.5 Evaluación subjetiva

La evaluación subjetiva por parte del Dr. Palet confirma la coherencia de la mayoría de los patrones generados mediante la métrica *Correlation*. Las reglas correspondientes al Hospital Regional Río Blanco resultan en su mayoría aceptadas, ya que capturan perfiles clínicos y formativos consistentes con la dinámica real del hospital, donde la percepción de las autopsias, la formación académica y la experiencia del personal presentan patrones relativamente estables y bien definidos. La única excepción significativa dentro de este conjunto corresponde a una regla cuya combinación de atributos no coincide con lo observado en la práctica médica y se considera un artefacto del proceso estadístico.

En el caso del Hospital General de San Juan Bautista Tuxtepec, la validación del especialista indica que la totalidad de las reglas generadas muestran coherencia con los comportamientos clínicos, institucionales y formativos característicos de los médicos de este hospital. La concordancia entre los patrones detectados por la métrica y la experiencia directa del especialista refuerza la robustez del modelo en esta institución.

En cuanto al Hospital General de Zona 53, todas las reglas fueron validadas positivamente por el Dr. Palet. La complejidad estructural de varias de estas reglas, así

como la diversidad de atributos involucrados, coincide con la heterogeneidad del personal médico en este hospital y con los patrones de experiencia, formación y percepción detectados en las encuestas. La única regla descartada en la evaluación corresponde a un patrón sobre extendido cuya combinación de atributos no refleja adecuadamente la dinámica clínica observada por el especialista.

En conjunto, la evaluación subjetiva del Dr. Palet indica que diecisiete de las dieciocho reglas generadas mediante la métrica *Correlation* poseen coherencia clínica, institucional y operativa, mientras que únicamente una regla fue descartada por su falta de correspondencia con la práctica médica real. Este resultado confirma que la métrica *Correlation* no solo es sólida desde un punto de vista estadístico, sino también desde la interpretación experta.

4.5 Vista Minable: Resultados aplicando la función de calidad *Klosgen*

La métrica de *Klosgen* combina el soporte de una regla con la ganancia de precisión respecto al comportamiento promedio, ponderando la mejora en la probabilidad del consecuente por la frecuencia con la que se observa el patrón en los datos. En términos prácticos, privilegia reglas que, además de ser frecuentes dentro del grupo objetivo, incrementan de manera apreciable la precisión de clasificación frente al resto de la población, lo que la hace especialmente útil para identificar conjuntos de contraste con buen equilibrio entre cobertura y valor informativo. En este trabajo se aplicó Klosgen sobre el Conjunto C (Vista Minable) en el nivel de análisis específico, utilizando como atributo objetivo el centro hospitalario e incorporando los mismos parámetros de configuración empleados para las demás funciones de calidad en esta etapa del experimento.

4.5.1 Reglas sin interpretar

Las reglas sin interpretar obtenidas mediante *Klosgen* muestran un comportamiento numérico consistente con el propósito de la métrica de balancear soporte y ganancia de precisión como se observa en la Tabla 4.7. En el bloque correspondiente al Hospital

Regional Río Blanco se observan patrones con soportes comprendidos aproximadamente entre 0.48 y 0.92, con diferencias de soporte siempre positivas y superiores a 0.48. En particular, las reglas iniciales combinan condiciones sobre desacuerdo parcial con las implicaciones legales de la autopsia, formación médica en instituciones específicas y ausencia de determinados motivos administrativos para solicitarla o rechazarla. La precisión de estas reglas oscila entre valores muy altos, iguales a 1.0, y valores intermedios cercanos a 0.60, mientras que la diferencia de precisión se mantiene positiva salvo en la regla que presenta el soporte más elevado, donde la ganancia se vuelve negativa debido al incremento de contraejemplos.

El conjunto de reglas asociado al Hospital General de San Juan Bautista Tuxtepec se caracteriza por soportes muy altos, cercanos o superiores a 0.89 en la mayoría de los casos, con diferencias de soporte que sobrepasan el 0.73 y precisiones que se sitúan sistemáticamente por encima de 0.90. Esto se traduce en diferencias de precisión muy elevadas, cercanas a 1.0, y en valores de redundancia que superan los diez mil, lo que evidencia una fuerte estabilidad interna en los perfiles de respuesta de este hospital.

Por su parte, las reglas vinculadas al Hospital General de Zona 53 presentan soportes que van desde aproximadamente 0.68 hasta casi 1.0, con diferencias de soporte igualmente altas y positivas. La precisión mantiene valores elevados en algunas reglas, superiores a 0.34, pero muestra descensos relevantes en aquellas donde el soporte negativo aumenta, llegando incluso a valores por debajo de 0.40. En estas últimas reglas la diferencia de precisión se torna negativa, lo que indica que, pese a su elevada cobertura, no mejoran la precisión global del modelo. La redundancia, nuevamente, se sitúa en rangos superiores a 40 000, lo que sugiere la existencia de múltiples patrones equivalentes en el hospital.

Tabla 4.7. *Klosgen* – Vista Minable: Reglas sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	IF hall_arb = !{8a} AND esc_med_gral = !{c40} AND mcc_no_aut = !{16p} AND esc_esp = !{c111} THEN centro_hospitalario = {H1}	0.747114	0.038441	0.708672	0.851895	0.703790	5.0
2	IF hall_arb = !{8a} AND com_sug_op = !{22e} AND med_aut = !{20i} THEN centro_hospitalario = {H1}	0.730074	0.108766	0.621309	0.665168	0.330336	5.0
3	IF hall_dem = !{11a} AND esc_med_gral = !{c40} AND anios_prac = !{3c} THEN centro_hospitalario = {H1}	0.732402	0.145664	0.586738	0.598084	0.196168	5.0
4	IF categoria = !{c1} THEN centro_hospitalario = {H1}	0.482402	0.000000	0.482402	1.000000	1.000000	5.0
5	IF esc_med_gral = {c1} AND per_sol_aut = !{19c} AND fmr_sol_aut = !{21h} THEN centro_hospitalario = {H1}	0.597020	0.076745	0.520275	0.697184	0.394368	5.0
6	IF esc_esp = !{c103} AND rechazo_fam = !{17e} AND hall_dem = !{11e} AND mcc_no_aut = !{16p} THEN centro_hospitalario = {H1}	0.919367	0.404861	0.514506	0.401938	-0.196125	5.0
7	IF hall_arb = {8a} AND hall_disc = !{7b} AND categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c85} AND mcc_no_aut = !{16i} AND com_sug_op = !{22h} THEN centro_hospitalario = {H11}	0.896097	0.026305	0.869793	0.984547	0.969094	10457.0
8	IF hall_dem = {11a} AND ult_grado = !{g8} AND anios_prac = !{3e} AND mcc_aut = !{14d} AND fmr_sol_aut = !{21e} AND casos = !{4e} AND no_hosp = !{18z} AND med_aut = !{20f} AND per_sol_aut = !{19e} THEN centro_hospitalario = {H11}	0.808380	0.048093	0.760287	0.969171	0.938341	10939.0
9	IF hall_arb = {8a} AND esc_med_gral = !{c125} AND categoria = {c1} AND esc_esp = !{c6} THEN centro_hospitalario = {H11}	0.904027	0.072933	0.831094	0.958648	0.917295	10457.0

Tabla 4.7. *Klosgen* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
10	IF hall_disc = {7a} AND categoria = {c1} AND esc_esp = !{c6} AND rechazo_fam = !{17z} THEN centro_hospitalario = {H11}	0.906083	0.083674	0.822409	0.952946	0.905893	10457.0
11	IF hall_disc = !{7b} AND categoria = {c1} AND hall_arb = !{8b} AND hall_dem = !{11d} THEN centro_hospitalario = {H11}	0.922203	0.098627	0.823577	0.945910	0.891820	10457.0
12	IF hall_disc = !{7b} AND esc_esp = !{c6} AND mcc_aut = !{14i} AND ult_grado = !{98} AND anios_prac = !{3e} THEN centro_hospitalario = {H11}	0.889505	0.145011	0.744495	0.919822	0.839644	10457.0
13	IF esc_esp = !{c6} AND hall_arb = !{8c} AND com_sug_op = !{22a} AND casos = !{4e} AND ult_grado = !{g8} THEN centro_hospitalario = {H11}	0.926837	0.193836	0.733001	0.899424	0.798847	10939.0
14	IF hall_arb = !{8b} AND categoria = {c1} AND ult_grado = !{g8} AND esc_med_gral = !{c85} AND hall_dem = !{11c} THEN centro_hospitalario = {H11}	0.875506	0.102655	0.772851	0.941005	0.882010	10457.0
15	IF hall_disc = !{7a} AND esc_med_gral = !{c1} AND esc_esp = !{c6} AND anios_prac = !{3e} AND rechazo_fam = !{17z} AND categoria = {c1} AND mcc_aut = !{14d} AND mcc_no_aut = !{16d} AND fmr_sol_aut = !{21h} AND casos = !{4e} THEN centro_hospitalario = {H12}	0.888220	0.004688	0.883532	0.962750	0.925499	41101.0
16	IF esc_med_gral = !{1} AND com_sug_op = !{22a} AND area = !{23} AND per_sol_aut = !{19b} AND casos = {4a} THEN centro_hospitalario = {H12}	0.682019	0.082737	0.599282	0.529282	0.058565	42424.0

Tabla 4.7. *Klosgen* – Vista Minable: Reglas sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
17	IF anios_prac = {3a} AND esc_esp = !{c6} AND hall_dem = !{11d} AND hall_arb = !{8d} AND med_aut = !{20z} AND mcc_no_aut = !{16e} AND no_hosp = !{18z} THEN centro_hospitalario = {H12}	0.750044	0.172192	0.577852	0.372711	-0.254577	42424.0
18	IF hall_dem = !{11a} AND esc_med_gral = !{c1} AND categoria = {c1} AND esc_esp = !{c6} THEN centro_hospitalario = {H12}	0.687334	0.107747	0.579587	0.465284	-0.069433	41101.0
19	IF hall_disc = !{7a} AND esc_med_gral = !{c1} THEN centro_hospitalario = {H12}	0.888574	0.083631	0.804943	0.591719	0.183438	41101.0
20	IF esc_esp = !{c103} AND esc_med_gral = !{c1} AND com_sug_op = !{22a} AND fmr_sol_aut = !{21h} AND per_sol_aut = !{19b} THEN centro_hospitalario = {H12}	0.995394	0.260028	0.735366	0.343040	-0.313919	41101.0

4.5.2 Reglas interpretadas

La Tabla 4.8 de reglas interpretadas de *Klosgen* ofrece una visión textual clara de los perfiles que caracterizan a cada hospital. En el Hospital Regional Río Blanco las reglas describen médicos que, de manera reiterada, muestran desacuerdo parcial con que los hallazgos de las autopsias originen demandas o arbitrajes, no cursaron medicina general en la UNAM, no realizaron su especialidad en determinadas instituciones federales y no consideran ciertos motivos administrativos como desencadenantes de la solicitud de autopsias. También aparecen perfiles que integran la formación en la Universidad Veracruzana, la percepción sobre el personal adecuado para ordenar autopsias y la preferencia por ciertos canales institucionales para emitir la solicitud.

En el Hospital General de San Juan Bautista Tuxtepec las reglas interpretadas enfatizan tanto las percepciones sobre la relación entre autopsias, arbitrajes y diagnósticos discrepantes como el tipo de participación del médico en estudios previos, el rechazo o aceptación de motivos para solicitar el procedimiento y la identificación de la figura responsable de autorizarlo. Se combinan elementos de experiencia profesional, número de autopsias en las que se ha participado y postura frente a la aplicabilidad de los estudios, construyendo perfiles que reflejan una cultura institucional muy definida respecto al valor médico–legal de la autopsia.

En el Hospital General de Zona 53 las reglas generadas tienden a ser más extensas, incorporando simultáneamente formación académica, años de práctica, participación previa en estudios de autopsia, motivos para solicitar o no solicitar el procedimiento, percepción del desinterés como causa de omisión, importancia de la psicología en el proceso y rol de diferentes actores en la autorización. Estos patrones describen un hospital con mayor diversidad interna, donde los perfiles combinan múltiples dimensiones de experiencia y percepción, pero logran agruparse de forma consistente en torno a configuraciones específicas de respuestas.

Tabla 4.8. *Klosgen* – Vista Minable: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en UNAM, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza y no hicieron su especialidad en la Secretaría de Salubridad y Asistencia; entonces son del Hospital Regional Río Blanco.
2	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no opinan que son importantes las autopsias y no estiman como motivo para solicitar autopsia el acoso por delincuencia organizada; entonces son del Hospital Regional Río Blanco.
3	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estudiaron medicina general en UNAM y no tienen entre 11 y 15 años de práctica; entonces son del Hospital Regional Río Blanco.
4	Si los médicos no respondieron espontáneamente la encuesta; entonces son del Hospital Regional Río Blanco.
5	Si los médicos estudiaron medicina general en Universidad Veracruzana, alegan que no es el trabajador social el personal adecuado para ordenar autopsia y están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología; entonces son del Hospital Regional Río Blanco.
6	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, creen que la causa de rechazo por parte del familiar no es por inconformidad de los familiares con la atención prestada al paciente, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda y piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza; entonces son del Hospital Regional Río Blanco.
7	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no estudiaron medicina general en el Hospital General de Orizaba, piensan que una causa o motivo para no solicitar autopsia no es por cuestiones religiosas y no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.8. *Klosgen* – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
8	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no son estudiantes de medicina, no tienen más de 20 años de práctica, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, están de acuerdo en que el método eficiente para solicitar autopsia no es realizarlo a través de otra área, no han participado en más de 20 casos de autopsias, responden por qué en el hospital no se realizan suficientes autopsias, no estiman como motivo para solicitar autopsia la orientación diagnóstica y alegan que no es el familiar quien debe ordenan autopsia; entonces son del Hospital General de San Juan Bautista Tuxtepec.
9	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud, respondieron espontáneamente la encuesta y no estudiaron su especialidad en Hospital Regional de Río Blanco; entonces son del Hospital General de San Juan Bautista Tuxtepec.
10	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco y responden a cuál es la causa de rechazo por parte del familiar; entonces son del Hospital General de San Juan Bautista Tuxtepec.
11	Si los médicos no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, respondieron espontáneamente la encuesta, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de San Juan Bautista Tuxtepec.
12	Si los médicos no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron su especialidad en Hospital Regional de Río Blanco, no plantean como motivo para solicitar autopsia esclarecer diagnóstico difícil, no son estudiantes de medicina y no tienen más de 20 años de práctica; entonces son del Hospital General de San Juan Bautista Tuxtepec.
13	Si los médicos no estudiaron su especialidad en Hospital Regional de Río Blanco, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, comentan, no han participado en más de 20 casos de autopsias y no son estudiantes de medicina; entonces son del Hospital General de San Juan Bautista Tuxtepec.
14	Si los médicos no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, respondieron espontáneamente la encuesta, no son estudiantes de medicina, no estudiaron medicina general en el Hospital General de Orizaba y no consideran incierto que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.8. *Klosgen* – Vista Minable: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
15	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Hospital Regional de Río Blanco, no tienen más de 20 años de práctica, responden a cuál es la causa de rechazo por parte del familiar, respondieron espontáneamente la encuesta, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, piensan que una causa o motivo para no solicitar autopsia no es el desconocimiento de la práctica de autopsias, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología y no han participado en más de 20 casos de autopsias; entonces son del Hospital General de Zona 53.
16	Si los médicos no estudiaron medicina general en Universidad Veracruzana, comentan, no son adscritos, alegan que no es la enfermera el personal adecuado para ordenar autopsia y no han participado en ningún estudio de autopsia; entonces son del Hospital General de Zona 53.
17	Si los médicos tienen menos de 5 años de práctica, no estudiaron su especialidad en Hospital Regional de Río Blanco, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, especifican motivos para solicitar autopsia, piensan que una causa o motivo para no solicitar autopsia no es el desinterés y responden por qué en el hospital no se realizan suficientes autopsias; entonces son del Hospital General de Zona 53.
18	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estudiaron medicina general en Universidad Veracruzana, respondieron espontáneamente la encuesta y no estudiaron su especialidad en Hospital Regional de Río Blanco; entonces son del Hospital General de Zona 53.
19	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos y no estudiaron medicina general en Universidad Veracruzana; entonces son del Hospital General de Zona 53.
20	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Veracruzana, comentan, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología y alegan que no es la enfermera el personal adecuado para ordenar autopsia; entonces son del Hospital General de Zona 53.

4.5.3 Análisis Visual

El análisis visual de las métricas asociadas a *Klosgen* permite profundizar en la estructura cuantitativa de las reglas descubiertas. En la gráfica representada en la Figura

4.34, correspondiente al soporte, se aprecia una concentración clara de valores altos: la mayoría de las reglas se sitúa por encima de 0.70 y varias superan el 0.90, alcanzando incluso un valor cercano a 1.0 en la última regla. Esta distribución indica que la métrica tiende a seleccionar patrones con una cobertura muy amplia dentro del hospital objetivo, lo cual es coherente con su definición, que pondera explícitamente el soporte. La mediana se ubica en un rango elevado, lo que sugiere que la mayor parte de los médicos en cada hospital se agrupa dentro de perfiles bien definidos por las reglas.

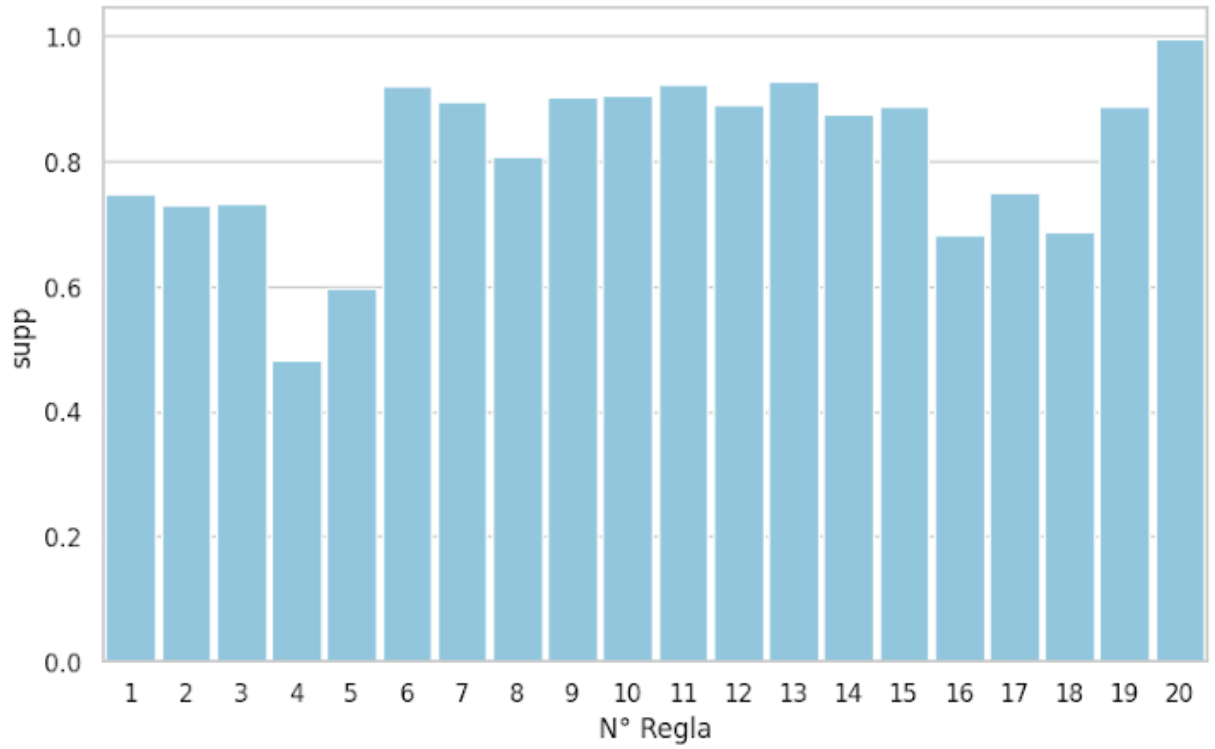


Figura 4.34 *Klosgen* – Vista Minable: Soporte por Regla

En la gráfica correspondiente al soporte negativo, mostrada en la Figura 4.35, la mayoría de las reglas presenta valores bajos, cercanos a cero, reflejando una delimitación adecuada de los subgrupos y un número reducido de contraejemplos. Sin embargo, se observan incrementos notables en algunas reglas, particularmente en las asociadas al Hospital Regional Río Blanco y al Hospital General de Zona 53, donde el soporte negativo supera el 0.20 e incluso alcanza valores cercanos a 0.40. Estas barras emergentes se correlacionan con las reglas que presentan precisiones más bajas y diferencias de

precisión negativas, revelando patrones amplios, pero con menor capacidad discriminativa.

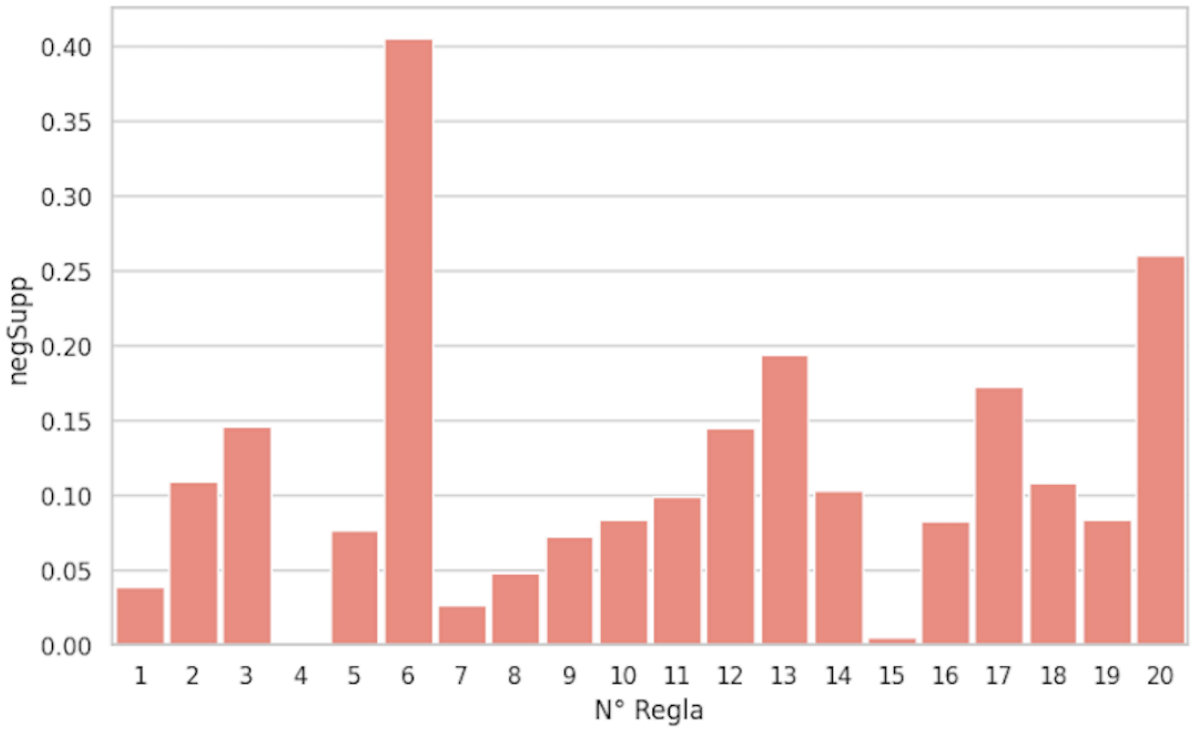


Figura 4.35 *Klosgen* – Vista Minable: Soporte Negativo por Regla

La diferencia de soporte, visible en la gráfica de la Figura 4.36, mantiene valores consistentemente positivos en todas las reglas, con un rango que oscila entre aproximadamente 0.48 y más de 0.88. Este comportamiento reafirma que todos los patrones representan verdaderos conjuntos de contraste, en los que la presencia del hospital objetivo es significativamente mayor que en el resto de la población. Los valores más altos se concentran en ciertas reglas de Tuxtepec y del Hospital General de Zona 53, lo cual confirma una separación clara entre los perfiles internos de estos hospitales y los de los demás centros.

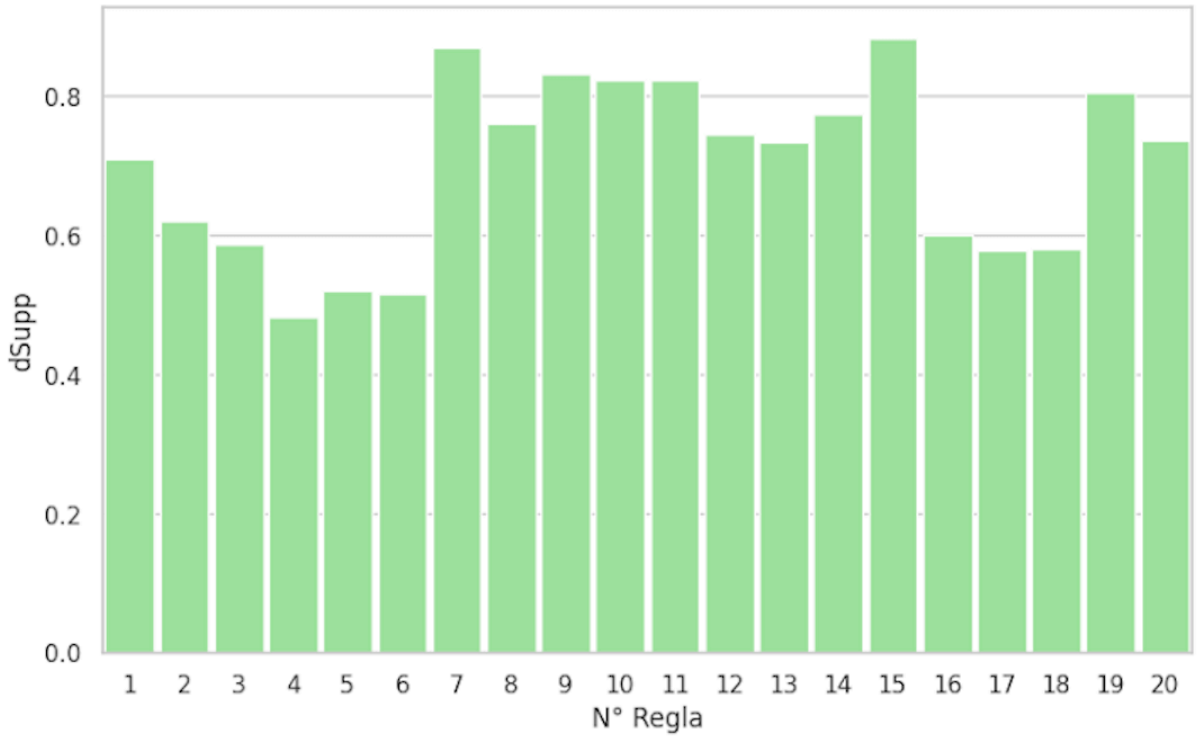


Figura 4.36 *Klosgen* – Vista Minable: Diferencia de Soporte por Regla

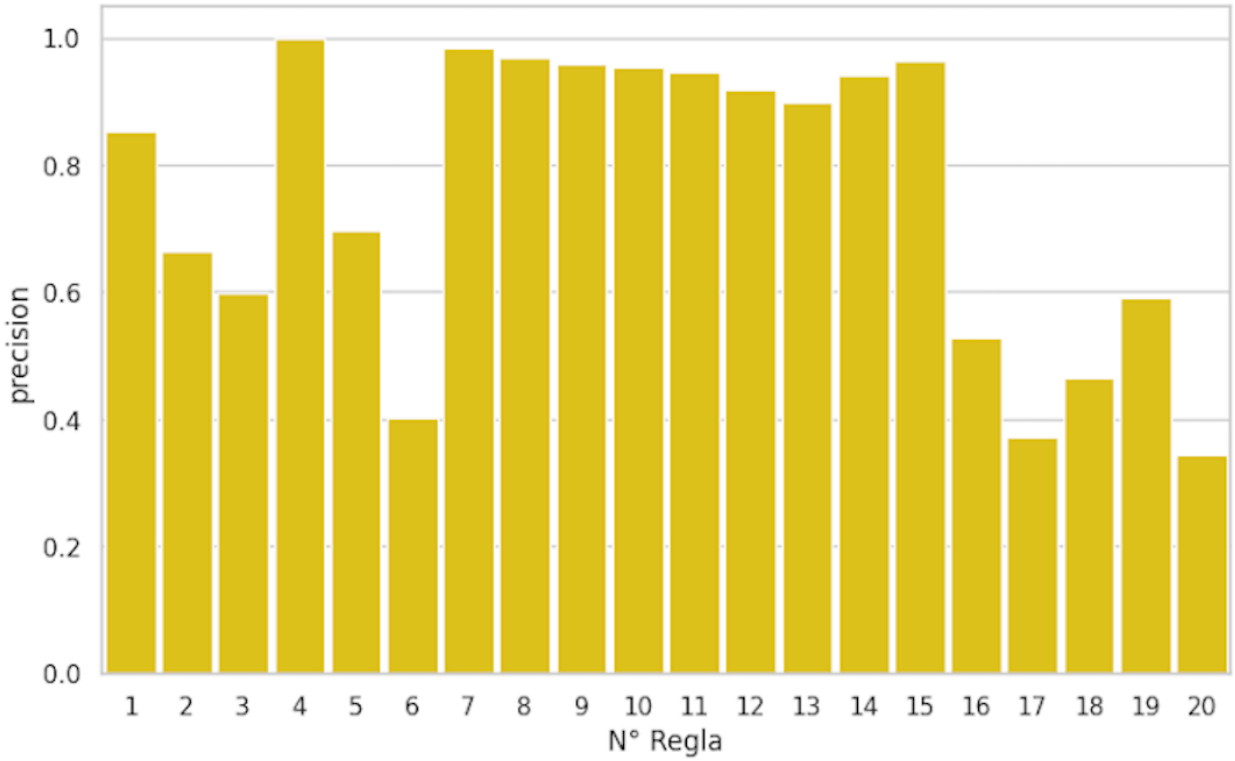


Figura 4.37 *Klosgen* – Vista Minable: Precisión por Regla

La gráfica de precisión, representada en la Figura 4.37, muestra un comportamiento heterogéneo, pero con predominio de valores elevados. Un subconjunto importante de reglas alcanza precisiones superiores a 0.94, especialmente en el hospital H11, mientras que otras se sitúan en rangos intermedios entre 0.55 y 0.70, y unas pocas descienden a valores cercanos a 0.34 o 0.37. Esta variación se explica directamente por los niveles de soporte negativo observados en la Figura 4.35. El modelo mantiene, no obstante, una proporción considerable de reglas con precisión alta, lo que indica que *Klosgen* logra capturar patrones informativos sin sacrificar en exceso la exactitud.

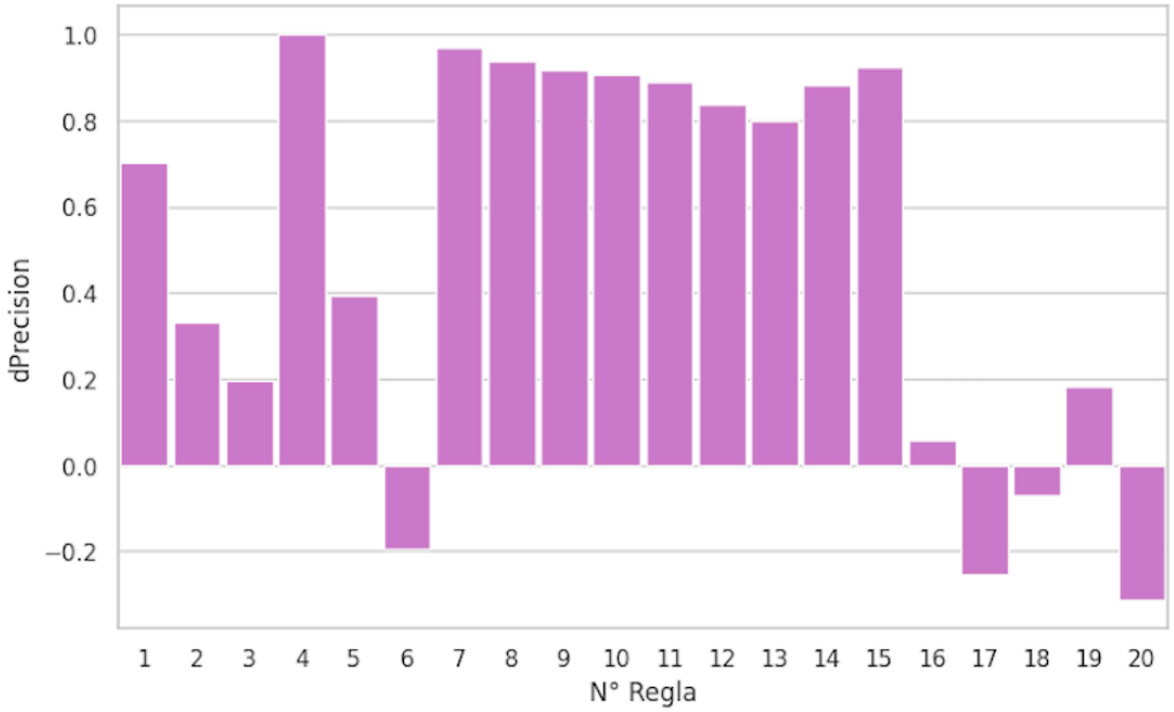


Figura 4.38 *Klosgen* – Vista Minable: Diferencia de Precisión por Regla

La diferencia de precisión, mostrada en la Figura 4.38, refuerza esta interpretación. La mayoría de las reglas exhibe mejoras significativas respecto al comportamiento promedio, con valores de ganancia que superan con frecuencia el 0.70 e incluso alcanzan valores cercanos a 1.0 en varias reglas asociadas a H11 y H12. En el extremo opuesto, algunas reglas presentan diferencias de precisión negativas, especialmente aquellas con soporte negativo elevado, lo que confirma que, aunque cubren una gran parte del hospital, no incrementan la precisión de la clasificación. Estas reglas

constituyen puntos de advertencia dentro del modelo y serán retomadas en la evaluación subjetiva.

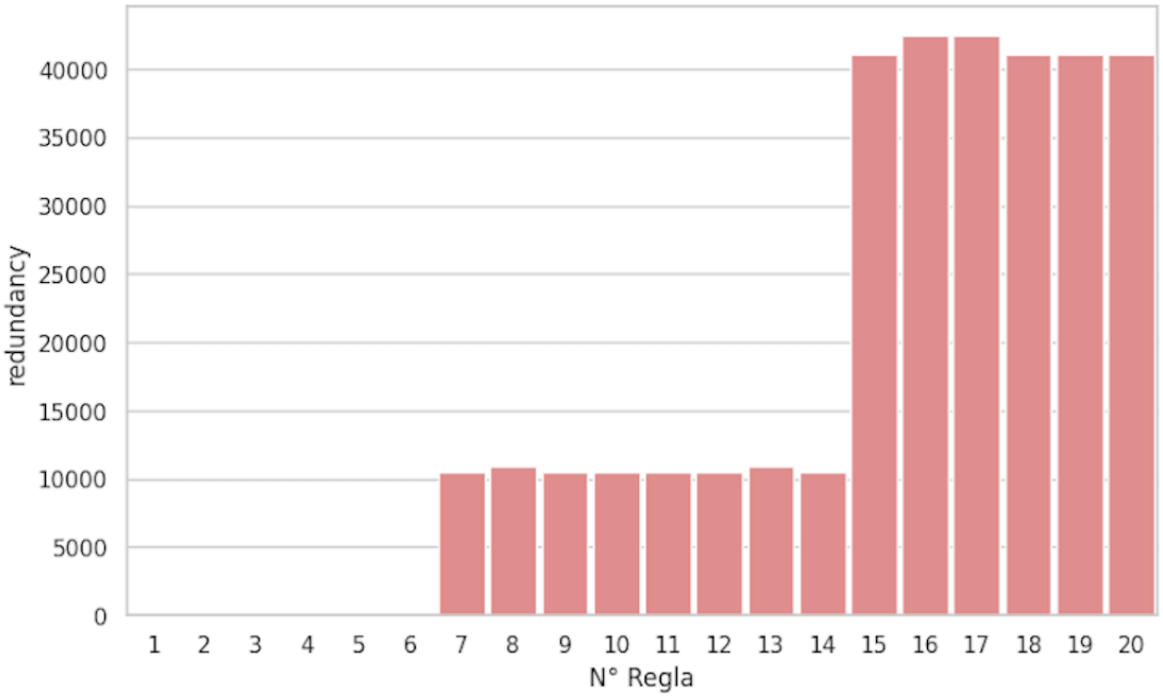


Figura 4.39 *Klosgen* – Vista Minable: Redundancia por Regla

Finalmente, la gráfica de redundancia representada en la Figura 4.39 muestra una separación muy marcada entre dos grupos. Las reglas del Hospital Regional Río Blanco se concentran en los valores mínimos de redundancia, alrededor de cinco unidades, lo que indica patrones relativamente únicos y con poco solapamiento entre sí. En contraste, las reglas asociadas a Tuxtepec y al Hospital General de Zona 53 exhiben valores que superan los diez mil y se acercan a los cuarenta mil, lo que revela la existencia de regiones densas en el espacio de búsqueda donde múltiples reglas comparten estructuras muy similares. Esta alta redundancia no invalida los patrones, pero sí sugiere la necesidad de considerar criterios adicionales de selección cuando se busque un subconjunto reducido de reglas representativas.

4.5.4 Evaluación Objetiva

Desde la perspectiva de la evaluación objetiva, el modelo construido con *Klosgen* muestra una base cuantitativa robusta. La diferencia de soporte es estrictamente positiva en todas las reglas, lo que confirma que cada patrón incrementa la frecuencia del hospital objetivo en comparación con el resto de la población, cumpliendo así con la noción de conjunto de contraste. Los valores de precisión son, en su mayoría, altos, particularmente en las reglas asociadas a Tuxtepec y a algunas reglas de Zona 53, donde se alcanzan probabilidades condicionales cercanas a uno.

Las reglas con diferencias de precisión negativas constituyen excepciones relevantes y se relacionan de forma directa con valores elevados de soporte negativo. Estas reglas, si bien cumplen con el criterio de contraste en términos de frecuencia, no mejoran la precisión general y, por tanto, deben ser interpretadas con cautela. El análisis conjunto de *supp, negSupp, dSupp* y *dPrecision* revela que Klosgen privilegia patrones amplios y muy contrastivos, pero permite la entrada de algunos patrones de menor calidad clasificatoria cuando su soporte es extremadamente alto. La redundancia, por su parte, confirma la existencia de núcleos muy densos de reglas en Tuxtepec y en el Hospital General de Zona 53, lo que refuerza la idea de que los perfiles internos de estos hospitales son altamente consistentes y se pueden describir mediante múltiples combinaciones de condiciones equivalentes.

4.5.5 Evaluación Subjetiva

La evaluación subjetiva realizada por el Dr. Palet sobre las reglas generadas por *Klosgen* permite distinguir con claridad los patrones que, además de ser estadísticamente sólidos, resultan clínicamente plausibles y coherentes con la dinámica institucional de los hospitales analizados. En el Hospital Regional Río Blanco el especialista consideró válidas las reglas que vinculan de manera razonable las percepciones sobre la demanda o el arbitraje, la formación en la Universidad Veracruzana, la forma de solicitar autopsias y ciertos elementos de la trayectoria profesional, mientras que descartó reglas donde la combinación de condiciones no se corresponde con perfiles observados en su

experiencia, particularmente aquellas que integran motivos de rechazo familiar o configuraciones formativas que no reflejan la realidad del hospital.

En el Hospital General de San Juan Bautista Tuxtepec la evaluación subjetiva confirmó la validez del 50% de las reglas, destacando que las combinaciones que describen percepciones sobre discrepancias diagnósticas, motivos para solicitar autopsias, número de procedimientos realizados y reconocimiento de la figura encargada de autorizarlas coinciden con patrones reales en la práctica clínica del hospital. El 50% de las reglas fueron rechazadas por considerarse demasiado artificiosas, al combinar estados de opinión que difícilmente se presentan simultáneamente en el personal médico de la institución.

Para el Hospital General de Zona 53 el Dr. Palet validó de forma positiva todas las reglas de la parte final del modelo, en las que se describen médicos con determinadas trayectorias de formación, años de experiencia, participación en autopsias y percepciones sobre los motivos de solicitud o rechazo del procedimiento. Estas reglas se alinean con la heterogeneidad propia del hospital, donde coexisten médicos con distintos niveles de experiencia y formación, pero que comparten ciertos patrones de respuesta frente a la práctica de autopsias.

En conjunto, la validación subjetiva sobre *Klosgen* indica que una fracción importante de las reglas generadas es clínicamente interpretable y consistente con la realidad de los hospitales estudiados, mientras que un subconjunto reducido debe considerarse con cautela o bien descartarse al momento de construir explicaciones o recomendaciones basadas en los resultados del modelo. Esta fase refuerza la complementariedad entre el análisis cuantitativo y el juicio experto, y confirma que la métrica de Klosgen, aplicada sobre la Vista Minable, produce patrones en su mayoría alineados con la práctica médica real.

4.6 Vista Minable: Resultados aplicando la función de calidad *RelativeRisk*

La métrica *RelativeRisk* evalúa la proporción en la que la presencia del antecedente incrementa o disminuye la probabilidad del consecuente en comparación con la población restante. A diferencia de otras funciones de calidad empleadas anteriormente, *RelativeRisk* hace explícita la relación proporcional entre los grupos, permitiendo identificar patrones en los que el antecedente amplifica la presencia del hospital objetivo o, por el contrario, la reduce de forma significativa. Su aplicación en el nivel específico de la Vista Minable genera modelos compactos con pocas reglas, lo que facilita su interpretación, pero al mismo tiempo revela las limitaciones inherentes del método cuando existen perfiles institucionales heterogéneos.

4.6.1 Reglas sin interpretar

Las reglas sin interpretar derivadas de *RelativeRisk* conforman un conjunto compacto de tres patrones y se muestra en la Tabla 4.9. En el caso del Hospital Regional Río Blanco se identificó una única regla que cubre la totalidad del hospital, lo que se evidencia en un soporte igual a 1.0. Sin embargo, el soporte negativo de aproximadamente 0.51 indica la presencia considerable de contraejemplos en otros hospitales que cumplen el mismo antecedente. Esto provoca que la precisión caiga a un valor cercano a 0.36 y que la diferencia de precisión sea negativa, lo cual refleja que, aunque la regla aparece en todos los médicos del hospital, también aparece con frecuencia relevante en los demás centros; por ello su capacidad discriminativa es limitada.

En el Hospital General de San Juan Bautista Tuxtepec la regla obtenida presenta también un soporte igual a 1.0, pero con un soporte negativo menor, de alrededor de 0.32, lo que reduce la presencia de contraejemplos y genera una diferencia de soporte notablemente alta de 0.678. Este comportamiento se traduce en una precisión elevada de 0.85 y una ganancia positiva de precisión que supera el 0.70. La redundancia, superior a diez mil unidades, indica que este patrón se encuentra dentro de una región densa del espacio de búsqueda donde múltiples reglas comparten estructuras similares.

La tercera regla, correspondiente al Hospital General de Zona 53, también cubre todo el conjunto del hospital, con un soporte de 1.0 y un soporte negativo de 0.44. La diferencia de soporte se mantiene en valores positivos cercanos a 0.55, pero la precisión desciende hasta aproximadamente 0.23 y la diferencia de precisión se vuelve marcadamente negativa. Esta caída está directamente relacionada con la elevada presencia de contraejemplos en otros hospitales, lo que convierte a este patrón en uno de los menos discriminativos dentro del modelo. La redundancia, cercana a cuarenta mil unidades, confirma nuevamente la existencia de múltiples reglas similares en la estructura interna del hospital.

Tabla 4.9. *RelativeRisk* – Vista Minable: Reglas sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	IF esc_esp = !{c103} AND esc_med_gral = !{c104} THEN centro_hospitalario = {H1}	1.0	0.517843	0.482157	0.363673	-0.272653	5.0
2	IF categoria = {c1} AND esc_esp = !{c6} AND esc_med_gral = !{c125} AND mcc_aut = !{14e} AND mcc_no_aut = !{16m} AND com_sug_op = !{22j} AND no_hosp = !{18h} AND med_aut = !{20h} THEN centro_hospitalario = {H11}	1.0	0.321575	0.678425	0.853284	0.706569	10457.0
3	IF esc_med_gral = !{c1} AND mcc_no_aut = !{16p} AND fmr_sol_aut = !{21h} AND esc_esp = !{c6} AND categoria = {c1} AND com_sug_op = !{22p} AND rechazo_fam = !{17z} AND mcc_aut = !{14d} AND no_hosp = !{18h} AND med_aut = !{20f} THEN centro_hospitalario = {H12}	1.0	0.444955	0.555045	0.234632	-0.530737	41101.0

4.6.2 Reglas Interpretadas

La regla interpretada correspondiente al Hospital Regional Río Blanco describe un perfil marcado por la ausencia de formación tanto en la Secretaría de Salud de Oaxaca como en la Secretaría de Salud del Estado de Yucatán. Este patrón se presenta en todos los médicos del hospital, lo que explica su soporte máximo, aunque la baja precisión indica que también aparece en profesionales de otros hospitales.

Tabla 4.10. *RelativeRisk* – Vista Minable: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos no hicieron su especialidad en la Secretaría de Salud de Oaxaca y no estudiaron medicina general en la Secretaría de Salud del Estado de Yucatán; entonces son del Hospital Regional Río Blanco.
2	Si los médicos respondieron espontáneamente la encuesta, no estudiaron su especialidad en Hospital Regional de Río Blanco, no estudiaron medicina general en el Centro Interdisciplinario de Ciencias de la Salud, no plantean como motivo para solicitar autopsia la falta de correlación entre los datos clínicos y los de laboratorio, piensan que una causa o motivo para no solicitar autopsia no es la condición del finado, no opinan que para este estudio deberían participar todos los servicios, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares y no estiman como motivo para solicitar autopsia un problema médico legal; entonces son del Hospital General de San Juan Bautista Tuxtepec.
3	Si los médicos no estudiaron medicina general en Universidad Veracruzana, piensan que una causa o motivo para no solicitar autopsia no es que se haga sin fines de enseñanza, están de acuerdo en que el método eficiente para solicitar autopsia no es a través del área de psicología, no estudiaron su especialidad en Hospital Regional de Río Blanco, respondieron espontáneamente la encuesta, no opinan que se debe solicitar el consentimiento, responden a cuál es la causa de rechazo por parte del familiar, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares y no estiman como motivo para solicitar autopsia la orientación diagnóstica; entonces son del Hospital General de Zona 53.

En el caso del Hospital General de San Juan Bautista Tuxtepec, la regla interpretada caracteriza a médicos que respondieron espontáneamente la encuesta, no realizaron su formación en instituciones específicas, no consideran que la falta de correlación clínico–laboratorio sea un motivo para solicitar autopsia, no ven la condición del finado como una barrera, no consideran necesaria la participación de todos los servicios y tampoco perciben la negación familiar o los problemas médico-legales como factores determinantes para solicitar el procedimiento. Este conjunto de atributos conforma un perfil semánticamente complejo y clínicamente plausible para la institución, y la métrica lo respalda con alta precisión.

La regla interpretada para el Hospital General de Zona 53 describe un patrón más amplio en términos de percepciones y formación. En este caso se combinan atributos

relacionados con desacuerdos sobre la enseñanza en autopsias, posicionamientos sobre la utilidad de áreas como psicología para la solicitud del estudio, ausencia de formación en la Universidad Veracruzana y respuesta espontánea al instrumento. Su amplitud coincide con el comportamiento heterogéneo del hospital, aunque la métrica confirma que este patrón no es particularmente preciso.

4.6.3 Análisis Visual

El análisis visual del modelo basado en *RelativeRisk* ofrece una lectura clara del comportamiento cuantitativo de los tres patrones generados. En la gráfica representada en la Figura 4.40, correspondiente al soporte, se confirma que todas las reglas poseen un valor igual a 1.0, lo que significa que cada antecedente está presente en la totalidad del grupo hospitalario que predice. Esta uniformidad es característica de métricas basadas en proporciones relativas, ya que la maximización del riesgo relativo tiende a seleccionar patrones completos.

En la gráfica mostrada en la Figura 4.41, que representa el soporte negativo, se observa una variación sustancial entre las reglas. La primera presenta el valor más alto, alrededor de 0.52, mientras que las restantes muestran valores cercanos a 0.32 y 0.44, respectivamente. Esta discrepancia es fundamental porque determina la precisión de las reglas; a mayor soporte negativo, mayor presencia de contraejemplos, lo que reduce la capacidad clasificatoria del patrón.

El comportamiento de la diferencia de soporte, visible en la Figura 4.42, muestra valores positivos que oscilan entre 0.48 y aproximadamente 0.67. La segunda regla, correspondiente al Hospital General de San Juan Bautista Tuxtepec, presenta el valor más alto, confirmando su mayor capacidad de diferenciación entre su hospital objetivo y el resto de la población. Las otras reglas mantienen valores positivos, pero más moderados.

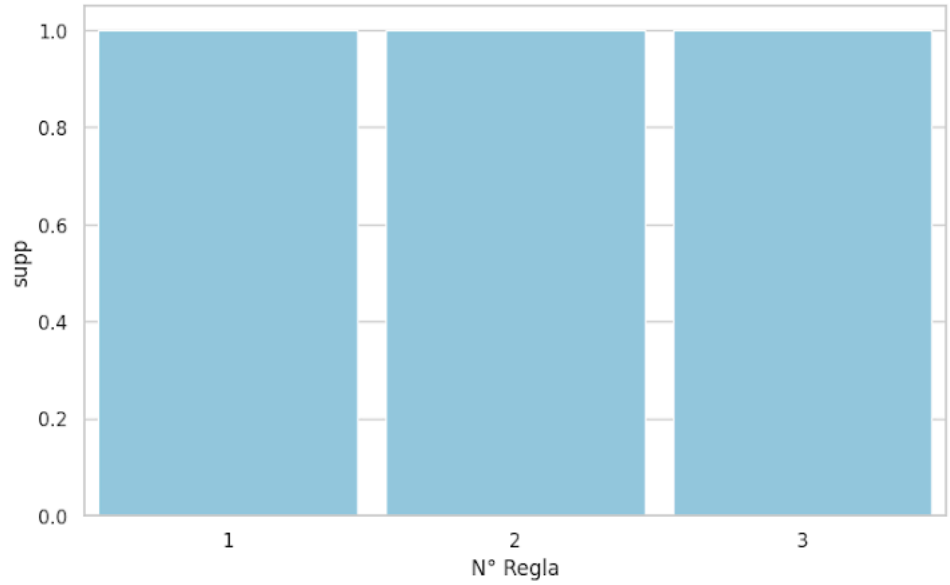


Figura 4.40 *RelativeRisk* – Vista Minable: Soporte por Regla

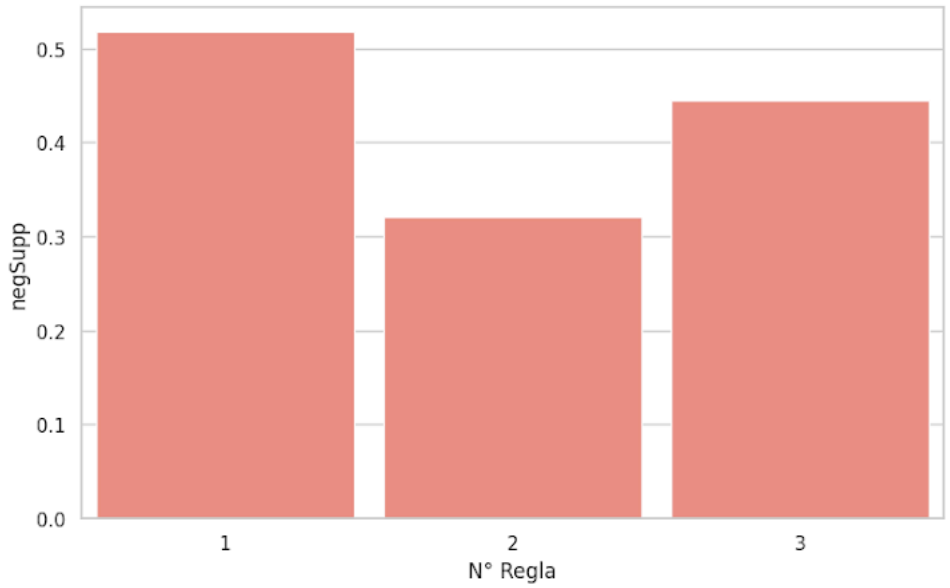


Figura 4.41 *RelativeRisk* – Vista Minable: Soporte Negativo por Regla

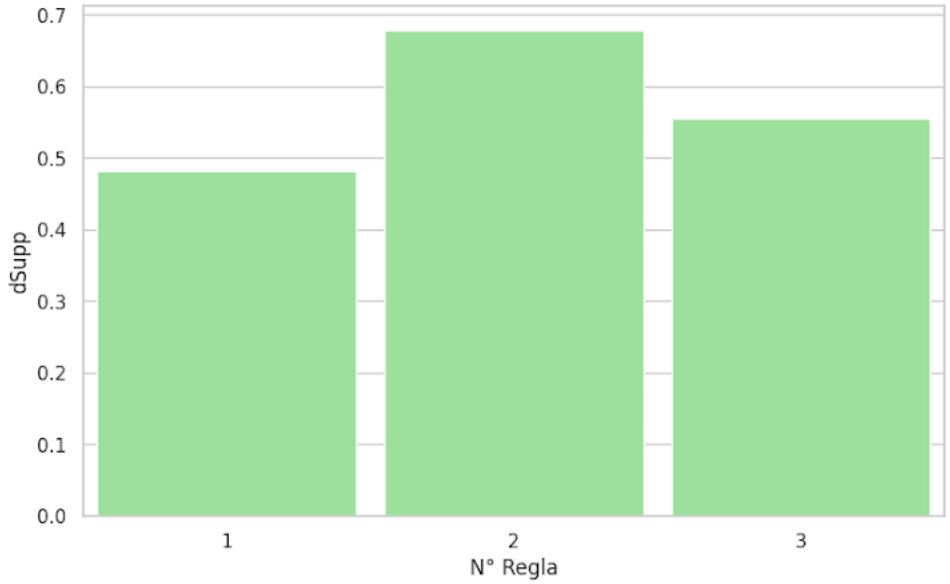


Figura 4.42 *RelativeRisk* – Vista Minable: Diferencia de Soporte por Regla

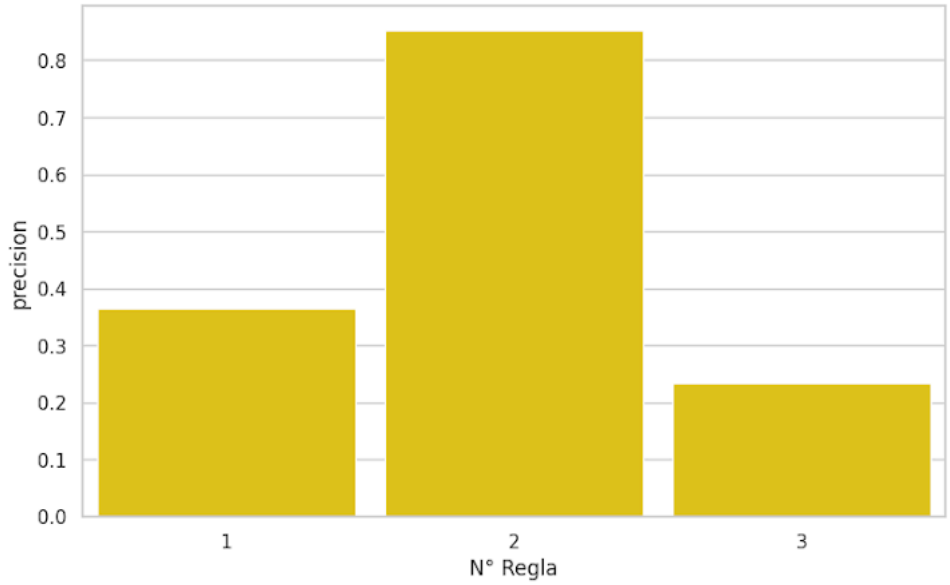


Figura 4.43 *RelativeRisk* – Vista Minable: Precisión por Regla

La gráfica de precisión representada en la Figura 4.43 ilustra con claridad el impacto directo del soporte negativo. La segunda regla muestra una precisión notablemente mayor, por encima de 0.85, mientras que la primera y la tercera descienden a valores cercanos a 0.36 y 0.23, respectivamente. Este comportamiento se explica de manera

directa por la proporción de contraejemplos presentes en cada regla, reforzando la relación entre *negSupp* y precisión.

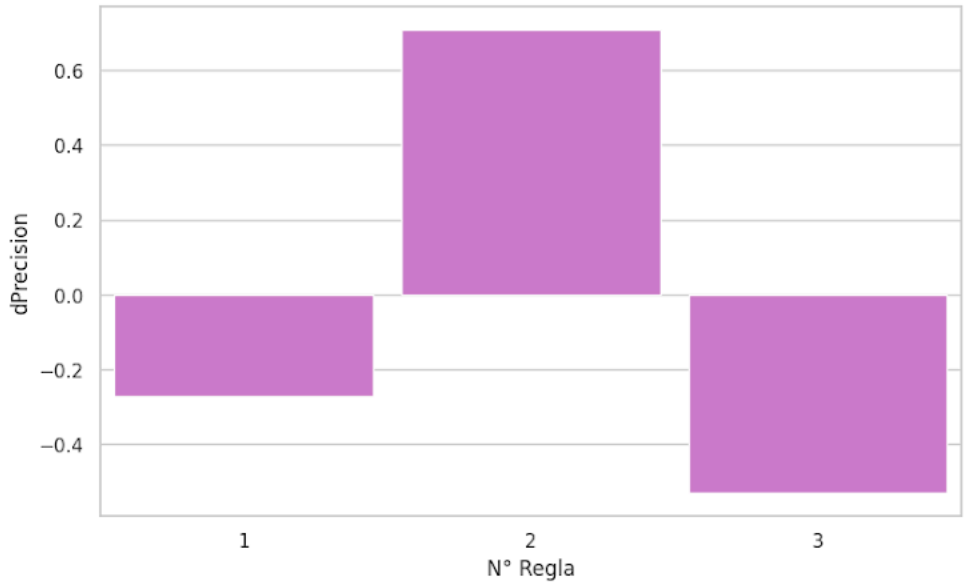


Figura 4.44 *RelativeRisk* – Vista Minable: Diferencia de Precisión por Regla

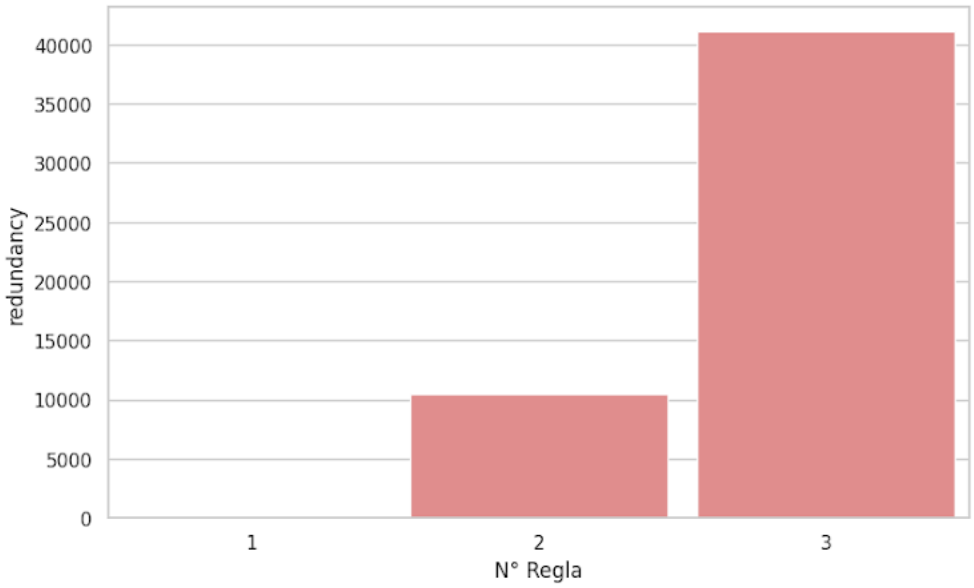


Figura 4.45 *RelativeRisk* – Vista Minable: Redundancia por Regla

En la gráfica correspondiente a la diferencia de precisión, mostrada en la Figura 4.44, se observa una ganancia positiva significativa únicamente en la regla de Tuxtepec, cuyo

incremento supera 0.70. En contraste, las reglas de Río Blanco y del Hospital General de Zona 53 exhiben diferencias negativas, reflejando que su precisión es inferior al comportamiento general del conjunto de datos. Este resultado sugiere que, si bien *RelativeRisk* logra identificar patrones contrastivos, también incorpora reglas con bajo valor discriminativo cuando su soporte es extremadamente alto.

Finalmente, la redundancia representada en la Figura 4.45 muestra una separación clara entre la primera regla, con un valor mínimo de cinco, y las reglas restantes, con valores que superan ampliamente las diez mil unidades y alcanzan picos cercanos a cuarenta mil. Esta diferencia indica que los patrones de H11 y H12 se encuentran en regiones densas del espacio de reglas donde múltiples configuraciones similares producen resultados equivalentes, lo que es consistente con las métricas analizadas previamente.

4.6.4 Evaluación Objetiva

La evaluación objetiva confirma que las tres reglas generadas cumplen con el criterio fundamental de la métrica: todas presentan diferencias de soporte positivas, lo que indica que los patrones son más frecuentes en el hospital correspondiente que en el resto de las instituciones. Sin embargo, la calidad clasificatoria es heterogénea. La regla correspondiente al Hospital General de San Juan Bautista Tuxtepec presenta el mejor desempeño, con precisión elevada, diferencia de precisión positiva y soporte negativo moderado. En contraste, las reglas correspondientes al Hospital Regional Río Blanco y al Hospital General de Zona 53 exhiben diferencias de precisión negativas, evidenciando que la métrica permite la entrada de reglas cuyo soporte elevado no se traduce en un incremento real de precisión. Esto sugiere que, si bien *RelativeRisk* identifica patrones con fuerte presencia estructural, su capacidad discriminativa puede verse limitada en contextos donde los atributos se distribuyen de manera similar entre hospitales.

4.6.5 Evaluación Subjetiva

La evaluación subjetiva realizada por el Dr. Palet reveló diferencias importantes entre la solidez estadística y la plausibilidad clínica de los patrones generados. La regla

correspondiente al Hospital Regional Río Blanco fue descartada al no reflejar un perfil clínico real ni una combinación de formación y características profesionales coherentes con la experiencia del especialista. Aunque el modelo la selecciona por su soporte absoluto, el patrón carece de valor interpretativo en el contexto del hospital.

La regla del Hospital General de San Juan Bautista Tuxtepec fue validada positivamente. Su estructura incorpora múltiples elementos que coinciden con patrones reales observados en la institución, incluyendo la iniciativa del personal para responder la encuesta, la percepción sobre la participación institucional necesaria y los motivos considerados válidos o inválidos para solicitar autopsias. La combinación de estos atributos representa un perfil clínico reconocible, por lo que el especialista la considera un patrón válido.

La regla correspondiente al Hospital General de Zona 53 también fue validada. Pese a su baja precisión, el patrón refleja adecuadamente las posturas y características formativas del personal del hospital, especialmente en lo relativo a la percepción de motivos para solicitar o rechazar autopsias, la ausencia de formación en instituciones específicas y la postura general respecto a la participación institucional. Para el especialista, la amplitud del patrón coincide con la diversidad de perfiles clínicos del hospital, lo que justifica su aceptación.

4.7 Matriz Binaria: Resultados aplicando la función de calidad *Lift*

En este apartado se presentan los resultados del análisis de Minería de Conjuntos de Contraste aplicado al conjunto de datos Matriz Binaria, en nivel específico, utilizando la función de calidad *Lift*. La Matriz Binaria representa cada respuesta categórica del instrumento como variables binarias independientes, lo que permite una descomposición más fina de los atributos y un análisis detallado de la co-ocurrencia entre características específicas y la pertenencia a un centro hospitalario determinado.

La métrica *Lift* se emplea con el objetivo de identificar asociaciones no aleatorias entre los antecedentes y el hospital objetivo, evaluando en qué medida la presencia conjunta

de determinados atributos incrementa la probabilidad de pertenencia a un hospital en comparación con su ocurrencia esperada bajo un supuesto de independencia estadística. De esta forma, el análisis se centra en detectar patrones altamente contrastivos que revelen perfiles institucionales diferenciados a partir de combinaciones binarias de variables, manteniendo la misma estructura metodológica empleada en los análisis previos para garantizar consistencia, trazabilidad y comparabilidad de los resultados.

4.7.1 Reglas sin interpretar

La Tabla 4.11, correspondiente a las reglas sin interpretar generadas mediante la función de calidad *Lift* sobre la Matriz Binaria, muestra un conjunto de cincuenta y tres patrones que se caracterizan por un comportamiento cuantitativamente extremo en términos de precisión y soporte negativo. Las reglas se distribuyen entre los tres hospitales analizados y evidencian una capacidad discriminativa perfecta desde el punto de vista algorítmico, derivada de la naturaleza binaria del conjunto de datos y del criterio de selección impuesto por la métrica *Lift*.

Las reglas asociadas al Hospital Regional Río Blanco conforman un primer bloque integrado por veinticinco patrones, con valores de soporte que oscilan desde niveles bajos, cercanos a 0.07, hasta valores moderados que alcanzan aproximadamente 0.33. En todos los casos, el soporte negativo es igual a cero, lo que implica la ausencia total de contraejemplos en los demás hospitales. Como consecuencia directa, las reglas presentan valores de precisión y diferencia de precisión iguales a 1.0, reflejando una clasificación perfecta del hospital objetivo siempre que se cumple el antecedente. Los valores de redundancia en este conjunto se mantienen relativamente bajos en comparación con los otros hospitales, lo que sugiere un espacio de reglas más compacto y menos solapado.

El conjunto de reglas vinculado al Hospital General de San Juan Bautista Tuxtepec está conformado por doce patrones, con soportes que alcanzan valores cercanos a 0.30 en las reglas más representativas. Al igual que en el bloque anterior, todas las reglas

presentan soporte negativo nulo y precisión perfecta, lo que confirma que las combinaciones binarias identificadas son exclusivas del hospital. No obstante, este conjunto presenta valores de redundancia superiores a los observados en el Hospital Regional Río Blanco, alcanzando niveles intermedios dentro del modelo. Este incremento evidencia la presencia de múltiples reglas con estructuras parcialmente equivalentes, lo que sugiere una mayor densidad combinatoria y un mayor grado de solapamiento interno entre patrones.

Finalmente, las reglas correspondientes al Hospital General de Zona 53 conforman un tercer bloque compuesto por dieciséis patrones, dentro del cual se observan los valores de soporte más altos del modelo, alcanzando aproximadamente 0.37 en la regla de mayor cobertura. Todas las reglas mantienen soporte negativo igual a cero y precisión máxima, confirmando nuevamente la exclusividad de los patrones para el hospital objetivo. En este bloque destacan los valores de redundancia, que superan ampliamente los observados en los otros hospitales y alcanzan niveles muy elevados, lo que evidencia una fuerte concentración de reglas con estructuras similares y una elevada densidad combinatoria en la representación binaria del hospital.

En conjunto, la estructura numérica reflejada en la Tabla 4.11 muestra que la aplicación de *Lift* sobre la Matriz Binaria produce un modelo compuesto por reglas perfectamente discriminativas desde una perspectiva cuantitativa, pero con distintos grados de redundancia entre hospitales. Esta tabla constituye el insumo base para la posterior interpretación semántica de los patrones, donde será posible evaluar si la precisión perfecta observada se traduce en perfiles institucionales clínicamente plausibles o si responde principalmente a efectos derivados de la codificación binaria de los datos.

Tabla 4.11. *Lift* – Matriz Binaria: Reglas Sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF p17_10 = {S} THEN centro_hospitalario = {H1}"	0.236025	0.000000	0.236025	1.000000	1.000000	66.000000
2	"IF p11_6 = {S} THEN centro_hospitalario = {H1}"	0.099379	0.000000	0.099379	1.000000	1.000000	40.000000
3	"IF p6_4 = {S} THEN centro_hospitalario = {H1}"	0.074534	0.000000	0.074534	1.000000	1.000000	56.000000
4	"IF p16_8 = {S} AND p1_3 = !{S} THEN centro_hospitalario = {H1}"	0.329193	0.000000	0.329193	1.000000	1.000000	8.000000
5	"IF p13_1 = {S} AND p8_5 = !{S} THEN centro_hospitalario = {H1}"	0.322981	0.000000	0.322981	1.000000	1.000000	10.000000
6	"IF p5_4 = {S} AND p14_8 = !{S} THEN centro_hospitalario = {H1}"	0.099379	0.000000	0.099379	1.000000	1.000000	31.000000
7	"IF p9_1 = !{S} AND p2_3 = {S} AND p17_107 = !{S} THEN centro_hospitalario = {H1}"	0.086957	0.000000	0.086957	1.000000	1.000000	38.000000
8	"IF p1_5 = {S} AND p12_3 = {S} AND p2_1 = !{S} THEN centro_hospitalario = {H1}"	0.105590	0.000000	0.105590	1.000000	1.000000	36.000000
9	"IF p9_1 = !{S} AND p6_9 = {S} AND p13_8 = !{S} THEN centro_hospitalario = {H1}"	0.080745	0.000000	0.080745	1.000000	1.000000	67.000000
10	"IF p9_1 = !{S} AND p17_23 = {S} AND p5_3 = {S} THEN centro_hospitalario = {H1}"	0.105590	0.000000	0.105590	1.000000	1.000000	42.000000
11	"IF p1_5 = {S} AND p10_-1 = !{S} AND p13_11 = !{S} THEN centro_hospitalario = {H1}"	0.167702	0.000000	0.167702	1.000000	1.000000	25.000000
12	"IF p9_1 = !{S} AND p7_3 = {S} AND p12_2 = !{S} THEN centro_hospitalario = {H1}"	0.118012	0.000000	0.118012	1.000000	1.000000	26.000000
13	"IF p16_13 = {S} AND p17_23 = {S} AND p12_4 = !{S} THEN centro_hospitalario = {H1}"	0.080745	0.000000	0.080745	1.000000	1.000000	33.000000

Tabla 4.11. *Lift* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
14	"IF p9_1 = !{S} AND p7_3 = {S} AND p10_-1 = !{S} THEN centro_hospitalario = {H1}"	0.167702	0.000000	0.167702	1.000000	1.000000	16.000000
15	"IF p1_5 = {S} AND p10_-1 = !{S} AND p6_2 = !{S} THEN centro_hospitalario = {H1}"	0.161491	0.000000	0.161491	1.000000	1.000000	35.000000
16	"IF p9_1 = !{S} AND p17_23 = {S} AND p3_2 = !{S} THEN centro_hospitalario = {H1}"	0.198758	0.000000	0.198758	1.000000	1.000000	26.000000
17	"IF p4_2 = {S} AND p16_13 = {S} AND p17_103 = !{S} AND p9_6 = !{S} THEN centro_hospitalario = {H1}"	0.080745	0.000000	0.080745	1.000000	1.000000	35.000000
18	"IF p9_1 = !{S} AND p9_4 = {S} AND p6_1 = {S} AND p8_6 = !{S} THEN centro_hospitalario = {H1}"	0.192547	0.000000	0.192547	1.000000	1.000000	22.000000
19	"IF p11_4 = !{S} AND p7_5 = {S} AND p8_4 = !{S} THEN centro_hospitalario = {H1}"	0.093168	0.000000	0.093168	1.000000	1.000000	32.000000
20	"IF p9_1 = !{S} AND p1_1 = {S} AND p17_115 = !{S} AND p9_3 = !{S} THEN centro_hospitalario = {H1}"	0.236025	0.000000	0.236025	1.000000	1.000000	50.000000
21	"IF p4_2 = {S} AND p5_1 = {S} AND p1_4 = !{S} AND p16_38 = !{S} THEN centro_hospitalario = {H1}"	0.074534	0.000000	0.074534	1.000000	1.000000	88.000000
22	"IF p11_4 = !{S} AND p12_3 = {S} AND p14_2 = !{S} AND p16_112 = !{S} THEN centro_hospitalario = {H1}"	0.242236	0.000000	0.242236	1.000000	1.000000	7.000000
23	"IF p4_2 = {S} AND p9_4 = {S} AND p15_3 = !{S} AND p17_106 = !{S} THEN centro_hospitalario = {H1}"	0.223602	0.000000	0.223602	1.000000	1.000000	31.000000
24	"IF p11_4 = !{S} AND p17_23 = {S} AND p5_3 = {S} THEN centro_hospitalario = {H1}"	0.118012	0.000000	0.118012	1.000000	1.000000	32.000000

Tabla 4.11. *Lift* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
25	"IF p12_7 = !{S} AND p13_11 = !{S} AND p9_3 = !{S} AND p5_2 = {S} AND p17_38 = !{S} THEN centro_hospitalario = {H1}"	0.204969	0.000000	0.204969	1.000000	1.000000	19.000000
26	"IF p17_103 = {S} THEN centro_hospitalario = {H11}"	0.140625	0.000000	0.140625	1.000000	1.000000	203.000000
27	"IF p17_107 = {S} AND p4_3 = !{S} THEN centro_hospitalario = {H11}"	0.296875	0.000000	0.296875	1.000000	1.000000	196.000000
28	"IF p16_99 = {S} AND p13_3 = !{S} THEN centro_hospitalario = {H11}"	0.187500	0.000000	0.187500	1.000000	1.000000	180.000000
29	"IF p16_93 = {S} AND p3_2 = !{S} THEN centro_hospitalario = {H11}"	0.234375	0.000000	0.234375	1.000000	1.000000	187.000000
30	"IF p12_8 = {S} AND p8_4 = {S} THEN centro_hospitalario = {H11}"	0.281250	0.000000	0.281250	1.000000	1.000000	186.000000
31	"IF p10_2 = {S} AND p9_1 = {S} THEN centro_hospitalario = {H11}"	0.265625	0.000000	0.265625	1.000000	1.000000	183.000000
32	"IF p7_17 = {S} AND p11_4 = {S} THEN centro_hospitalario = {H11}"	0.093750	0.000000	0.093750	1.000000	1.000000	90.000000
33	"IF p12_7 = {S} AND p1_3 = {S} THEN centro_hospitalario = {H11}"	0.140625	0.000000	0.140625	1.000000	1.000000	210.000000
34	"IF p12_7 = {S} AND p1_5 = {S} THEN centro_hospitalario = {H11}"	0.078125	0.000000	0.078125	1.000000	1.000000	90.000000
35	"IF p11_3 = {S} AND p5_1 = {S} AND p10_4 = {S} THEN centro_hospitalario = {H11}"	0.234375	0.000000	0.234375	1.000000	1.000000	181.000000
36	"IF p8_5 = {S} AND p10_4 = {S} AND p11_2 = {S} THEN centro_hospitalario = {H11}"	0.187500	0.000000	0.187500	1.000000	1.000000	186.000000

Tabla 4.11. *Lift* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
37	"IF p9_1 = {S} AND p13_2 = {S} AND p10_-1 = {S} AND p1_1 = !{S} THEN centro_hospitalario = {H11}"	0.078125	0.000000	0.078125	1.000000	1.000000	7.000000
38	"IF p16_131 = {S} THEN centro_hospitalario = {H12}"	0.232143	0.000000	0.232143	1.000000	1.000000	271.000000
39	"IF p16_129 = {S} THEN centro_hospitalario = {H12}"	0.107143	0.000000	0.107143	1.000000	1.000000	175.000000
40	"IF p17_115 = {S} AND p9_6 = !{S} THEN centro_hospitalario = {H12}"	0.375000	0.000000	0.375000	1.000000	1.000000	251.000000
41	"IF p16_112 = {S} AND p3_2 = {S} THEN centro_hospitalario = {H12}"	0.214286	0.000000	0.214286	1.000000	1.000000	239.000000
42	"IF p17_38 = {S} AND p2_2 = !{S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	90.000000
43	"IF p16_38 = {S} AND p14_8 = {S} AND p2_1 = {S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	90.000000
44	"IF p11_-1 = {S} AND p14_1 = {S} AND p13_11 = {S} THEN centro_hospitalario = {H12}"	0.107143	0.000000	0.107143	1.000000	1.000000	90.000000
45	"IF p16_38 = {S} AND p2_1 = {S} AND p10_4 = !{S} THEN centro_hospitalario = {H12}"	0.107143	0.000000	0.107143	1.000000	1.000000	90.000000
46	"IF p11_-1 = {S} AND p11_1 = !{S} AND p10_-1 = !{S} AND p4_4 = !{S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	7.000000
47	"IF p16_38 = {S} AND p12_1 = !{S} AND p3_1 = !{S} THEN centro_hospitalario = {H12}"	0.107143	0.000000	0.107143	1.000000	1.000000	90.000000
48	"IF p11_1 = !{S} AND p3_4 = {S} AND p5_2 = {S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	90.000000
49	"IF p11_1 = !{S} AND p8_3 = {S} AND p15_2 = {S} AND p3_1 = !{S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	7.000000

Tabla 4.11. *Lift* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
50	"IF p3_2 = {S} AND p9_6 = {S} AND p8_5 = {S} THEN centro_hospitalario = {H12}"	0.089286	0.000000	0.089286	1.000000	1.000000	90.0000000
51	"IF p3_2 = {S} AND p1_3 = {S} AND p14_2 = {S} AND p15_3 = {S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	7.0000000
52	"IF p7_6 = {S} AND p9_6 = {S} AND p5_3 = {S} AND p1_1 = !{S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	7.0000000
53	"IF p7_6 = {S} AND p1_2 = {S} AND p15_3 = {S} AND p6_2 = !{S} THEN centro_hospitalario = {H12}"	0.071429	0.000000	0.071429	1.000000	1.000000	7.0000000

4.7.2 Reglas Interpretadas

La interpretación semántica de las reglas generadas mediante la función de calidad *Lift*, aplicada a la Matriz Binaria en nivel específico, permite transformar los patrones algorítmicos en descripciones comprensibles desde una perspectiva clínica e institucional. A partir de este proceso se obtuvieron cincuenta y tres reglas interpretadas, las cuales describen perfiles contrastivos claramente diferenciados entre los tres hospitales analizados y reflejan combinaciones coherentes de formación académica, experiencia profesional y percepciones en torno a la práctica de la autopsia.

Las reglas asociadas al Hospital Regional Río Blanco conforman el bloque más amplio, integrado por veinticinco patrones interpretados. En este conjunto destacan reglas simples, de una sola condición, vinculadas directamente con la formación académica y los motivos para solicitar autopsia, como la especialidad cursada en determinadas instituciones o la orientación diagnóstica y la aplicabilidad del estudio como razones centrales para su solicitud. Estas reglas de estructura corta coexisten con patrones más complejos que incorporan múltiples elementos relacionados con la experiencia

profesional, la participación en estudios de autopsia, las percepciones sobre demandas y arbitraje, así como la negativa de los familiares y la disponibilidad de recursos humanos o financieros. En conjunto, este bloque describe un perfil institucional en el que la práctica de la autopsia se asocia tanto con trayectorias formativas específicas como con una postura activa frente a su utilidad diagnóstica y medicolegal.

El conjunto de reglas interpretadas correspondiente al Hospital General de San Juan Bautista Tuxtepec está conformado por doce patrones que, en su mayoría, presentan estructuras de complejidad moderada. Estas reglas enfatizan aspectos relacionados con la institución donde se estudió la especialidad o medicina general, así como con las opiniones sobre el método más eficiente para solicitar la autopsia, incluyendo la participación de áreas como psicología o patología, y el rol de la institución o del personal de enfermería en dicho proceso. Asimismo, aparecen de manera recurrente elementos vinculados con la negligencia médica, el error médico y la percepción sobre la suficiencia de recursos humanos. Este conjunto de reglas sugiere un perfil institucional caracterizado por una visión normativa y procedimental de la práctica de la autopsia, donde los mecanismos formales y las responsabilidades institucionales ocupan un papel central.

Finalmente, las reglas interpretadas asociadas al Hospital General de Zona 53 conforman un bloque de dieciséis patrones que se distinguen por la diversidad de combinaciones y por una mayor heterogeneidad en los perfiles descritos. En este conjunto destacan reglas relacionadas con la formación en instituciones como UNAM, FES Zaragoza, Hospital Juárez o Universidad del Valle de México, así como con el rol profesional de los médicos (residentes, adscritos, médicos generales o especialistas). De manera consistente, estas reglas incorporan percepciones sobre la discrepancia entre hallazgos de autopsia y diagnósticos clínicos, la participación limitada en estudios de autopsia, y diversos motivos para no solicitar el procedimiento, como el interés, las enfermedades de base conocidas o consideraciones morales por parte de los familiares. Este conjunto refleja un perfil institucional amplio y diverso, coherente con la variedad de posturas y trayectorias presentes en el hospital.

En conjunto, como se muestra en la Tabla 4.12 las reglas interpretadas muestran que la aplicación de *Lift* sobre la Matriz Binaria permite identificar patrones semánticamente claros y perfectamente discriminativos desde el punto de vista cuantitativo, los cuales describen perfiles institucionales diferenciados basados en combinaciones específicas de formación, experiencia y percepciones clínicas. Esta interpretación constituye el paso previo indispensable para el análisis visual y la evaluación objetiva y subjetiva del modelo, donde se valorará si la precisión perfecta observada se traduce en conocimiento útil y clínicamente significativo.

Tabla 4.12. *Lift* – Matriz Binaria: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos estudiaron su especialidad en Facultad de Medicina; entonces son del Hospital Regional Río Blanco.
2	Si los médicos estiman como motivo para solicitar autopsia la orientación diagnóstica; entonces son del Hospital Regional Río Blanco.
3	Si los médicos plantean como motivo para solicitar autopsia la aplicabilidad del estudio; entonces son del Hospital Regional Río Blanco.
4	Si los médicos estudiaron medicina general en Alta Especialidad en IMSS y no tienen 11-15 años de práctica; entonces son del Hospital Regional Río Blanco.
5	Si los médicos no comentan y no creen que la causa de rechazo por parte del familiar es por inconformidad de los familiares con la atención prestada al paciente; entonces son del Hospital Regional Río Blanco.
6	Si los médicos están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda y no son estudiantes de medicina; entonces son del Hospital Regional Río Blanco.
7	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, han participado entre 6 y 10 casos de autopsia, y sí especifican dónde realizaron su especialidad; entonces son del Hospital Regional Río Blanco.
8	Si los médicos tienen más de 20 años de práctica, están de acuerdo en que el método eficiente para solicitar autopsia es que la institución a nivel regional o estatal reglamente la práctica de autopsias de forma obligatoria a los pacientes hospitalizados que fallezcan, y sí han participado en algún estudio de autopsia; entonces son del Hospital Regional Río Blanco.
9	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, plantean como motivo para solicitar autopsia el diagnóstico difícil, y no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general; entonces son del Hospital Regional Río Blanco.

Tabla 4.12. *Lift* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
10	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, estudiaron su especialidad en La Raza IMSS, y consideran incierto que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital Regional Río Blanco.
11	Si los médicos tienen más de 20 años de práctica y no agradecen por la encuesta; entonces son del Hospital Regional Río Blanco.
12	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, piensan que una causa o motivo para no solicitar autopsia es la negativa de los familiares, y no están de acuerdo en que el método eficiente para solicitar autopsia es solicitar el consentimiento de autopsia a los familiares inmediatamente después de la defunción; entonces son del Hospital Regional Río Blanco.
13	Si los médicos estudiaron medicina general en Universidad de San Carlos, estudiaron su especialidad en La Raza IMSS, y no están de acuerdo en que el método eficiente para solicitar autopsia es que el médico tratante sea quién la solicite; entonces son del Hospital Regional Río Blanco.
14	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos y piensan que una causa o motivo para no solicitar autopsia es la negativa de los familiares; entonces son del Hospital Regional Río Blanco.
15	Si los médicos tienen más de 20 años de práctica y no plantean como motivo para solicitar autopsia cuando existe relación con temas relacionados con medicina forense; entonces son del Hospital Regional Río Blanco.
16	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, estudiaron su especialidad en La Raza IMSS, y no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital Regional Río Blanco.
17	Si los médicos están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, estudiaron medicina general en Universidad de San Carlos, no estudiaron su especialidad en IMSS, y no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica; entonces son del Hospital Regional Río Blanco.

Tabla 4.12. *Lift* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
18	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, consideran que no se realizan suficientes autopsias en el hospital debido a que no se solicitan, plantean como motivo para solicitar autopsia establecer diagnósticos definitivos, y no creen que la causa de rechazo por parte del familiar es por deficiente comunicación del médico con el paciente y sus familiares; entonces son del Hospital Regional Río Blanco.
19	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, piensan que una causa o motivo para no solicitar autopsia es el desinterés, y no creen que la causa de rechazo por parte del familiar es por no haber solicitado la autorización en forma adecuada; entonces son del Hospital Regional Río Blanco.
20	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, tienen menos de 5 años de práctica, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, y no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos financieros; entonces son del Hospital Regional Río Blanco.
21	Si los médicos están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no tienen 16-20 años de práctica, y no estudiaron medicina general en UNAM; entonces son del Hospital Regional Río Blanco.
22	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, están de acuerdo en que el método eficiente para solicitar autopsia es que la institución a nivel regional o estatal reglamente la práctica de autopsias de forma obligatoria a los pacientes hospitalizados que fallezcan, no son médicos generales, y no estudiaron medicina general en Universidad del Valle de México; entonces son del Hospital Regional Río Blanco.
23	Si los médicos están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, consideran que no se realizan suficientes autopsias en el hospital debido a que no se solicitan, no son adscritos, y no estudiaron su especialidad en Hospital General Balbuena; entonces son del Hospital Regional Río Blanco.
24	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, estudiaron su especialidad en La Raza IMSS, y consideran incierto que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital Regional Río Blanco.
25	Si los médicos no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no agradecen por la encuesta, no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos financieros, están de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, y no realizaron su especialidad en la UNAM; entonces son del Hospital Regional Río Blanco.

Tabla 4.12. *Lift* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
26	Si los médicos estudiaron su especialidad en IMSS; entonces son del Hospital General de San Juan Bautista Tuxtepec.
27	Si los médicos no especifican dónde realizaron su especialidad y no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje; entonces son del Hospital General de San Juan Bautista Tuxtepec.
28	Si los médicos estudiaron medicina general en Secretaría de Salud de Oaxaca y no opinan que es inadecuada la enseñanza de la patología general; entonces son del Hospital General de San Juan Bautista Tuxtepec.
29	Si los médicos estudiaron medicina general en Universidad Anáhuac y no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de San Juan Bautista Tuxtepec.
30	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología y creen que la causa de rechazo por parte del familiar es por no haber solicitado la autorización en forma adecuada; entonces son del Hospital General de San Juan Bautista Tuxtepec.
31	Si los médicos consideran que quien debe solicitar la autopsia es una enfermera y consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos; entonces son del Hospital General de San Juan Bautista Tuxtepec.
32	Si los médicos piensan que una causa o motivo para no solicitar autopsia es que se haga sin fines de enseñanza y estiman como motivo para solicitar autopsia la negligencia médica; entonces son del Hospital General de San Juan Bautista Tuxtepec.
33	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología y tienen 11-15 años de práctica; entonces son del Hospital General de San Juan Bautista Tuxtepec.
34	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología y tienen más de 20 años de práctica; entonces son del Hospital General de San Juan Bautista Tuxtepec.
35	Si los médicos estiman como motivo para solicitar autopsia el consumo de medicamentos mal recetados, están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, y consideran que quien debe solicitar la autopsia es una institución; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.12. *Lift* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
36	Si los médicos creen que la causa de rechazo por parte del familiar es por inconformidad de los familiares con la atención prestada al paciente, consideran que quien debe solicitar la autopsia es una institución, y estiman como motivo para solicitar autopsia un error médico; entonces son del Hospital General de San Juan Bautista Tuxtepec.
37	Si los médicos consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, opinan que este estudio debería ser rutinario, y no tienen menos de 5 años de práctica; entonces son del Hospital General de San Juan Bautista Tuxtepec.
38	Si los médicos realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de Zona 53.
39	Si los médicos estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de Zona 53.
40	Si los médicos estudiaron su especialidad en Hospital Hermanos Ameijeiras y no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica; entonces son del Hospital General de Zona 53.
41	Si los médicos estudiaron medicina general en Universidad del Valle de México y están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de Zona 53.
42	Si los médicos realizaron su especialidad en la UNAM y no han participado en menos de 5 casos de autopsias; entonces son del Hospital General de Zona 53.
43	Si los médicos estudiaron medicina general en UNAM, son estudiantes de medicina, y no han participado en algún estudio de autopsia; entonces son del Hospital General de Zona 53.
44	Si los médicos son especialistas y agradecen por la encuesta; entonces son del Hospital General de Zona 53.
45	Si los médicos estudiaron medicina general en UNAM, no han participado en algún estudio de autopsia, y no consideran que quien debe solicitar la autopsia es una institución; entonces son del Hospital General de Zona 53.
46	Si los médicos no estiman como motivo para solicitar autopsia el interés y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje; entonces son del Hospital General de Zona 53.
47	Si los médicos estudiaron medicina general en UNAM, no están de acuerdo en que el método eficiente para solicitar autopsia es pedir el consentimiento al familiar del paciente, desde que ingresa al hospital, y no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de Zona 53.

Tabla 4.12. *Lift* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
48	Si los médicos no estiman como motivo para solicitar autopsia el interés, están en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, y están de acuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.
49	Si los médicos no estiman como motivo para solicitar autopsia el interés, creen que la causa de rechazo por parte del familiar es por motivos morales, son residentes, y no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de Zona 53.
50	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, y creen que la causa de rechazo por parte del familiar es por inconformidad de los familiares con la atención prestada al paciente; entonces son del Hospital General de Zona 53.
51	Si los médicos están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, tienen 11-15 años de práctica, son médicos generales, y son adscritos; entonces son del Hospital General de Zona 53.
52	Si los médicos piensan que una causa o motivo para no solicitar autopsia es por enfermedad de base conocida, consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, consideran incierto que los hallazgos de las autopsias dan origen a casos de demanda, y no tienen menos de 5 años de práctica; entonces son del Hospital General de Zona 53.
53	Si los médicos piensan que una causa o motivo para no solicitar autopsia es por enfermedad de base conocida, tienen 5-10 años de práctica, son adscritos, y no plantean como motivo para solicitar autopsia cuando existe relación con temas relacionados con medicina forense; entonces son del Hospital General de Zona 53.

4.7.3 Análisis visual

El análisis visual complementa la evaluación numérica de las reglas generadas mediante la función de calidad *Lift*, permitiendo observar de manera gráfica el comportamiento de las métricas asociadas a cada patrón. A través de estas representaciones se identifican

tendencias, concentraciones y variaciones en la cobertura, discriminación y redundancia de las reglas, facilitando una lectura global del modelo obtenido sobre la Matriz Binaria.

La Figura 4.46 muestra la distribución de los valores de soporte para las cincuenta y tres reglas generadas mediante la función de calidad *Lift*. Se observa una variabilidad considerable en la cobertura de los patrones, con valores que oscilan desde niveles bajos, cercanos a 0.07, hasta valores más elevados que alcanzan aproximadamente 0.37. Este comportamiento indica la coexistencia de reglas que describen subgrupos muy específicos junto con otras que representan segmentos más amplios de la población médica dentro de los hospitales analizados.

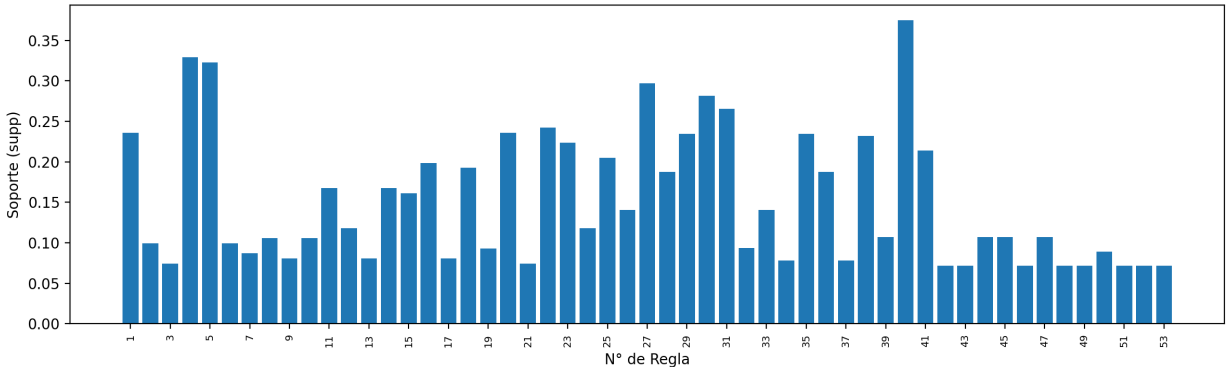


Figura 4.46 *Lift* – Matriz Binaria: Soporte por Regla

Destacan un conjunto reducido de reglas con soportes moderados y altos, lo que sugiere la presencia de combinaciones binarias particularmente representativas en ciertos hospitales. Al mismo tiempo, la mayoría de las reglas presenta soportes bajos a intermedios, reflejando el carácter altamente específico de muchos de los patrones identificados en la representación binaria del conjunto de datos.

La Figura 4.47 presenta el comportamiento del soporte negativo para las cincuenta y tres reglas generadas mediante la función de calidad *Lift* sobre la Matriz Binaria. En la gráfica se observa que todas las reglas presentan valores de soporte negativo iguales a cero, lo que indica la ausencia total de contraejemplos en los hospitales distintos al hospital objetivo.

Este comportamiento refleja que ninguna de las combinaciones binarias identificadas en los antecedentes de las reglas aparece en médicos pertenecientes a otros hospitales, lo que confirma el carácter exclusivo de los patrones descubiertos. La inexistencia de soporte negativo refuerza la capacidad discriminativa del modelo y explica los valores máximos de precisión observados en las reglas, evidenciando una separación perfecta entre los perfiles hospitalarios a partir de la representación binaria de los datos.

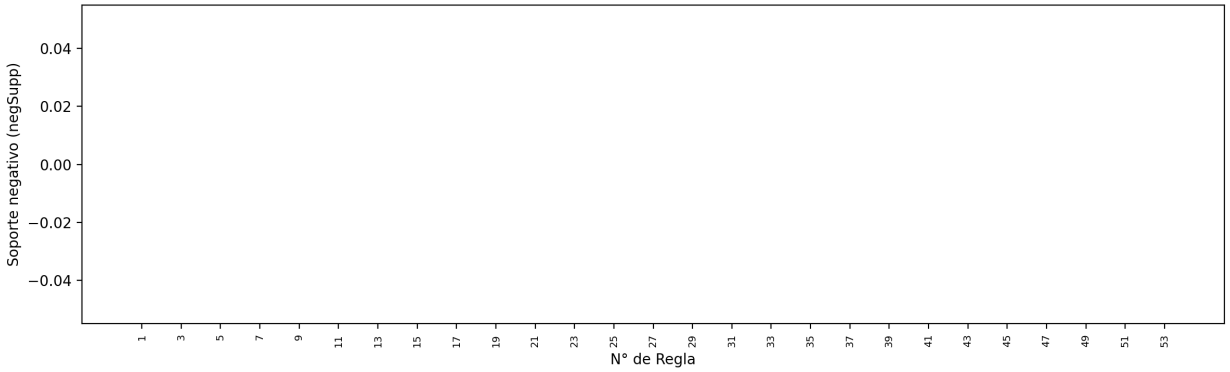


Figura 4.47 *Lift* – Matriz Binaria: Soporte Negativo por Regla

La Figura 4.48 muestra la diferencia de soporte asociada a cada una de las cincuenta y tres reglas generadas mediante la función de calidad *Lift*. Dado que el soporte negativo es nulo en todos los casos, la diferencia de soporte coincide directamente con los valores de soporte observados para cada regla.

En la gráfica se aprecia una variabilidad clara entre patrones, con valores que van desde niveles bajos, cercanos a 0.07, hasta valores más elevados que alcanzan aproximadamente 0.37. Este comportamiento indica que algunas reglas representan subgrupos altamente característicos del hospital objetivo, mientras que otras describen segmentos más específicos de la población. La presencia sistemática de diferencias de soporte positivas confirma que todos los patrones identificados contribuyen de manera efectiva a la discriminación entre hospitales dentro de la representación binaria del conjunto de datos.

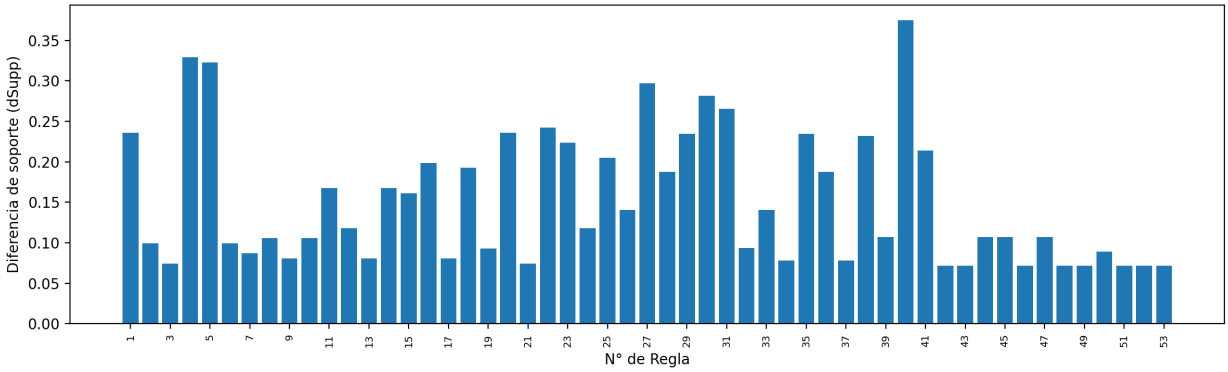


Figura 4.48 *Lift* – Matriz Binaria: Diferencia de Soporte por Regla

La Figura 4.49 muestra los valores de precisión asociados a cada una de las cincuenta y tres reglas generadas mediante la función de calidad *Lift*. En la gráfica se observa que todas las reglas alcanzan un valor de precisión igual a 1.0, lo que indica que, en todos los casos, la activación del antecedente conduce de manera inequívoca a la correcta identificación del hospital objetivo.

Este comportamiento confirma la ausencia total de clasificaciones erróneas dentro del modelo y es consistente con los valores nulos de soporte negativo observados previamente. La precisión perfecta refleja una separación completa entre los hospitales en el espacio binario de atributos, evidenciando la elevada capacidad discriminativa del modelo bajo esta representación de los datos.

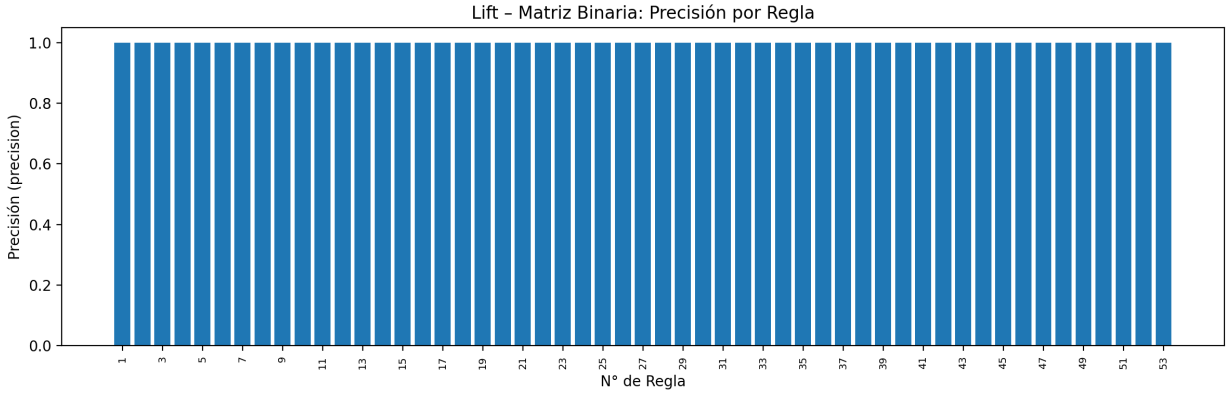


Figura 4.49 *Lift* – Matriz Binaria: Precisión por Regla

La Figura 4.50 presenta los valores de diferencia de precisión correspondientes a cada una de las cincuenta y tres reglas generadas mediante la función de calidad *Lift*. En la gráfica se observa que todas las reglas alcanzan un valor de diferencia de precisión igual a 1.0, lo que indica una ganancia máxima respecto al comportamiento global del conjunto de datos.

Este resultado confirma que la precisión de cada regla supera de manera absoluta la precisión base, reforzando el carácter perfectamente discriminativo del modelo. La uniformidad observada en la gráfica refleja una separación total entre los hospitales cuando se cumplen los antecedentes de las reglas, consolidando la coherencia entre las métricas de soporte negativo, precisión y diferencia de precisión dentro del análisis basado en la Matriz Binaria.

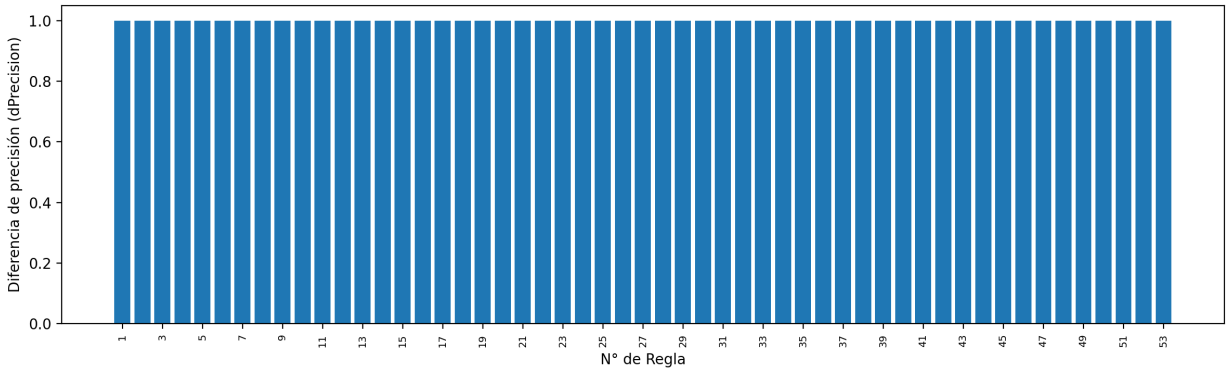


Figura 4.50 *Lift* – Matriz Binaria: Diferencia de Precisión por Regla

La Figura 4.51 muestra el comportamiento de la redundancia en las cincuenta y tres reglas generadas mediante la función de calidad Lift, evidenciando una distribución diferenciada en varios rangos de magnitud. Los valores oscilan desde niveles muy bajos, cercanos a 5 unidades, hasta máximos que superan los 270, lo que refleja una variabilidad considerable en el grado de solapamiento estructural entre patrones.

Las reglas asociadas al Hospital Regional Río Blanco se concentran principalmente en el rango bajo, con valores mayoritariamente inferiores a 50, indicando un conjunto relativamente compacto y con menor redundancia combinatoria. En contraste, las reglas

correspondientes al Hospital General de San Juan Bautista Tuxtepec se ubican en un rango intermedio–alto, mientras que el Hospital General de Zona 53 concentra los valores más elevados del modelo, incluyendo los picos máximos observados, lo que evidencia una mayor densidad de reglas estructuralmente similares.

En conjunto, aunque el modelo alcanza una discriminación perfecta en términos de precisión, la distribución de la redundancia revela distintos niveles de solapamiento entre hospitales, consolidándose como un criterio relevante para la depuración posterior del conjunto de reglas.

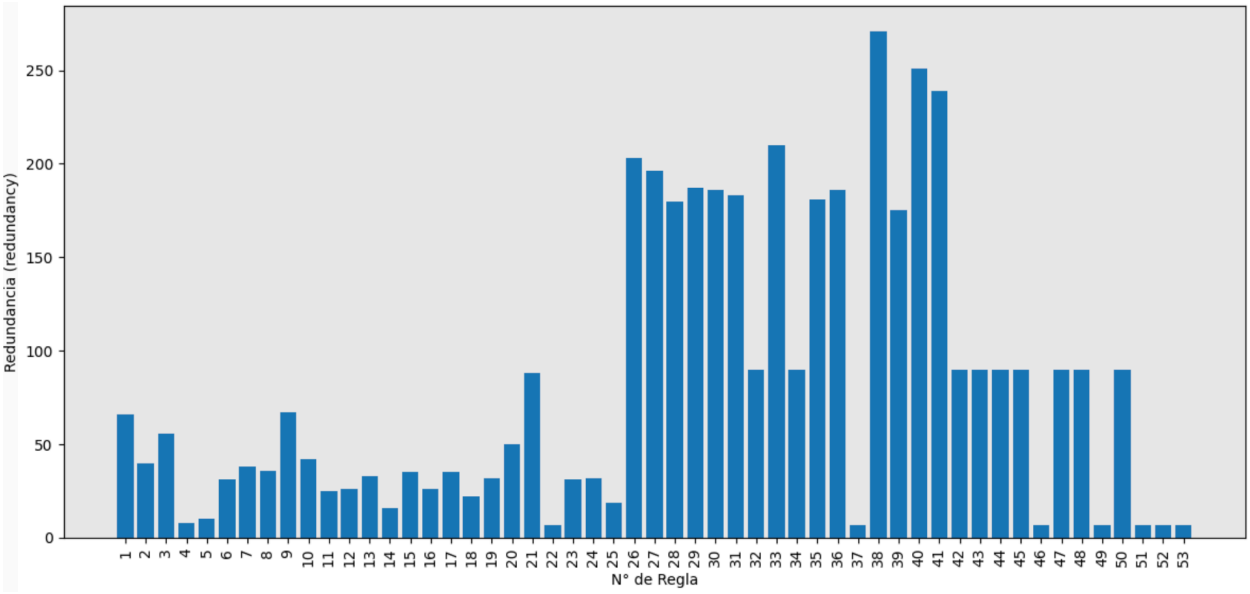


Figura 4.51 *Lift* – Matriz Binaria: Redundancia por Regla

4.7.4 Evaluación objetiva

La evaluación objetiva del modelo generado mediante la función de calidad *Lift*, aplicada al conjunto de datos Matriz Binaria en nivel específico, se centra en el análisis cuantitativo de las métricas asociadas a las reglas descubiertas, con el propósito de determinar su capacidad discriminativa, consistencia interna y estabilidad estadística.

En primer lugar, un rasgo distintivo del modelo es que todas las reglas presentan un soporte negativo igual a cero, lo que indica la ausencia total de contraejemplos. Este

comportamiento implica que ninguna de las combinaciones binarias identificadas en los antecedentes de las reglas aparece en médicos pertenecientes a hospitales distintos del hospital objetivo, confirmando la exclusividad de los patrones descubiertos.

Como consecuencia directa de lo anterior, la precisión y la diferencia de precisión alcanzan el valor máximo (1.0) en todas las reglas, lo que evidencia una clasificación perfecta del hospital objetivo siempre que se cumple el antecedente. Desde una perspectiva estrictamente cuantitativa, este resultado refleja una separación completa entre los hospitales en el espacio de atributos binarios, lo que constituye un desempeño extremo en términos de discriminación.

El análisis del soporte y la diferencia de soporte muestra una variabilidad significativa entre reglas, con valores que oscilan desde niveles bajos, cercanos a 0.07, hasta valores más elevados que alcanzan aproximadamente 0.37. Esta variabilidad indica que, aunque todas las reglas son perfectamente precisas, no todas poseen el mismo grado de representatividad dentro del hospital correspondiente, coexistiendo patrones altamente específicos con otros de mayor cobertura.

Finalmente, la métrica de redundancia introduce un matiz importante en la evaluación objetiva del modelo. Mientras que algunas reglas presentan valores de redundancia bajos, otras alcanzan valores considerablemente elevados, especialmente en los hospitales H11 y H12. Este comportamiento sugiere la presencia de múltiples reglas que describen regiones similares del espacio de búsqueda, lo que, si bien no afecta la capacidad discriminativa del modelo, sí señala la existencia de solapamiento informativo entre patrones.

En conjunto, la evaluación objetiva muestra que el modelo *Lift* aplicado a la Matriz Binaria posee una capacidad discriminativa perfecta desde el punto de vista cuantitativo, sustentada en la ausencia de contraejemplos y en la precisión máxima de todas las reglas. No obstante, la variabilidad en el soporte y los niveles elevados de redundancia en una parte del conjunto evidencian la necesidad de complementar este análisis con

una evaluación subjetiva que permita valorar la relevancia y utilidad práctica de los patrones identificados.

4.7.5 Evaluación subjetiva

La evaluación subjetiva realizada por el Dr. Palet evidenció que la función de calidad *Lift* genera un volumen amplio de patrones con alta significancia estadística, aunque con una variabilidad considerable en su plausibilidad clínica. Una proporción mayoritaria de las reglas fue considerada válida, al describir perfiles clínicos coherentes con la práctica observada en el Hospital Regional Río Blanco, el Hospital General de San Juan Bautista Tuxtepec y el Hospital General de Zona 53, integrando de forma consistente la formación médica en escuelas públicas, la experiencia profesional y las percepciones institucionales sobre la práctica de la autopsia.

No obstante, el especialista identificó un conjunto reducido de reglas que fueron descartadas al no reflejar perfiles clínicos reconocibles o al presentar combinaciones excesivamente generales y redundantes. En estos casos, aunque los patrones destacan por su valor estadístico, carecen de utilidad interpretativa dentro del contexto hospitalario específico. Desde la perspectiva clínica, *Lift* resulta particularmente útil como herramienta exploratoria, siempre que se complemente con un proceso de validación experta que permita priorizar las reglas con mayor significado institucional.

4.8 Matriz Binaria: Resultados aplicando la función de calidad *WRAcc*

En este apartado se presentan los resultados del análisis de Minería de Conjuntos de Contraste aplicado al conjunto de datos Matriz Binaria, en nivel específico, utilizando la función de calidad *WRAcc*. Esta métrica permite identificar patrones contrastivos que, además de discriminar entre hospitales, favorecen reglas con una cobertura representativa y una contribución real a la separación entre clases. Bajo este enfoque, las reglas obtenidas se analizan de forma consistente mediante su presentación sin interpretación, su interpretación semántica, el análisis visual de métricas y la evaluación objetiva y subjetiva del modelo.

4.8.1 Reglas sin interpretar

La Tabla 4.13, correspondiente a las reglas sin interpretar generadas mediante *WRAcc*, muestra un conjunto de veintidós patrones distribuidos entre los tres hospitales analizados. En términos generales, el modelo presenta reglas con soportes predominantemente moderados y altos, diferencias de soporte consistentemente positivas y precisiones elevadas, lo que evidencia un comportamiento discriminativo sólido, aunque no extremo como en otros escenarios donde el soporte negativo resulta nulo para todas las reglas.

Las reglas asociadas al Hospital Regional Río Blanco conforman un primer bloque integrado por seis patrones, con soportes que oscilan aproximadamente entre 0.68 y 0.92. En este conjunto se observa un comportamiento mixto: existen reglas con soporte negativo menor a 0.08 y precisión superior a 0.93, pero también reglas con valores de soporte negativo relativamente mayores (hasta alrededor de 0.22), lo que reduce la precisión a rangos cercanos a 0.85. Aun así, la diferencia de soporte se mantiene elevada (en torno a 0.64–0.79), reflejando patrones marcadamente característicos del hospital. La redundancia en este bloque se mantiene baja (aproximadamente entre 5 y 9), lo que sugiere un conjunto compacto con menor solapamiento informativo.

El conjunto de reglas vinculado al Hospital General de San Juan Bautista Tuxtepec constituye el bloque más amplio, conformado por nueve patrones, con soportes que van aproximadamente de 0.42 a 0.88. En este hospital el soporte negativo se mantiene generalmente bajo (hasta alrededor de 0.06), lo que permite precisiones altas, aunque existen reglas donde la precisión desciende a valores cercanos a 0.81. Las diferencias de soporte permanecen consistentemente positivas y, en los casos más representativos, alcanzan valores cercanos a 0.82. Un rasgo notable en este bloque es la redundancia persistentemente alta (aproximadamente entre 166 y 181), lo que indica que múltiples reglas describen regiones muy similares del espacio de búsqueda y generan patrones con información parcialmente equivalente.

Finalmente, las reglas asociadas al Hospital General de Zona 53 conforman un bloque de siete patrones, con el mayor rango de variabilidad en soporte, desde valores cercanos a 0.11 hasta valores máximos que alcanzan 1.00. En este conjunto se observan reglas completamente representativas (con soporte máximo), pero también reglas de cobertura reducida. El soporte negativo puede alcanzar valores cercanos a 0.08, lo que implica que no todas las reglas son exclusivas del hospital y que la precisión puede descender hasta valores aproximados de 0.75. A pesar de ello, las diferencias de soporte se mantienen positivas y pueden alcanzar valores cercanos a 0.92, confirmando un comportamiento contrastivo fuerte. La redundancia en este bloque es también elevada (aproximadamente entre 229 y 271), evidenciando un solapamiento estructural importante entre reglas.

En conjunto, la Tabla 4.13 evidencia que el modelo *WRAcc* sobre Matriz Binaria produce reglas con alta capacidad discriminativa y con cobertura relevante, aunque con distintos niveles de exclusividad entre hospitales y con redundancias notablemente altas en H11 y H12. Esta tabla constituye el insumo base para su posterior interpretación semántica y para el análisis visual de las métricas del modelo.

Tabla 4.13. *WRAcc* – Matriz Binaria: Reglas Sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF p9_1 = !{S} AND p17_115 = !{S} AND p17_107 = !{S} AND p16_112 = !{S} AND p17_38 = !{S} AND p16_86 = !{S} AND p2_4 = !{S} AND p12_-1 = !{S} AND p17_102 = !{S} AND p16_121 = !{S} AND p17_123 = !{S} AND p16_124 = !{S} AND p16_18 = !{S} AND p16_129 = !{S} AND p16_23 = !{S} THEN centro_hospitalario = {H1}"	0.720497	0.000000	0.720497	1.000000	1.000000	8.000000
2	"IF p17_115 = !{S} AND p17_103 = !{S} AND p17_117 = !{S} AND p17_122 = !{S} AND p17_126 = !{S} AND p12_6 = !{S} AND p7_16 = !{S} AND p12_8 = !{S} AND p17_100 = !{S} AND p17_101 = !{S} AND p17_105 = !{S} AND p16_108 = !{S} AND p17_109 = !{S} AND p17_110 = !{S} AND p17_111 = !{S} AND p17_114 = !{S} AND p17_119 = !{S} AND p17_124 = !{S} AND p17_128 = !{S} AND p17_129 = !{S} AND p17_130 = !{S} AND p17_125 = !{S} AND p16_131 = !{S} AND p17_131 = !{S} AND p17_133 = !{S} THEN centro_hospitalario = {H1}"	0.919255	0.216667	0.702588	0.850575	0.701149	5.000000
3	"IF p9_1 = !{S} AND p17_115 = !{S} AND p9_6 = !{S} AND p17_38 = !{S} AND p16_112 = !{S} AND p16_114 = !{S} AND p13_15 = !{S} THEN centro_hospitalario = {H1}"	0.677019	0.041667	0.635352	0.956140	0.912281	8.000000
4	"IF p17_115 = !{S} AND p12_7 = !{S} AND p16_112 = !{S} AND p17_107 = !{S} AND p16_129 = !{S} AND p16_99 = !{S} AND p12_-1 = !{S} AND p17_104 = !{S} AND p17_106 = !{S} AND p16_120 = !{S} THEN centro_hospitalario = {H1}"	0.913043	0.200000	0.713043	0.859649	0.719298	5.000000

Tabla 4.13. *WRAcc* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
5	"IF p11_4 = !{S} AND p16_112 = !{S} AND p16_99 = !{S} AND p1_3 = !{S} AND p17_118 = !{S} AND p7_1 = !{S} AND p10_2 = !{S} AND p17_101 = !{S} AND p17_111 = !{S} AND p16_18 = !{S} AND p16_131 = !{S} AND p16_125 = !{S} AND p16_115 = !{S} THEN centro_hospitalario = {H1}"	0.776398	0.066667	0.709731	0.939850	0.879699	9.000000
6	"IF p8_5 = !{S} AND p17_115 = !{S} AND p17_107 = !{S} AND p16_112 = !{S} AND p11_3 = !{S} AND p16_99 = !{S} AND p16_93 = !{S} AND p17_103 = !{S} AND p17_38 = !{S} AND p17_106 = !{S} AND p16_114 = !{S} AND p13_9 = !{S} AND p7_15 = !{S} AND p16_121 = !{S} AND p17_126 = !{S} AND p17_132 = !{S} AND p16_86 = !{S} THEN centro_hospitalario = {H1}"	0.832298	0.041667	0.790631	0.964029	0.928058	5.000000
7	"IF p12_8 = {S} AND p17_10 = !{S} AND p12_-1 = !{S} THEN centro_hospitalario = {H11}"	0.421875	0.000000	0.421875	1.000000	1.000000	181.000000
8	"IF p12_7 = {S} AND p3_1 = {S} AND p17_115 = !{S} AND p4_3 = !{S} AND p6_4 = !{S} AND p16_8 = !{S} THEN centro_hospitalario = {H11}"	0.578125	0.000000	0.578125	1.000000	1.000000	174.000000
9	"IF p4_1 = {S} AND p10_4 = {S} AND p9_1 = {S} AND p7_9 = !{S} AND p16_33 = !{S} AND p17_38 = !{S} AND p16_91 = !{S} THEN centro_hospitalario = {H11}"	0.640625	0.000000	0.640625	1.000000	1.000000	174.000000
10	"IF p12_7 = {S} AND p3_2 = !{S} AND p3_4 = !{S} AND p16_89 = !{S} AND p16_121 = !{S} AND p17_123 = !{S} AND p16_131 = !{S} AND p17_131 = !{S} THEN centro_hospitalario = {H11}"	0.625000	0.013825	0.611175	0.930233	0.860465	174.000000

Tabla 4.13. *WRAcc* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
11	"IF p14_8 = !{S} AND p4_3 = !{S} AND p17_125 = !{S} AND p17_10 = !{S} AND p17_13 = !{S} AND p5_4 = !{S} AND p13_10 = !{S} AND p16_9 = !{S} AND p16_11 = !{S} AND p17_2 = !{S} AND p11_6 = !{S} AND p17_3 = !{S} AND p16_27 = !{S} AND p16_16 = !{S} AND p6_5 = !{S} AND p10_5 = !{S} AND p16_17 = !{S} AND p16_47 = !{S} AND p7_12 = !{S} AND p6_7 = !{S} AND p7_13 = !{S} AND p13_8 = !{S} AND p16_35 = !{S} AND p17_15 = !{S} AND p3_5 = !{S} AND p16_12 = !{S} AND p17_43 = !{S} AND p17_31 = !{S} AND p17_22 = !{S} AND p17_17 = !{S} AND p16_1 = !{S} AND p8_7 = !{S} AND p16_41 = !{S} AND p17_42 = !{S} AND p17_32 = !{S} AND p17_51 = !{S} AND p17_8 = !{S} AND p17_82 = !{S} AND p9_-1 = !{S} AND p17_85 = !{S} AND p16_97 = !{S} AND p17_117 = !{S} AND p17_118 = !{S} AND p17_119 = !{S} AND p17_122 = !{S} AND p17_124 = !{S} AND p16_125 = !{S} AND p17_128 = !{S} AND p17_129 = !{S} AND p17_130 = !{S} AND p16_129 = !{S} AND p17_132 = !{S} AND p17_133 = !{S} THEN centro_hospitalario = {H11}"	0.875000	0.059908	0.815092	0.811594	0.623188	170.000000
12	"IF p10_4 = {S} AND p11_4 = {S} AND p16_8 = !{S} AND p16_23 = !{S} AND p16_33 = !{S} AND p4_3 = !{S} AND p16_98 = !{S} AND p7_13 = !{S} AND p16_91 = !{S} AND p16_129 = !{S} THEN centro_hospitalario = {H11}"	0.687500	0.009217	0.678283	0.956522	0.913043	175.000000

Tabla 4.13. *WRAcc* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
13	"IF p9_1 = {S} AND p3_2 = !{S} AND p4_2 = !{S} AND p17_115 = !{S} AND p13_1 = !{S} AND p17_10 = !{S} AND p6_6 = !{S} AND p17_16 = !{S} AND p5_5 = !{S} AND p17_22 = !{S} AND p17_38 = !{S} AND p17_82 = !{S} AND p16_33 = !{S} THEN centro_hospitalario = {H11}"	0.718750	0.000000	0.718750	1.000000	1.000000	166.000000
14	"IF p9_1 = {S} AND p4_2 = !{S} AND p14_8 = !{S} AND p17_115 = !{S} AND p17_10 = !{S} AND p17_125 = !{S} AND p4_4 = !{S} AND p7_15 = !{S} AND p5_4 = !{S} AND p17_22 = !{S} AND p17_82 = !{S} THEN centro_hospitalario = {H11}"	0.703125	0.009217	0.693908	0.957447	0.914894	172.000000
15	"IF p9_6 = {S} AND p3_2 = !{S} AND p6_9 = !{S} AND p4_2 = !{S} AND p13_16 = !{S} AND p16_86 = !{S} AND p16_35 = !{S} AND p9_-1 = !{S} AND p17_130 = !{S} THEN centro_hospitalario = {H11}"	0.484375	0.000000	0.484375	1.000000	1.000000	174.000000
16	"IF p17_115 = {S} AND p9_6 = !{S} THEN centro_hospitalario = {H12}"	0.375000	0.000000	0.375000	1.000000	1.000000	251.000000
17	"IF p16_131 = {S} THEN centro_hospitalario = {H12}"	0.232143	0.000000	0.232143	1.000000	1.000000	271.000000
18	"IF p16_112 = {S} AND p3_2 = {S} THEN centro_hospitalario = {H12}"	0.214286	0.000000	0.214286	1.000000	1.000000	239.000000
19	"IF p16_129 = {S} THEN centro_hospitalario = {H12}"	0.107143	0.000000	0.107143	1.000000	1.000000	

Tabla 4.13. *WRAcc* – Matriz Binaria: Reglas Sin Interpretar – Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
20	"IF p16_8 = !{S} AND p3_1 = !{S} AND p17_10 = !{S} AND p13_1 = !{S} AND p16_13 = !{S} AND p1_5 = !{S} AND p2_5 = !{S} AND p8_-1 = !{S} AND p6_4 = !{S} AND p16_41 = !{S} AND p16_9 = !{S} AND p16_11 = !{S} AND p7_7 = !{S} AND p17_16 = !{S} AND p16_35 = !{S} AND p16_12 = !{S} AND p12_8 = !{S} AND p16_88 = !{S} AND p17_13 = !{S} THEN centro_hospitalario = {H12}"	0.767857	0.000000	0.767857	1.000000	1.000000	229.000000
21	"IF p16_8 = !{S} AND p13_1 = !{S} AND p17_10 = !{S} AND p16_13 = !{S} AND p16_99 = !{S} AND p16_93 = !{S} AND p16_27 = !{S} AND p16_16 = !{S} AND p16_33 = !{S} AND p7_17 = !{S} AND p17_105 = !{S} AND p17_106 = !{S} AND p17_107 = !{S} AND p16_1 = !{S} AND p17_88 = !{S} AND p17_92 = !{S} AND p16_95 = !{S} AND p16_97 = !{S} AND p16_96 = !{S} AND p17_14 = !{S} AND p17_21 = !{S} AND p17_22 = !{S} AND p13_15 = !{S} AND p13_17 = !{S} AND p17_1 = !{S} AND p16_81 = !{S} AND p17_84 = !{S} AND p17_85 = !{S} AND p16_83 = !{S} AND p17_71 = !{S} AND p16_108 = !{S} AND p17_114 = !{S} AND p13_8 = !{S} AND p5_4 = !{S} THEN centro_hospitalario = {H12}"	0.928571	0.026667	0.901905	0.896552	0.793103	229.000000

Tabla 4.13. *WRAcc* – Matriz Binaria: Reglas Sin Interpretar – Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
22	"IF p16_8 = !{S} AND p12_8 = !{S} AND p17_10 = !{S} AND p16_13 = !{S} AND p16_93 = !{S} AND p16_99 = !{S} AND p16_9 = !{S} AND p10_5 = !{S} AND p16_47 = !{S} AND p13_16 = !{S} AND p16_89 = !{S} AND p16_98 = !{S} AND p16_71 = !{S} AND p16_114 = !{S} AND p7_7 = !{S} AND p17_106 = !{S} AND p17_3 = !{S} AND p16_17 = !{S} AND p7_12 = !{S} AND p7_14 = !{S} AND p16_26 = !{S} AND p17_31 = !{S} AND p17_49 = !{S} AND p17_17 = !{S} AND p7_22 = !{S} AND p16_25 = !{S} AND p17_82 = !{S} AND p17_109 = !{S} AND p16_113 = !{S} AND p16_27 = !{S} AND p16_35 = !{S} AND p16_41 = !{S} AND p17_13 = !{S} THEN centro_hospitalario = {H12}"	1.000000	0.084444	0.915556	0.746667	0.493333	229.000000

4.8.2 Reglas interpretadas

La Tabla 4.14, correspondiente a las reglas interpretadas generadas mediante *WRAcc*, muestra la traducción semántica de los veintidós patrones obtenidos, permitiendo comprender los perfiles contrastivos desde una perspectiva clínica e institucional. En conjunto, las reglas interpretadas integran atributos relacionados con formación académica, métodos percibidos como eficientes para solicitar autopsia, posturas sobre discrepancias diagnósticas, y factores de rechazo o barreras vinculadas al proceso.

Las reglas asociadas al Hospital Regional Río Blanco describen un perfil dominado por combinaciones que incluyen elementos de trayectoria formativa y condiciones relacionadas con la percepción de suficiencia o insuficiencia de autopsias por motivos organizacionales. También aparecen componentes vinculados con el método eficiente para solicitar autopsia,

así como elementos que involucran el rol de actores específicos dentro del proceso. En este bloque, el contenido semántico tiende a ser consistente y relativamente compacto, lo cual es congruente con los valores bajos de redundancia observados en el modelo.

En el Hospital General de San Juan Bautista Tuxtepec, las reglas interpretadas enfatizan con mayor frecuencia elementos procedimentales, destacando la postura de los médicos respecto al método eficiente para solicitar autopsia (por ejemplo, a través de áreas específicas), así como afirmaciones de alta intensidad relacionadas con arbitraje, demanda y discrepancias entre hallazgos de autopsia y diagnósticos clínicos. Este patrón interpretativo sugiere un perfil institucional altamente estructurado, donde las percepciones sobre implicaciones médico-legales y la ruta procedimental de solicitud aparecen de manera recurrente, lo cual también es consistente con la redundancia elevada del bloque.

Para el Hospital General de Zona 53, las reglas interpretadas muestran una combinación de elementos de formación (incluyendo instituciones específicas) y posturas respecto a discrepancias diagnósticas, así como condiciones asociadas a percepciones del proceso y su aceptación. En este bloque se identifican reglas con alta representatividad, pero también reglas más específicas, reflejando la variabilidad observada en el soporte. La presencia de redundancia elevada sugiere que varias reglas describen el mismo perfil institucional desde combinaciones alternativas de atributos.

En conjunto, la Tabla 4.14 evidencia que *WRAcc* genera patrones semánticamente interpretables y coherentes con diferencias institucionales, permitiendo vincular los resultados cuantitativos con descripciones comprensibles del fenómeno analizado. Este paso habilita el análisis visual y la evaluación integral del modelo desde una perspectiva objetiva y subjetiva.

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad del Valle de México, no realizaron su especialidad en la UNAM, no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo, no han participado entre 11 y 20 casos de autopsias, no estudiaron su especialidad en Hospital Ángeles, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no realizaron su formación de medicina general en la Escuela Superior de Medicina, no estudiaron medicina general en Universidad Juárez Autónoma, no estudiaron medicina general en Hospital Juárez de México, y no estudiaron medicina general en La Raza IMSS; entonces son del Hospital Regional Río Blanco.
2	Si los médicos no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estudiaron su especialidad en IMSS, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital general, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de trabajo social, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos financieros, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, sí especifican dónde realizaron su especialidad, no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron medicina general en Instituto Politécnico Nacional, no estudiaron su especialidad en Universidad de Nuevo León, no estudiaron su especialidad en Escuela Militar de Medicina, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron su especialidad en Estudios de Posgrado de Excelencia, no estudiaron su especialidad en Hospital General La Villa Urgencias, no estudiaron su especialidad en Centro Médico Nacional La Raza, no estudiaron su especialidad en Hospital Juárez de México, no estudiaron su especialidad en Universidad Autónoma de México, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, y no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza; entonces son del Hospital Regional Río Blanco.

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
3	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, no realizaron su especialidad en la UNAM, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en Estudios de Posgrado de Excelencia, y no opinan que se debe obligar a los residentes; entonces son del Hospital Regional Río Blanco.
4	Si los médicos no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no estudiaron medicina general en Universidad del Valle de México, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Hospital Juárez de México, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron su especialidad en Universidad de Guanajuato, no estudiaron su especialidad en Hospital General Balbuena, y no estudiaron medicina general en Escuela Superior de Medicina; entonces son del Hospital Regional Río Blanco.
5	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no tienen 11-15 años de práctica, no estudiaron su especialidad en Hospital General Palenque, no piensan que una causa o motivo para no solicitar autopsia es que no exista el servicio, no consideran que quien debe solicitar la autopsia es una enfermera, no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron medicina general en Universidad Juárez Autónoma, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, no estudiaron medicina general en Universidad Autónoma de México, y no estudiaron medicina general en Hospital Hermanos Ameijeiras; entonces son del Hospital Regional Río Blanco.

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
6	Si los médicos no creen que la causa de rechazo por parte del familiar es por inconformidad de los familiares con la atención prestada al paciente, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad del Valle de México, no estiman como motivo para solicitar autopsia el consumo de medicamentos mal recetados, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Anáhuac, no estudiaron su especialidad en IMSS, no realizaron su especialidad en la UNAM, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron medicina general en Estudios de Posgrado de Excelencia, no opinan que el estudio aclara las causas de muerte, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos materiales, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no estudiaron su especialidad en Centro Médico Nacional La Raza, y no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo; entonces son del Hospital Regional Río Blanco.
7	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología y no estudiaron su especialidad en Facultad de Medicina; entonces son del Hospital General de San Juan Bautista Tuxtepec.
8	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, y no estudiaron medicina general en Alta Especialidad en IMSS; entonces son del Hospital General de San Juan Bautista Tuxtepec.
9	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, consideran que quien debe solicitar la autopsia es una institución, consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no piensan que una causa o motivo para no solicitar autopsia es por cuestiones religiosas, no estudiaron medicina general en Regional de Occidente, no realizaron su especialidad en la UNAM, y no estudiaron medicina general en la Universidad Xochicolco; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
10	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Regional del Sureste, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, sí especifican dónde realizaron su especialidad, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, y no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de San Juan Bautista Tuxtepec.
11	Si los médicos no son estudiantes de medicina, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron su especialidad en Universidad Autónoma de México, no estudiaron su especialidad en Facultad de Medicina, no estudiaron su especialidad en Universidad de San Carlos, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no opinan que para este estudio deberían participar todos los servicios, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron medicina general en Instituto Nacional de Cancerología, no estudiaron su especialidad en Nachón/CEM, no estiman como motivo para solicitar autopsia la orientación diagnóstica, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron medicina general en Universidad Popular Autónoma, no plantean como motivo para solicitar autopsia cuando existe falta de correlación entre los datos clínicos y los de laboratorio, no consideran que quien debe solicitar la autopsia es un familiar, no estudiaron medicina general en Universidad Benito Juárez, no estudiaron medicina general en Hospital General Regional IMSS, no piensan que una causa o motivo para no solicitar autopsia es cuando el cuerpo se encuentra en descomposición, no plantean como motivo para solicitar autopsia cuando es necesario determinar si fue correcta o no la aplicación del tratamiento empleado, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron su especialidad en Instituto Nacional de Pediatría, no están totalmente en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Villa Rica, no estudiaron su especialidad en Instituto de Salud Pública, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Secretaría de Salud, no estudiaron su especialidad en Universidad Benito Juárez, no estudiaron medicina general en Universidad Veracruzana, no creen que la causa de rechazo por parte del familiar es circunstancial, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron su

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	especialidad en Centro Médico de Occidente, no estudiaron su especialidad en Hospital de la Mujer, no estudiaron su especialidad en Hospital Clínico Quirúrgico Hermanos Ameijeiras, no estudiaron su especialidad en Alta Especialidad en IMSS, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital General Palenque, no estudiaron su especialidad en Hospital General La Villa Urgencias, no estudiaron su especialidad en Hospital general, no estudiaron medicina general en Universidad Autónoma de México, no estudiaron su especialidad en Centro Médico Nacional La Raza, no estudiaron su especialidad en Hospital Juárez de México, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, y no estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de San Juan Bautista Tuxtepec.
12	Si los médicos consideran que quien debe solicitar la autopsia es una institución, estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron medicina general en Alta Especialidad en IMSS, no estudiaron medicina general en La Raza IMSS, no estudiaron medicina general en Regional de Occidente, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en una Universidad Autónoma, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no estudiaron medicina general en la Universidad Xochicolco, y no estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de San Juan Bautista Tuxtepec.
13	Si los médicos consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí comentan, no estudiaron su especialidad en Facultad de Medicina, no plantean como motivo para solicitar autopsia cuando es necesario proporcionar información esencial a familiares en casos de enfermedades infecciosas o congénitas, no estudiaron su especialidad en Universidad Popular Autónoma, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estudiaron su especialidad en Secretaría de Salud, no realizaron su especialidad en la UNAM, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, y no estudiaron medicina general en Regional de Occidente; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
14	Si los médicos consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no son estudiantes de medicina, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estudiaron su especialidad en Facultad de Medicina, no estudiaron su especialidad en Universidad Autónoma de México, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos materiales, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estudiaron su especialidad en Secretaría de Salud, y no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla; entonces son del Hospital General de San Juan Bautista Tuxtepec.
15	Si los médicos consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no plantean como motivo para solicitar autopsia el diagnóstico difícil, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no opinan que nunca le han realizado autopsias a sus pacientes, no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo, no estudiaron medicina general en San Nicolás de Hidalgo, y no estudiaron su especialidad en Hospital General Dr. Miguel Silva; entonces son del Hospital General de San Juan Bautista Tuxtepec.
16	Si los médicos estudiaron su especialidad en Hospital Hermanos Ameijeiras y no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica; entonces son del Hospital General de Zona 53.
17	Si los médicos realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de Zona 53.
18	Si los médicos estudiaron medicina general en Universidad del Valle de México y están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de Zona 53.
19	Si los médicos estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de Zona 53.
20	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron su especialidad en Facultad de Medicina, sí comentan, no estudiaron medicina general en Universidad de San Carlos, no tienen más de 20 años de práctica, no han participado en más de 20 casos de autopsias, no plantean como motivo para solicitar autopsia la

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	aplicabilidad del estudio, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron medicina general en Instituto Nacional de Cancerología, no piensan que una causa o motivo para no solicitar autopsia es por los trámites administrativos, no estudiaron su especialidad en Universidad Popular Autónoma, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron medicina general en Universidad Villa Rica, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron medicina general en Hospital General Dr. Manuel Gea González, y no estudiaron su especialidad en Universidad de San Carlos; entonces son del Hospital General de Zona 53.
21	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, sí comentan, no estudiaron su especialidad en Facultad de Medicina, no estudiaron medicina general en Universidad de San Carlos, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron medicina general en Universidad Popular Autónoma, no estudiaron medicina general en Regional de Occidente, no piensan que una causa o motivo para no solicitar autopsia es que se haga sin fines de enseñanza, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron su especialidad en Hospital General Balbuena, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Hospital General Dr. Manuel Gea González, no estudiaron su especialidad en Universidad Autónoma de Guadalajara, no estudiaron medicina general en Universidad Autónoma Benito Juárez de Oaxaca, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron medicina general en Hospital Regional De Alta Especialidad de Veracruz, no estudiaron su especialidad en Universidad La Salle, no estudiaron su especialidad en IMSS Adolfo Ruiz Cortines, no estudiaron su especialidad en Secretaría de Salud, no opinan que se debe obligar a los residentes, no opinan que se debe solicitar el consentimiento, no estudiaron su especialidad en Universidad Veracruzana, no estudiaron medicina general en Hospital General de Orizaba, no estudiaron su especialidad en Universidad de Guadalajara, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no estudiaron medicina general en la Universidad Popular Autónoma de Puebla, no estudiaron su especialidad en Hospital general Miguel Enríquez, no estudiaron medicina general en Instituto Politécnico Nacional, no estudiaron su especialidad en Estudios de Posgrado de Excelencia, no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general, y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.

Tabla 4.14. *WRAcc* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
22	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron su especialidad en Facultad de Medicina, no estudiaron medicina general en Universidad de San Carlos, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Hospital del Departamento del D.F., no consideran que quien debe solicitar la autopsia es un familiar, no estudiaron medicina general en Hospital General Regional IMSS, no opinan que nunca le han realizado autopsias a sus pacientes, no estudiaron medicina general en Universidad Regional del Sureste, no estudiaron medicina general en una Universidad Autónoma, no estudiaron medicina general en la Facultad de Ciencias Médicas, no estudiaron medicina general en Estudios de Posgrado de Excelencia, no piensan que una causa o motivo para no solicitar autopsia es por los trámites administrativos, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron medicina general en Universidad Benito Juárez, no piensan que una causa o motivo para no solicitar autopsia es cuando el cuerpo se encuentra en descomposición, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos humanos, no estudiaron medicina general en Hospital General, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Universidad de Colima, no estudiaron su especialidad en Universidad Benito Juárez, no piensan que una causa o motivo para no solicitar autopsia es el olvido en la práctica de este estudio, no estudiaron medicina general en Ignacio Chávez, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Universidad de Nuevo León, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron medicina general en Universidad Cristóbal Colón, y no estudiaron su especialidad en Universidad de San Carlos; entonces son del Hospital General de Zona 53.

4.8.3 Análisis visual

El análisis visual permite observar el comportamiento global de las métricas asociadas a las reglas generadas mediante *WRAcc*, facilitando la identificación de rangos, concentraciones y contrastes entre patrones. A continuación, se presentan las gráficas correspondientes a soporte, soporte negativo, diferencia de soporte, precisión, diferencia de precisión y redundancia.

La Figura 4.52 muestra la distribución del soporte para las veintidós reglas generadas. Se observa un rango amplio, desde valores bajos cercanos a 0.11 hasta valores máximos de 1.00, lo que indica la coexistencia de reglas altamente representativas con otras más específicas. La concentración principal se ubica en valores moderados y altos, reflejando que *WRAcc* favorece patrones con cobertura relevante dentro del hospital objetivo.

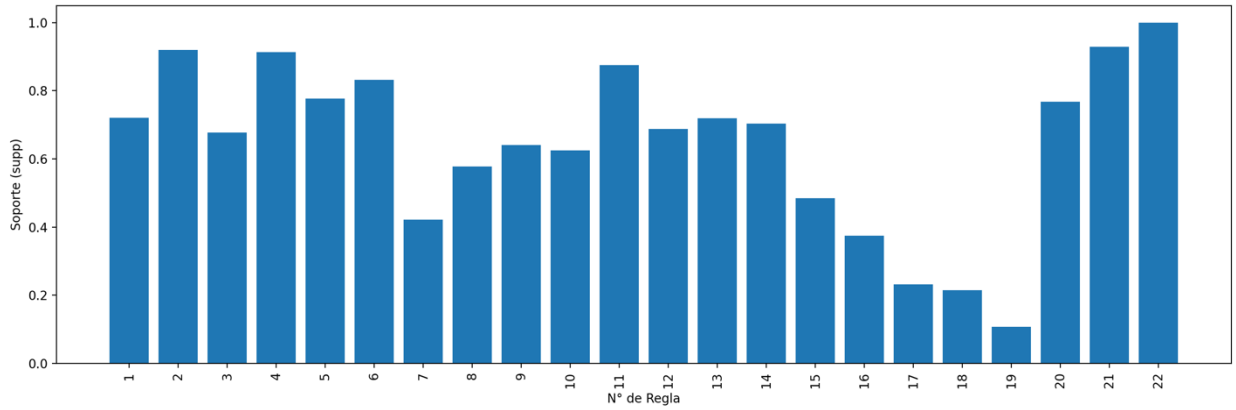


Figura 4.52 *WRAcc* – Matriz Binaria: Soporte por Regla

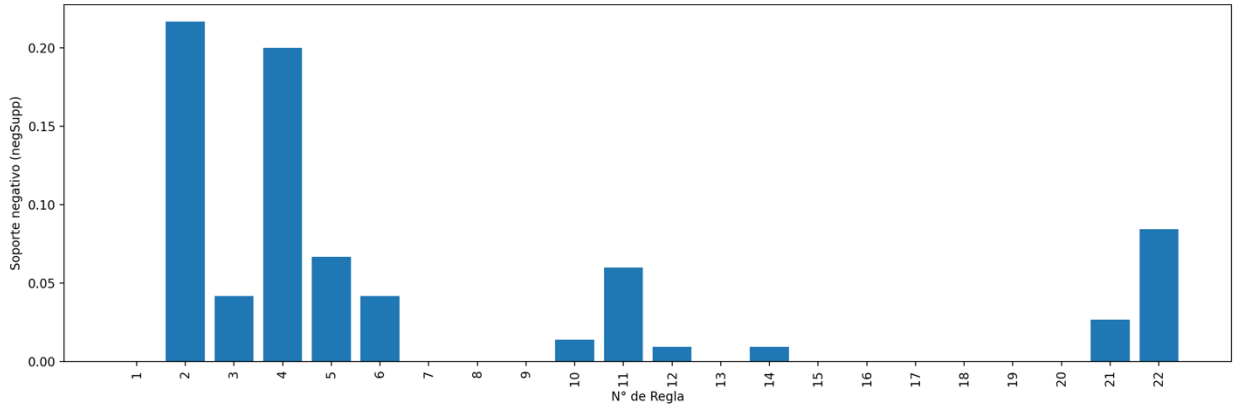


Figura 4.53 *WRAcc* – Matriz Binaria: Soporte Negativo por Regla

La Figura 4.53 presenta el comportamiento del soporte negativo, donde se observa que una parte del conjunto mantiene valores nulos, mientras que otras reglas exhiben valores positivos que alcanzan aproximadamente 0.22. Este comportamiento confirma que el modelo incluye tanto patrones completamente exclusivos del hospital objetivo como

reglas donde existen contraejemplos en otros hospitales, lo que incide directamente en la reducción de la precisión en ciertas reglas.

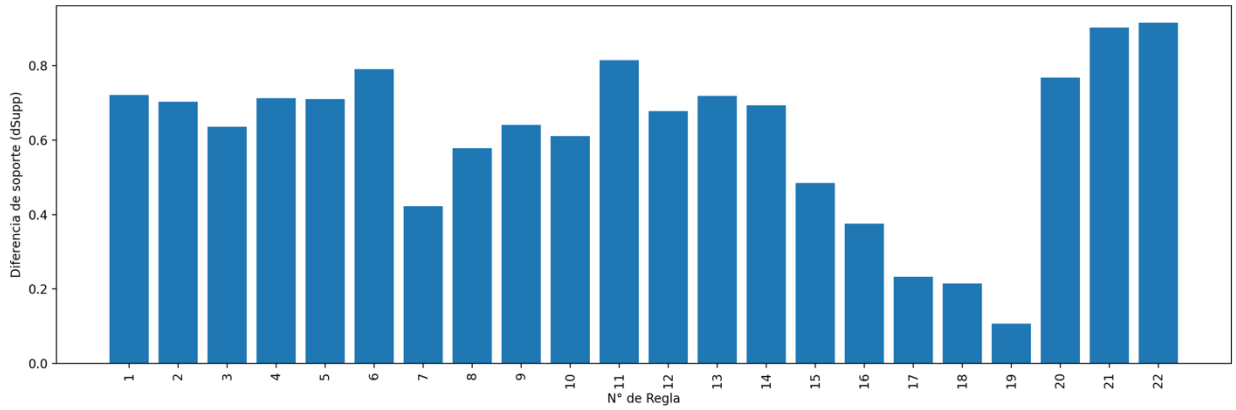


Figura 4.54 *WRAcc* – Matriz Binaria: Diferencia de Soporte por Regla

La Figura 4.54 muestra la diferencia de soporte, con valores consistentemente positivos que oscilan aproximadamente entre 0.11 y 0.92. Esto indica que todos los patrones aportan discriminación real a favor del hospital objetivo. Se observan reglas con diferencias de soporte muy elevadas, lo que refleja perfiles fuertemente contrastivos dentro del conjunto binario.

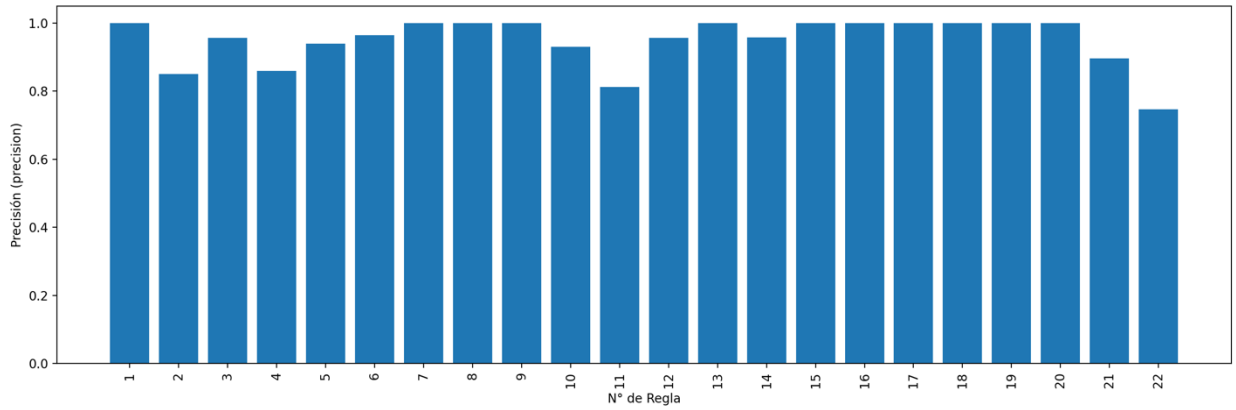


Figura 4.55 *WRAcc* – Matriz Binaria: Precisión por Regla

La Figura 4.55 evidencia que la precisión se mantiene en rangos altos, desde valores cercanos a 0.75 hasta 1.00. Se aprecia un subconjunto importante de reglas con

precisión perfecta, mientras que otras presentan descensos moderados asociados al incremento del soporte negativo. En conjunto, la gráfica confirma un modelo altamente discriminativo, pero con variaciones de pureza entre reglas.

La Figura 4.56 presenta la diferencia de precisión, con un rango que va aproximadamente de 0.49 a 1.00. La presencia de valores máximos en varias reglas confirma patrones con ganancia total respecto al comportamiento base, mientras que los valores menores reflejan reglas con desempeño relativo todavía positivo, pero afectado por la existencia de contraejemplos.

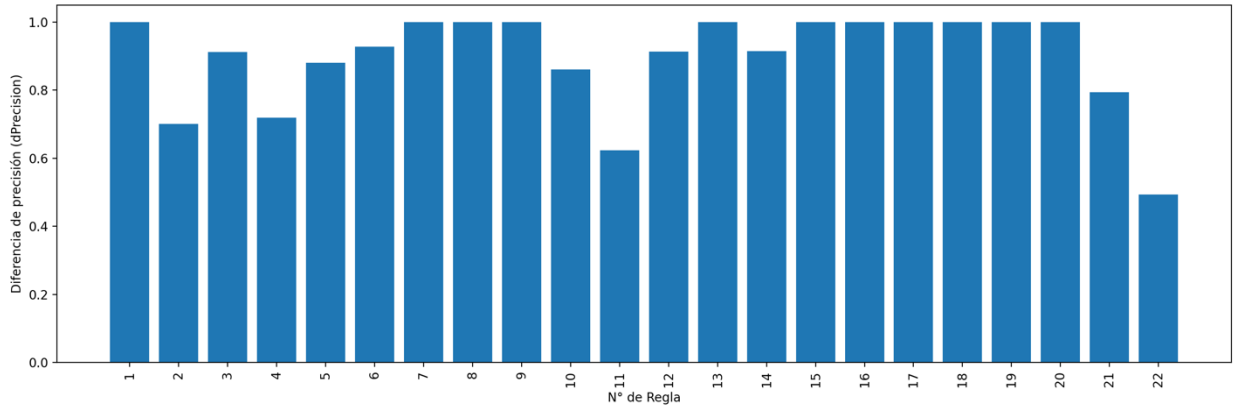


Figura 4.56 *WRAcc* – Matriz Binaria: Diferencia de Precisión por Regla

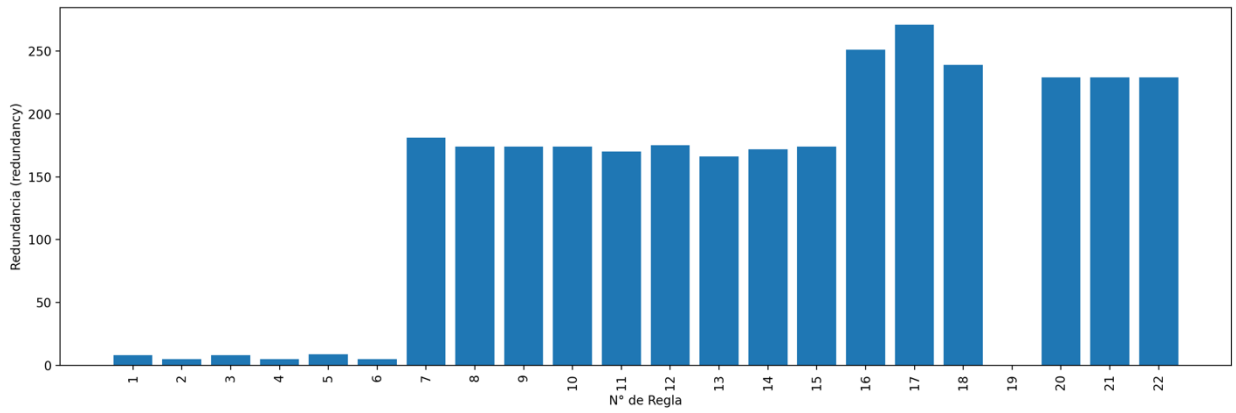


Figura 4.57 *WRAcc* – Matriz Binaria: Redundancia por Regla

4.8.4 Evaluación objetiva

La evaluación objetiva del modelo *WRAcc* aplicado a la Matriz Binaria muestra un comportamiento cuantitativamente consistente, caracterizado por diferencias de soporte positivas en todas las reglas, lo que confirma que cada patrón incrementa su presencia relativa en el hospital objetivo frente al resto de la población. El soporte se concentra mayormente en valores moderados y altos, indicando que *WRAcc* privilegia reglas con cobertura relevante y no únicamente patrones raros.

A diferencia de escenarios donde el soporte negativo es nulo para todo el conjunto, en este modelo se observa una mezcla de reglas con soporte negativo cero y reglas con soporte negativo positivo, alcanzando valores cercanos a 0.22. Esto implica que existen tanto patrones completamente exclusivos como reglas con contraejemplos, lo cual se refleja en precisiones que, aunque altas en general, pueden descender hasta rangos cercanos a 0.75. En consecuencia, el modelo mantiene un desempeño discriminativo fuerte, pero no uniformemente perfecto en todas las reglas.

La redundancia introduce el elemento más crítico de la evaluación objetiva. Se observa un contraste claro entre reglas con redundancia baja y un bloque con redundancia elevada, especialmente asociado a hospitales donde múltiples combinaciones generan patrones equivalentes. Esto indica que el modelo contiene conocimiento repetido a nivel estructural y que, para fines de síntesis, resulta razonable priorizar reglas con alta diferencia de soporte y menor redundancia, preservando la diversidad informativa del conjunto.

4.8.5 Evaluación subjetiva

La evaluación subjetiva del conjunto de reglas generado mediante *WRAcc* mostró un alto nivel de aceptación clínica. La mayoría de los patrones fue validada por el Dr. Palet al describir perfiles médicos coherentes con la práctica institucional observada, especialmente en el Hospital General de San Juan Bautista Tuxtepec y el Hospital Regional Río Blanco.

Las reglas descartadas corresponden a casos puntuales donde, pese a cumplir con los criterios estadísticos de la métrica, los patrones no aportan información clínica diferenciadora desde la experiencia del especialista. En general, *WRAcc* se percibe como una función de calidad que logra un equilibrio adecuado entre cobertura y capacidad discriminativa, produciendo reglas con alto valor interpretativo y una necesidad limitada de depuración posterior.

4.9 Matriz Binaria: Resultados aplicando la función de calidad *Correlation*

En este apartado se presentan los resultados del análisis de Minería de Conjuntos de Contraste aplicado al conjunto de datos Matriz Binaria, en nivel específico, utilizando la función de calidad *Correlation*. Esta métrica permite identificar patrones contrastivos al medir la asociación entre la ocurrencia del antecedente y la pertenencia al hospital objetivo, favoreciendo reglas que reflejen una relación estadística consistente. Bajo este enfoque, los resultados se reportan mediante la presentación de reglas sin interpretar, su interpretación semántica, el análisis visual de métricas y la evaluación objetiva y subjetiva del modelo, manteniendo el mismo estándar metodológico empleado en los apartados anteriores.

4.9.1 Reglas sin interpretar

La Tabla 4.15, correspondiente a las reglas sin interpretar generadas mediante *Correlation*, muestra un conjunto de veintidós patrones distribuidos entre los tres hospitales analizados. En términos generales, el modelo presenta reglas con soportes predominantemente moderados y altos, diferencias de soporte positivas y precisiones elevadas, lo que evidencia un comportamiento discriminativo sólido y estable.

Las reglas asociadas al Hospital Regional Río Blanco conforman un primer bloque integrado por seis patrones, con valores de soporte que oscilan aproximadamente entre 0.68 y 0.92. En este conjunto, la diferencia de soporte se mantiene elevada, alcanzando valores cercanos a 0.79, lo que confirma patrones altamente característicos del hospital. La precisión se conserva en rangos altos, desde aproximadamente 0.85 hasta 1.0, y la

diferencia de precisión alcanza valores de hasta 1.0. Se observa, sin embargo, que algunas reglas presentan soporte negativo positivo (hasta alrededor de 0.22), lo que explica la existencia de precisiones ligeramente menores en ciertos patrones. La redundancia en este bloque se mantiene baja (entre 5 y 9), indicando un conjunto compacto con menor solapamiento informativo.

El conjunto de reglas vinculado al Hospital General de San Juan Bautista Tuxtepec constituye el bloque más amplio, con nueve patrones, cuyos soportes se ubican aproximadamente entre 0.42 y 0.88. En este caso, el soporte negativo se mantiene generalmente bajo (hasta alrededor de 0.06), permitiendo que la precisión se conserve en valores altos, aunque con descensos moderados en reglas específicas (hasta alrededor de 0.81). Las diferencias de soporte alcanzan valores elevados cercanos a 0.82, reflejando una estructura de patrones contrastivos robusta. Un rasgo distintivo del bloque es la redundancia persistentemente alta (aproximadamente entre 166 y 181), lo que sugiere la existencia de múltiples reglas con estructuras similares que describen regiones equivalentes del espacio de búsqueda.

Finalmente, las reglas asociadas al Hospital General de Zona 53 conforman un bloque compuesto por siete patrones, donde se observa la mayor variabilidad en soporte, desde valores cercanos a 0.11 hasta valores máximos de 1.00. En este conjunto, el soporte negativo puede alcanzar valores cercanos a 0.08, lo que implica que no todas las reglas son exclusivas del hospital y que la precisión puede descender a valores aproximados de 0.75. A pesar de ello, las diferencias de soporte se mantienen positivas y pueden alcanzar valores cercanos a 0.92, confirmando un comportamiento contrastivo fuerte. La redundancia en este bloque también es elevada (aproximadamente entre 229 y 271), evidenciando un solapamiento estructural importante entre reglas.

Tabla 4.15. *Correlation* – Matriz Binaria: Reglas Sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF p9_1 = !{S} AND p17_115 = !{S} AND p17_107 = !{S} AND p16_112 = !{S} AND p17_38 = !{S} AND p16_86 = !{S} AND p2_4 = !{S} AND p12_-1 = !{S} AND p17_102 = !{S} AND p16_121 = !{S} AND p17_123 = !{S} AND p16_124 = !{S} AND p16_18 = !{S} AND p16_129 = !{S} AND p16_23 = !{S} THEN centro_hospitalario = {H1}"	0.720497	0.000000	0.720497	1.000000	1.000000	8.000000
2	"IF p17_115 = !{S} AND p17_103 = !{S} AND p17_117 = !{S} AND p17_122 = !{S} AND p17_126 = !{S} AND p12_6 = !{S} AND p7_16 = !{S} AND p12_8 = !{S} AND p17_100 = !{S} AND p17_101 = !{S} AND p17_105 = !{S} AND p16_108 = !{S} AND p17_109 = !{S} AND p17_110 = !{S} AND p17_111 = !{S} AND p17_114 = !{S} AND p17_119 = !{S} AND p17_124 = !{S} AND p17_128 = !{S} AND p17_129 = !{S} AND p17_130 = !{S} AND p17_125 = !{S} AND p16_131 = !{S} AND p17_131 = !{S} AND p17_133 = !{S} THEN centro_hospitalario = {H1}"	0.919255	0.216667	0.702588	0.850575	0.701149	5.000000
3	"IF p9_1 = !{S} AND p17_115 = !{S} AND p9_6 = !{S} AND p17_38 = !{S} AND p16_112 = !{S} AND p16_114 = !{S} AND p13_15 = !{S} THEN centro_hospitalario = {H1}"	0.677019	0.041667	0.635352	0.956140	0.912281	8.000000
4	"IF p17_115 = !{S} AND p12_7 = !{S} AND p16_112 = !{S} AND p17_107 = !{S} AND p16_129 = !{S} AND p16_99 = !{S} AND p12_-1 = !{S} AND p17_104 = !{S} AND p17_106 = !{S} AND p16_120 = !{S} THEN centro_hospitalario = {H1}"	0.913043	0.200000	0.713043	0.859649	0.719298	5.000000

Tabla 4.15. *Correlation* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
5	"IF p11_4 = !{S} AND p16_112 = !{S} AND p16_99 = !{S} AND p1_3 = !{S} AND p17_118 = !{S} AND p7_1 = !{S} AND p10_2 = !{S} AND p17_101 = !{S} AND p17_111 = !{S} AND p16_18 = !{S} AND p16_131 = !{S} AND p16_125 = !{S} AND p16_115 = !{S} THEN centro_hospitalario = {H1}"	0.776398	0.066667	0.709731	0.939850	0.879699	9.000000
6	"IF p8_5 = !{S} AND p17_115 = !{S} AND p17_107 = !{S} AND p16_112 = !{S} AND p11_3 = !{S} AND p16_99 = !{S} AND p16_93 = !{S} AND p17_103 = !{S} AND p17_38 = !{S} AND p17_106 = !{S} AND p16_114 = !{S} AND p13_9 = !{S} AND p7_15 = !{S} AND p16_121 = !{S} AND p17_126 = !{S} AND p17_132 = !{S} AND p16_86 = !{S} THEN centro_hospitalario = {H1}"	0.832298	0.041667	0.790631	0.964029	0.928058	5.000000
7	"IF p12_8 = {S} AND p17_10 = !{S} AND p12_-1 = !{S} THEN centro_hospitalario = {H11}"	0.421875	0.000000	0.421875	1.000000	1.000000	181.000000
8	"IF p12_7 = {S} AND p3_1 = {S} AND p17_115 = !{S} AND p4_3 = !{S} AND p6_4 = !{S} AND p16_8 = !{S} THEN centro_hospitalario = {H11}"	0.578125	0.000000	0.578125	1.000000	1.000000	174.000000
9	"IF p4_1 = {S} AND p10_4 = {S} AND p9_1 = {S} AND p7_9 = !{S} AND p16_33 = !{S} AND p17_38 = !{S} AND p16_91 = !{S} THEN centro_hospitalario = {H11}"	0.640625	0.000000	0.640625	1.000000	1.000000	174.000000
10	"IF p12_7 = {S} AND p3_2 = !{S} AND p3_4 = !{S} AND p16_89 = !{S} AND p16_121 = !{S} AND p17_123 = !{S} AND p16_131 = !{S} AND p17_131 = !{S} THEN centro_hospitalario = {H11}"	0.625000	0.013825	0.611175	0.930233	0.860465	174.000000

Tabla 4.15. *Correlation* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
11	"IF p14_8 = !{S} AND p4_3 = !{S} AND p17_125 = !{S} AND p17_10 = !{S} AND p17_13 = !{S} AND p5_4 = !{S} AND p13_10 = !{S} AND p16_9 = !{S} AND p16_11 = !{S} AND p17_2 = !{S} AND p11_6 = !{S} AND p17_3 = !{S} AND p16_27 = !{S} AND p16_16 = !{S} AND p6_5 = !{S} AND p10_5 = !{S} AND p16_17 = !{S} AND p16_47 = !{S} AND p7_12 = !{S} AND p6_7 = !{S} AND p7_13 = !{S} AND p13_8 = !{S} AND p16_35 = !{S} AND p17_15 = !{S} AND p3_5 = !{S} AND p16_12 = !{S} AND p17_43 = !{S} AND p17_31 = !{S} AND p17_22 = !{S} AND p17_17 = !{S} AND p16_1 = !{S} AND p8_7 = !{S} AND p16_41 = !{S} AND p17_42 = !{S} AND p17_32 = !{S} AND p17_51 = !{S} AND p17_8 = !{S} AND p17_82 = !{S} AND p9_-1 = !{S} AND p17_85 = !{S} AND p16_97 = !{S} AND p17_117 = !{S} AND p17_118 = !{S} AND p17_119 = !{S} AND p17_122 = !{S} AND p17_124 = !{S} AND p16_125 = !{S} AND p17_128 = !{S} AND p17_129 = !{S} AND p17_130 = !{S} AND p16_129 = !{S} AND p17_132 = !{S} AND p17_133 = !{S} THEN centro_hospitalario = {H11}"	0.875000	0.059908	0.815092	0.811594	0.623188	170.000000
12	"IF p10_4 = {S} AND p11_4 = {S} AND p16_8 = !{S} AND p16_23 = !{S} AND p16_33 = !{S} AND p4_3 = !{S} AND p16_98 = !{S} AND p7_13 = !{S} AND p16_91 = !{S} AND p16_129 = !{S} THEN centro_hospitalario = {H11}"	0.687500	0.009217	0.678283	0.956522	0.913043	175.000000

Tabla 4.15. *Correlation* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
13	"IF p9_1 = {S} AND p3_2 = !{S} AND p4_2 = !{S} AND p17_115 = !{S} AND p13_1 = !{S} AND p17_10 = !{S} AND p6_6 = !{S} AND p17_16 = !{S} AND p5_5 = !{S} AND p17_22 = !{S} AND p17_38 = !{S} AND p17_82 = !{S} AND p16_33 = !{S} THEN centro_hospitalario = {H11}"	0.718750	0.000000	0.718750	1.000000	1.000000	166.000000
14	"IF p9_1 = {S} AND p4_2 = !{S} AND p14_8 = !{S} AND p17_115 = !{S} AND p17_10 = !{S} AND p17_125 = !{S} AND p4_4 = !{S} AND p7_15 = !{S} AND p5_4 = !{S} AND p17_22 = !{S} AND p17_82 = !{S} THEN centro_hospitalario = {H11}"	0.703125	0.009217	0.693908	0.957447	0.914894	172.000000
15	"IF p9_6 = {S} AND p3_2 = !{S} AND p6_9 = !{S} AND p4_2 = !{S} AND p13_16 = !{S} AND p16_86 = !{S} AND p16_35 = !{S} AND p9_-1 = !{S} AND p17_130 = !{S} THEN centro_hospitalario = {H11}"	0.484375	0.000000	0.484375	1.000000	1.000000	174.000000
16	"IF p17_115 = {S} AND p9_6 = !{S} THEN centro_hospitalario = {H12}"	0.375000	0.000000	0.375000	1.000000	1.000000	251.000000
17	"IF p16_131 = {S} THEN centro_hospitalario = {H12}"	0.232143	0.000000	0.232143	1.000000	1.000000	271.000000
18	"IF p16_112 = {S} AND p3_2 = {S} THEN centro_hospitalario = {H12}"	0.214286	0.000000	0.214286	1.000000	1.000000	239.000000
19	"IF p16_129 = {S} THEN centro_hospitalario = {H12}"	0.107143	0.000000	0.107143	1.000000	1.000000	

Tabla 4.15. *Correlation* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
20	"IF p16_8 = !{S} AND p3_1 = !{S} AND p17_10 = !{S} AND p13_1 = !{S} AND p16_13 = !{S} AND p1_5 = !{S} AND p2_5 = !{S} AND p8_-1 = !{S} AND p6_4 = !{S} AND p16_41 = !{S} AND p16_9 = !{S} AND p16_11 = !{S} AND p7_7 = !{S} AND p17_16 = !{S} AND p16_35 = !{S} AND p16_12 = !{S} AND p12_8 = !{S} AND p16_88 = !{S} AND p17_13 = !{S} THEN centro_hospitalario = {H12}"	0.767857	0.000000	0.767857	1.000000	1.000000	229.000000
21	"IF p16_8 = !{S} AND p13_1 = !{S} AND p17_10 = !{S} AND p16_13 = !{S} AND p16_99 = !{S} AND p16_93 = !{S} AND p16_27 = !{S} AND p16_16 = !{S} AND p16_33 = !{S} AND p7_17 = !{S} AND p17_105 = !{S} AND p17_106 = !{S} AND p17_107 = !{S} AND p16_1 = !{S} AND p17_88 = !{S} AND p17_92 = !{S} AND p16_95 = !{S} AND p16_97 = !{S} AND p16_96 = !{S} AND p17_14 = !{S} AND p17_21 = !{S} AND p17_22 = !{S} AND p13_15 = !{S} AND p13_17 = !{S} AND p17_1 = !{S} AND p16_81 = !{S} AND p17_84 = !{S} AND p17_85 = !{S} AND p16_83 = !{S} AND p17_71 = !{S} AND p16_108 = !{S} AND p17_114 = !{S} AND p13_8 = !{S} AND p5_4 = !{S} THEN centro_hospitalario = {H12}"	0.928571	0.026667	0.901905	0.896552	0.793103	229.000000
22	"IF p16_8 = !{S} AND p12_8 = !{S} AND p17_10 = !{S} AND p16_13 = !{S} AND p16_93 = !{S} AND p16_99 = !{S} AND p16_9 = !{S} AND p10_5 = !{S} AND p16_47 = !{S} AND p13_16 = !{S} AND p16_89 = !{S} AND p16_98 = !{S} AND p16_71 = !{S} AND p16_114 = !{S} AND p7_7 = !{S} AND	1.000000	0.084444	0.915556	0.746667	0.493333	229.000000

Tabla 4.15. *Correlation* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
	p17_106 = !{S} AND p17_3 = !{S} AND p16_17 = !{S} AND p7_12 = !{S} AND p7_14 = !{S} AND p16_26 = !{S} AND p17_31 = !{S} AND p17_49 = !{S} AND p17_17 = !{S} AND p7_22 = !{S} AND p16_25 = !{S} AND p17_82 = !{S} AND p17_109 = !{S} AND p16_113 = !{S} AND p16_27 = !{S} AND p16_35 = !{S} AND p16_41 = !{S} AND p17_13 = !{S} THEN centro_hospitalario = {H12}"						

En conjunto, la Tabla 4.15 evidencia que *Correlation* sobre Matriz Binaria produce un conjunto de reglas con alta capacidad discriminativa y cobertura relevante, aunque con distintos niveles de exclusividad y con redundancias significativamente altas en H11 y H12. Esta tabla constituye el insumo base para su posterior interpretación semántica.

4.9.2 Reglas interpretadas

La Tabla 4.16, correspondiente a las reglas interpretadas generadas mediante *Correlation*, presenta la traducción semántica de los veintidós patrones obtenidos, permitiendo comprender los perfiles contrastivos desde una perspectiva clínica e institucional. En conjunto, las reglas interpretadas integran combinaciones relacionadas con formación académica, trayectoria profesional, percepciones sobre discrepancias diagnósticas, y condiciones vinculadas con la solicitud o rechazo de autopsias, permitiendo una lectura más clara del fenómeno analizado.

Para el Hospital Regional Río Blanco, las reglas interpretadas tienden a describir perfiles relativamente compactos, coherentes con su baja redundancia. En este bloque predominan combinaciones que agrupan elementos de formación y posturas institucionales, lo que sugiere patrones más delimitados y menos repetitivos.

En el Hospital General de San Juan Bautista Tuxtepec, las reglas interpretadas muestran un perfil más estructurado y reiterativo, donde múltiples reglas describen variaciones cercanas de un mismo patrón subyacente. Este comportamiento es consistente con la redundancia alta observada en las reglas sin interpretar, lo que sugiere una región densa del espacio de búsqueda donde diferentes combinaciones conducen a interpretaciones semánticamente similares.

Para el Hospital General de Zona 53, las reglas interpretadas reflejan perfiles con alta variabilidad, desde patrones altamente representativos (incluyendo reglas de soporte máximo) hasta reglas más específicas. La combinación de alta cobertura en algunas reglas y redundancia elevada en el bloque sugiere la existencia de perfiles institucionales dominantes, descritos por múltiples reglas equivalentes, junto con subgrupos de menor tamaño.

En conjunto, la Tabla 4.16 permite vincular los resultados cuantitativos con descripciones comprensibles, habilitando el análisis visual y la evaluación integral del modelo desde perspectivas objetiva y subjetiva.

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad del Valle de México, no realizaron su especialidad en la UNAM, no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo, no han participado entre 11 y 20 casos de autopsias, no estudiaron su especialidad en Hospital Ángeles, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no realizaron su formación de medicina general en la Escuela Superior de Medicina, no estudiaron medicina general en Universidad Juárez Autónoma, no estudiaron medicina general en Hospital Juárez de México, y no estudiaron medicina general en La Raza IMSS; entonces son del Hospital Regional Río Blanco.
2	Si los médicos no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estudiaron su especialidad en IMSS, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital general, no estudiaron su especialidad en Hospital General

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	Dr. Miguel Silva, no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de trabajo social, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos financieros, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, sí especifican dónde realizaron su especialidad, no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron medicina general en Instituto Politécnico Nacional, no estudiaron su especialidad en Universidad de Nuevo León, no estudiaron su especialidad en Escuela Militar de Medicina, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron su especialidad en Estudios de Posgrado de Excelencia, no estudiaron su especialidad en Hospital General La Villa Urgencias, no estudiaron su especialidad en Centro Médico Nacional La Raza, no estudiaron su especialidad en Hospital Juárez de México, no estudiaron su especialidad en Universidad Autónoma de México, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, y no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza; entonces son del Hospital Regional Río Blanco.
3	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, no realizaron su especialidad en la UNAM, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en Estudios de Posgrado de Excelencia, y no opinan que se debe obligar a los residentes; entonces son del Hospital Regional Río Blanco.
4	Si los médicos no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no estudiaron medicina general en Universidad del Valle de México, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Hospital Juárez de México, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron su especialidad en Universidad de Guanajuato, no estudiaron su especialidad en Hospital General Balbuena, y no estudiaron medicina general en Escuela Superior de Medicina; entonces son del Hospital Regional Río Blanco.
5	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no tienen 11-15 años de práctica, no estudiaron su especialidad en Hospital General Palenque, no piensan que una causa o motivo para no solicitar

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	autopsia es que no exista el servicio, no consideran que quien debe solicitar la autopsia es una enfermera, no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron medicina general en Universidad Juárez Autónoma, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, no estudiaron medicina general en Universidad Autónoma de México, y no estudiaron medicina general en Hospital Hermanos Ameijeiras; entonces son del Hospital Regional Río Blanco.
6	Si los médicos no creen que la causa de rechazo por parte del familiar es por inconformidad de los familiares con la atención prestada al paciente, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad del Valle de México, no estiman como motivo para solicitar autopsia el consumo de medicamentos mal recetados, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Anáhuac, no estudiaron su especialidad en IMSS, no realizaron su especialidad en la UNAM, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron medicina general en Estudios de Posgrado de Excelencia, no opinan que el estudio aclara las causas de muerte, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos materiales, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no estudiaron su especialidad en Centro Médico Nacional La Raza, y no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo; entonces son del Hospital Regional Río Blanco.
7	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología y no estudiaron su especialidad en Facultad de Medicina; entonces son del Hospital General de San Juan Bautista Tuxtepec.
8	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, y no estudiaron medicina general en Alta Especialidad en IMSS; entonces son del Hospital General de San Juan Bautista Tuxtepec.
9	Si los médicos están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, consideran que quien debe solicitar la autopsia es una institución, consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no piensan que una causa o motivo para no solicitar autopsia es por cuestiones religiosas, no estudiaron

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	medicina general en Regional de Occidente, no realizaron su especialidad en la UNAM, y no estudiaron medicina general en la Universidad Xochicolco; entonces son del Hospital General de San Juan Bautista Tuxtepec.
10	Si los médicos están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Regional del Sureste, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, sí especifican dónde realizaron su especialidad, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, y no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de San Juan Bautista Tuxtepec.
11	Si los médicos no son estudiantes de medicina, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron su especialidad en Universidad Autónoma de México, no estudiaron su especialidad en Facultad de Medicina, no estudiaron su especialidad en Universidad de San Carlos, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no opinan que para este estudio deberían participar todos los servicios, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron medicina general en Instituto Nacional de Cancerología, no estudiaron su especialidad en Nachón/CEM, no estiman como motivo para solicitar autopsia la orientación diagnóstica, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron medicina general en Universidad Popular Autónoma, no plantean como motivo para solicitar autopsia cuando existe falta de correlación entre los datos clínicos y los de laboratorio, no consideran que quien debe solicitar la autopsia es un familiar, no estudiaron medicina general en Universidad Benito Juárez, no estudiaron medicina general en Hospital General Regional IMSS, no piensan que una causa o motivo para no solicitar autopsia es cuando el cuerpo se encuentra en descomposición, no plantean como motivo para solicitar autopsia cuando es necesario determinar si fue correcta o no la aplicación del tratamiento empleado, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron su especialidad en Instituto Nacional de Pediatría, no están totalmente en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Villa Rica, no estudiaron su especialidad en Instituto de Salud Pública, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Secretaría de Salud, no

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	estudiaron su especialidad en Universidad Benito Juárez, no estudiaron medicina general en Universidad Veracruzana, no creen que la causa de rechazo por parte del familiar es circunstancial, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron su especialidad en Centro Médico de Occidente, no estudiaron su especialidad en Hospital de la Mujer, no estudiaron su especialidad en Hospital Clínico Quirúrgico Hermanos Ameijeiras, no estudiaron su especialidad en Alta Especialidad en IMSS, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital General Palenque, no estudiaron su especialidad en Hospital General La Villa Urgencias, no estudiaron su especialidad en Hospital general, no estudiaron medicina general en Universidad Autónoma de México, no estudiaron su especialidad en Centro Médico Nacional La Raza, no estudiaron su especialidad en Hospital Juárez de México, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, y no estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de San Juan Bautista Tuxtepec.
12	Si los médicos consideran que quien debe solicitar la autopsia es una institución, estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron medicina general en Alta Especialidad en IMSS, no estudiaron medicina general en La Raza IMSS, no estudiaron medicina general en Regional de Occidente, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en una Universidad Autónoma, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no estudiaron medicina general en la Universidad Xochicolco, y no estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de San Juan Bautista Tuxtepec.
13	Si los médicos consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí comentan, no estudiaron su especialidad en Facultad de Medicina, no plantean como motivo para solicitar autopsia cuando es necesario proporcionar información esencial a familiares en casos de enfermedades infecciosas o congénitas, no estudiaron su especialidad en Universidad Popular Autónoma, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estudiaron su especialidad en Secretaría de Salud, no realizaron su especialidad en la UNAM, no estudiaron su especialidad en Benemérita

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	Universidad Autónoma de Puebla, y no estudiaron medicina general en Regional de Occidente; entonces son del Hospital General de San Juan Bautista Tuxtepec.
14	Si los médicos consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no son estudiantes de medicina, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estudiaron su especialidad en Facultad de Medicina, no estudiaron su especialidad en Universidad Autónoma de México, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos materiales, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no estudiaron su especialidad en Secretaría de Salud, y no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla; entonces son del Hospital General de San Juan Bautista Tuxtepec.
15	Si los médicos consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no plantean como motivo para solicitar autopsia el diagnóstico difícil, no están de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no opinan que nunca le han realizado autopsias a sus pacientes, no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo, no estudiaron medicina general en San Nicolás de Hidalgo, y no estudiaron su especialidad en Hospital General Dr. Miguel Silva; entonces son del Hospital General de San Juan Bautista Tuxtepec.
16	Si los médicos estudiaron su especialidad en Hospital Hermanos Ameijeiras y no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica; entonces son del Hospital General de Zona 53.
17	Si los médicos realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de Zona 53.
18	Si los médicos estudiaron medicina general en Universidad del Valle de México y están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos; entonces son del Hospital General de Zona 53.
19	Si los médicos estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital General de Zona 53.
20	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, no están totalmente de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron su especialidad en Facultad de Medicina, sí comentan, no estudiaron

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	medicina general en Universidad de San Carlos, no tienen más de 20 años de práctica, no han participado en más de 20 casos de autopsias, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron medicina general en Instituto Nacional de Cancerología, no piensan que una causa o motivo para no solicitar autopsia es por los trámites administrativos, no estudiaron su especialidad en Universidad Popular Autónoma, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron medicina general en Universidad Villa Rica, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron medicina general en Hospital General Dr. Manuel Gea González, y no estudiaron su especialidad en Universidad de San Carlos; entonces son del Hospital General de Zona 53.
21	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, sí comentan, no estudiaron su especialidad en Facultad de Medicina, no estudiaron medicina general en Universidad de San Carlos, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron medicina general en Universidad Popular Autónoma, no estudiaron medicina general en Regional de Occidente, no piensan que una causa o motivo para no solicitar autopsia es que se haga sin fines de enseñanza, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron su especialidad en Hospital General Balbuena, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Hospital General Dr. Manuel Gea González, no estudiaron su especialidad en Universidad Autónoma de Guadalajara, no estudiaron medicina general en Universidad Autónoma Benito Juárez de Oaxaca, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron medicina general en Hospital Regional De Alta Especialidad de Veracruz, no estudiaron su especialidad en Universidad La Salle, no estudiaron su especialidad en IMSS Adolfo Ruiz Cortines, no estudiaron su especialidad en Secretaría de Salud, no opinan que se debe obligar a los residentes, no opinan que se debe solicitar el consentimiento, no estudiaron su especialidad en Universidad Veracruzana, no estudiaron medicina general en Hospital General de Orizaba, no estudiaron su especialidad en Universidad de Guadalajara, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no estudiaron medicina general en la Universidad Popular Autónoma de Puebla, no estudiaron su especialidad en Hospital general Miguel Enríquez, no estudiaron medicina general en Instituto Politécnico Nacional, no estudiaron su especialidad en Estudios de Posgrado de Excelencia, no opinan sobre fortalecer la relación

Tabla 4.16. *Correlation* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	entre anatomía patológica y cirugía general, y no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda; entonces son del Hospital General de Zona 53.
22	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron su especialidad en Facultad de Medicina, no estudiaron medicina general en Universidad de San Carlos, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Hospital del Departamento del D.F., no consideran que quien debe solicitar la autopsia es un familiar, no estudiaron medicina general en Hospital General Regional IMSS, no opinan que nunca le han realizado autopsias a sus pacientes, no estudiaron medicina general en Universidad Regional del Sureste, no estudiaron medicina general en una Universidad Autónoma, no estudiaron medicina general en la Facultad de Ciencias Médicas, no estudiaron medicina general en Estudios de Posgrado de Excelencia, no piensan que una causa o motivo para no solicitar autopsia es por los trámites administrativos, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron medicina general en Universidad Benito Juárez, no piensan que una causa o motivo para no solicitar autopsia es cuando el cuerpo se encuentra en descomposición, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos humanos, no estudiaron medicina general en Hospital General, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Universidad de Colima, no estudiaron su especialidad en Universidad Benito Juárez, no piensan que una causa o motivo para no solicitar autopsia es el olvido en la práctica de este estudio, no estudiaron medicina general en Ignacio Chávez, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Universidad de Nuevo León, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron medicina general en Universidad Cristóbal Colón, y no estudiaron su especialidad en Universidad de San Carlos; entonces son del Hospital General de Zona 53.

4.9.3 Análisis visual

El análisis visual permite observar el comportamiento global de las métricas asociadas a las reglas generadas mediante Correlation, facilitando la identificación de rangos, concentraciones y contrastes entre patrones. A continuación, se presentan las gráficas

correspondientes a soporte, soporte negativo, diferencia de soporte, precisión, diferencia de precisión y redundancia.

La Figura 4.58 muestra la distribución del soporte para las veintidós reglas generadas. Se observa un rango amplio, desde valores cercanos a 0.11 hasta valores máximos de 1.00, lo que indica la coexistencia de reglas altamente representativas con otras más específicas. La concentración principal se ubica en valores moderados y altos, reflejando patrones con cobertura relevante dentro del hospital objetivo.

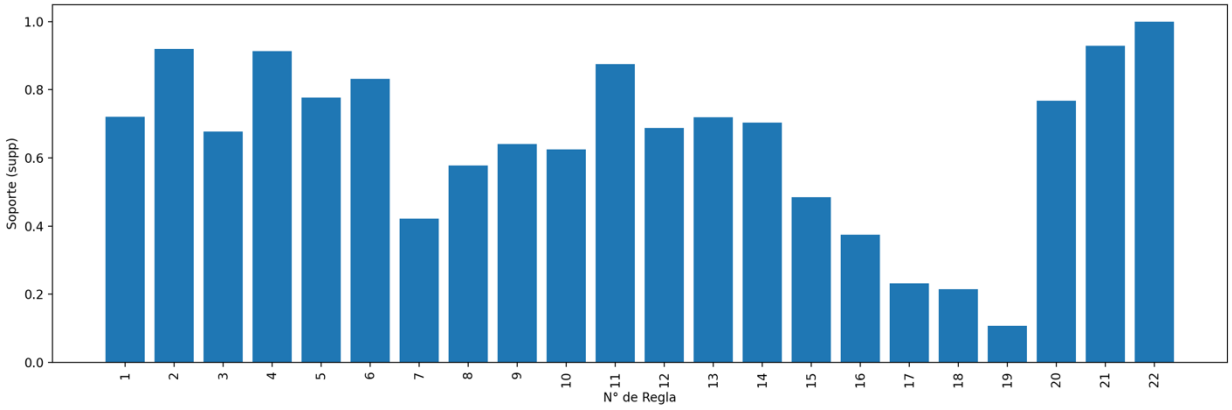


Figura 4.58 *Correlation* – Matriz Binaria: Soporte por Regla

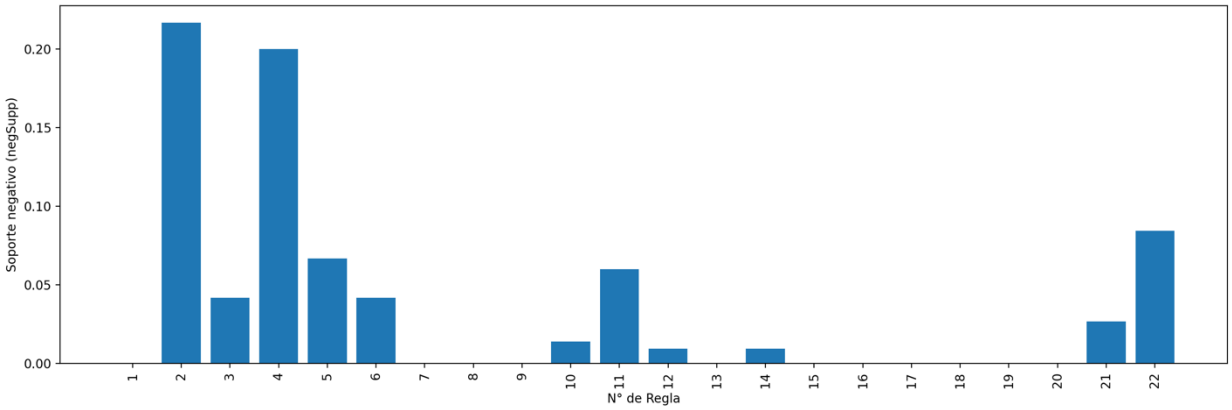


Figura 4.59 *Correlation* – Matriz Binaria: Soporte Negativo por Regla

La Figura 4.59 presenta el comportamiento del soporte negativo, donde se observa que una parte del conjunto mantiene valores nulos, mientras que otras reglas exhiben valores

positivos que alcanzan aproximadamente 0.22. Este comportamiento confirma la presencia de reglas exclusivas del hospital objetivo junto con reglas donde existen contraejemplos en otros hospitales, afectando la pureza de ciertos patrones.

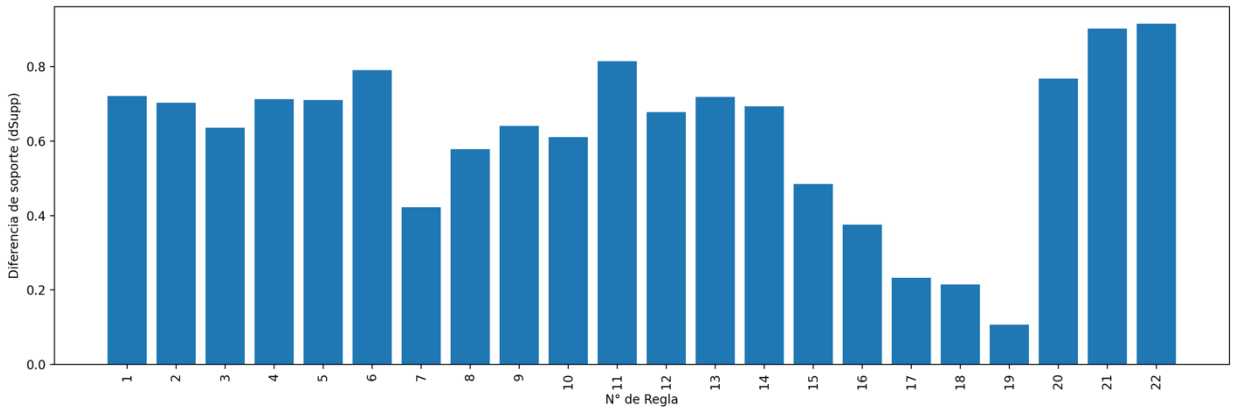


Figura 4.60 *Correlation* – Matriz Binaria: Diferencia de Soporte por Regla

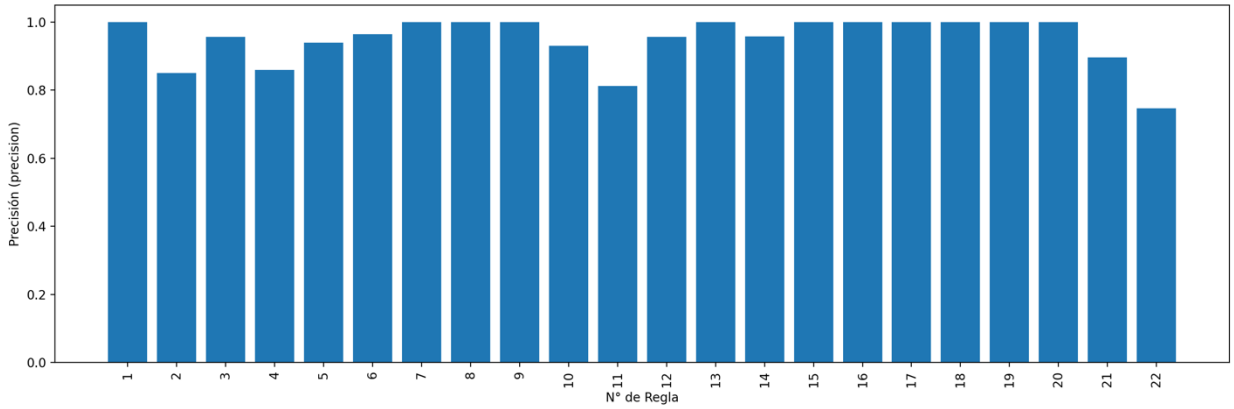


Figura 4.61 *Correlation* – Matriz Binaria: Precisión por Regla

La Figura 4.61 evidencia que la precisión se mantiene en rangos altos, desde valores cercanos a 0.75 hasta 1.00. Se observa un subconjunto de reglas con precisión perfecta, mientras que otras presentan descensos moderados asociados al incremento del soporte negativo, confirmando un modelo altamente discriminativo, pero no uniformemente perfecto.

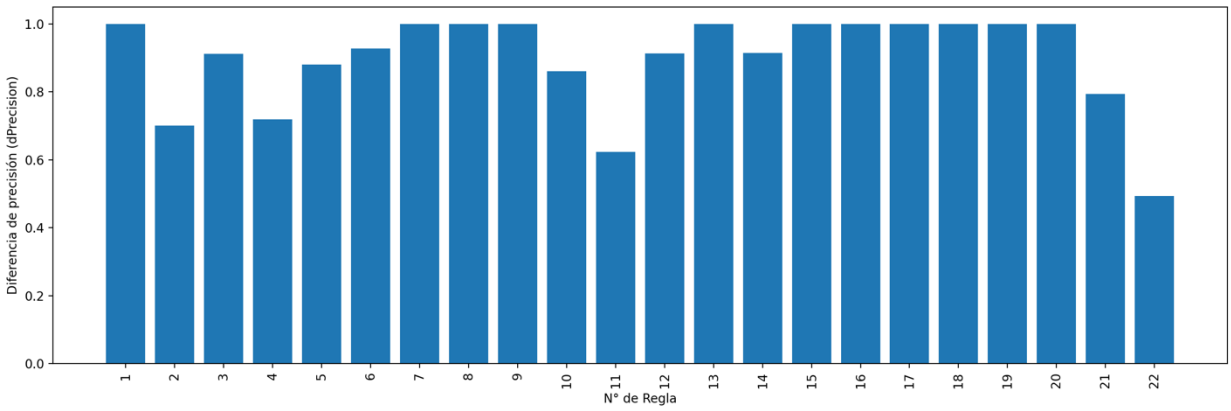


Figura 4.62 *Correlation* – Matriz Binaria: Diferencia de Precisión por Regla

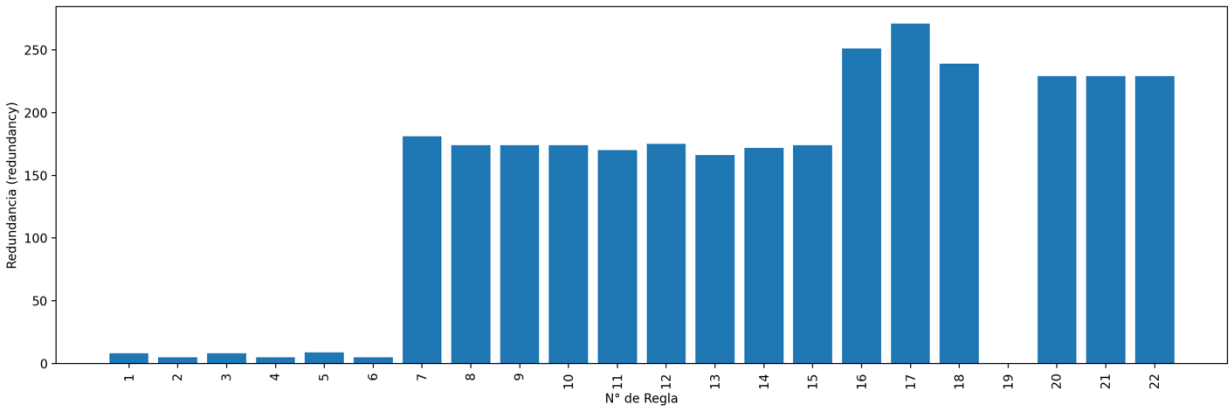


Figura 4.63 *Correlation* – Matriz Binaria: Redundancia por Regla

La Figura 4.63 muestra una separación marcada en la redundancia: un subconjunto de reglas presenta valores bajos (aproximadamente entre 5 y 9), mientras que otro bloque concentra valores muy altos (aproximadamente entre 166 y 271). Este comportamiento indica un solapamiento informativo elevado en una parte del conjunto, por lo que la redundancia se vuelve un criterio relevante para priorizar reglas en etapas posteriores de síntesis.

4.9.4 Evaluación objetiva

La evaluación objetiva del modelo *Correlation* aplicado a Matriz Binaria muestra un comportamiento cuantitativamente consistente, con diferencias de soporte positivas en todas las reglas, lo que confirma que cada patrón incrementa su presencia relativa en el

hospital objetivo frente al resto de la población. Los valores de soporte se concentran mayormente en rangos moderados y altos, indicando un modelo con cobertura relevante.

Se observa una combinación de reglas con soporte negativo nulo y reglas con soporte negativo positivo (hasta alrededor de 0.22), lo que implica la presencia de patrones tanto exclusivos como no exclusivos. En consecuencia, la precisión se mantiene alta en general, pero puede descender hasta rangos cercanos a 0.75 en algunas reglas, evidenciando variación de pureza entre patrones.

La redundancia representa el factor cuantitativo más crítico: coexisten reglas con redundancia baja y un bloque con redundancia alta, particularmente concentrado en H11 y H12. Esto sugiere que parte del conjunto contiene patrones estructuralmente equivalentes y que, para fines de síntesis, resulta adecuado priorizar reglas con alta diferencia de soporte y menor redundancia, preservando diversidad informativa.

4.9.5 Evaluación subjetiva

En el caso de la función de calidad *Correlation*, la evaluación subjetiva mostró una alta correspondencia entre los patrones generados y los perfiles clínicos reales. La mayoría de las reglas fue validada positivamente por el Dr. Palet, al reflejar combinaciones de atributos que coinciden con dinámicas institucionales observadas en los hospitales analizados, particularmente en lo relativo a la experiencia profesional, las percepciones diagnósticas y los criterios para la realización de autopsias.

Las pocas reglas descartadas corresponden a patrones que, si bien presentan correlación estadística, no aportan una diferenciación clínica clara entre hospitales. En términos generales, Correlation se percibe como una métrica que produce reglas sólidas, interpretables y alineadas con la realidad clínica, requiriendo un filtrado semántico mínimo en comparación con otras funciones de calidad.

4.10 Matriz Binaria: Resultados aplicando la función de calidad *Klosgen*

La función de calidad *Klosgen* se caracteriza por evaluar simultáneamente la frecuencia de un patrón y su capacidad discriminativa, penalizando aquellos conjuntos de condiciones que, aun siendo frecuentes, no aportan una ganancia clara respecto al comportamiento global del conjunto de datos. A diferencia de métricas centradas exclusivamente en la ganancia o la cobertura, *Klosgen* busca un equilibrio entre ambos factores, favoreciendo patrones que combinan representatividad y contraste.

En este apartado se presentan los resultados del análisis de Minería de Conjuntos de Contraste aplicado al conjunto de datos Matriz Binaria, en nivel específico, utilizando la función de calidad *Klosgen*. La estructura del análisis mantiene el mismo estándar metodológico seguido en los apartados previos, incluyendo la presentación de reglas sin interpretar, su interpretación semántica, el análisis visual de métricas y la evaluación objetiva y subjetiva del modelo generado.

4.10.1 Reglas sin interpretar

La Tabla 4.17, correspondiente a las reglas sin interpretar generadas mediante la función de calidad *Klosgen*, muestra un conjunto de dieciséis patrones distribuidos entre los tres hospitales analizados. En términos generales, el modelo presenta reglas con una amplia variabilidad en los valores de soporte y soporte negativo, lo que se refleja en diferencias de soporte tanto moderadas como elevadas y en un comportamiento menos homogéneo que el observado en métricas como *WRAcc* o *Correlation*.

Las reglas asociadas al Hospital Regional Río Blanco conforman un primer bloque caracterizado por soportes que oscilan desde valores moderados hasta valores muy elevados, con diferencias de soporte consistentemente positivas. En este grupo se observan reglas con precisiones altas, cercanas a 1.0, así como reglas donde la precisión desciende de forma notable debido al incremento del soporte negativo. Este comportamiento sugiere que, en Río Blanco, *Klosgen* identifica tanto patrones altamente representativos como otros que, aun siendo frecuentes, comparten características con

perfiles de otros hospitales. La redundancia en este bloque se mantiene mayoritariamente en niveles bajos o moderados, indicando que las reglas capturan regiones diferenciadas del espacio de búsqueda.

El conjunto de reglas vinculadas al Hospital General de San Juan Bautista Tuxtepec presenta soportes elevados y diferencias de soporte pronunciadas, lo que evidencia patrones fuertemente contrastivos. En este bloque, los valores de precisión se mantienen en rangos altos y las diferencias de precisión son consistentemente positivas, confirmando una ganancia clara respecto al comportamiento global. Sin embargo, la redundancia extremadamente alta observada en estas reglas indica que múltiples combinaciones de atributos conducen a patrones equivalentes, reflejando una estructura institucional altamente homogénea y densa desde el punto de vista del espacio de búsqueda.

Finalmente, las reglas correspondientes al Hospital General de Zona 53 conforman un bloque con una notable variabilidad en soporte, incluyendo reglas de cobertura amplia y otras más específicas. Aunque las diferencias de soporte se mantienen mayoritariamente positivas, se identifican reglas con precisiones moderadas e incluso con diferencias de precisión menores a 0.31. Este comportamiento revela que *Klosgen* también recupera patrones muy frecuentes que no siempre maximizan la discriminación entre hospitales. La redundancia elevada en este bloque reafirma la presencia de regiones densas del espacio de búsqueda y de perfiles institucionales descritos por múltiples reglas similares.

En conjunto, la Tabla 4.17 evidencia que *Klosgen* produce un conjunto de reglas con un equilibrio variable entre cobertura y discriminación, destacando patrones muy representativos, pero también reglas cuya utilidad discriminativa es más limitada en comparación con otras métricas.

Tabla 4.17. *Klosgen* – Matriz Binaria: Reglas Sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF p9_1 = !{S} AND p17_115 = !{S} AND p17_107 = !{S} AND p16_112 = !{S} AND p17_38 = !{S} AND p16_86 = !{S} AND p12_-1 = !{S} AND p17_102 = !{S} AND p17_103 = !{S} AND p17_106 = !{S} AND p16_114 = !{S} AND p16_121 = !{S} AND p17_123 = !{S} AND p16_124 = !{S} AND p16_18 = !{S} AND p16_129 = !{S} AND p16_23 = !{S} THEN centro_hospitalario = {H1}"	0.739130	0.000000	0.739130	1.000000	1.000000	8.000000
2	"IF p9_1 = !{S} AND p16_112 = !{S} AND p17_111 = !{S} AND p17_122 = !{S} AND p17_124 = !{S} AND p17_126 = !{S} AND p17_130 = !{S} AND p13_13 = !{S} AND p16_99 = !{S} AND p13_15 = !{S} AND p17_114 = !{S} AND p16_131 = !{S} AND p16_115 = !{S} THEN centro_hospitalario = {H1}"	0.745342	0.083333	0.662008	0.923077	0.846154	8.000000
3	"IF p11_4 = !{S} AND p17_115 = !{S} AND p9_6 = !{S} AND p7_16 = !{S} AND p17_110 = !{S} AND p17_117 = !{S} AND p17_38 = !{S} AND p7_1 = !{S} AND p17_101 = !{S} THEN centro_hospitalario = {H1}"	0.751553	0.066667	0.684886	0.937984	0.875969	8.000000
4	"IF p11_4 = !{S} AND p17_115 = !{S} AND p11_-1 = !{S} AND p11_2 = !{S} AND p16_99 = !{S} AND p17_105 = !{S} AND p17_128 = !{S} AND p17_129 = !{S} AND p17_125 = !{S} AND p17_133 = !{S} THEN centro_hospitalario = {H1}"	0.763975	0.108333	0.655642	0.904412	0.808824	6.000000
5	"IF p11_4 = !{S} AND p16_112 = !{S} AND p16_99 = !{S} AND p17_107 = !{S} AND p7_16 = !{S} AND p16_98 = !{S} AND p17_118 = !{S} AND p16_125 = !{S} AND p16_121 = !{S} AND p16_131 = !{S} THEN centro_hospitalario = {H1}"	0.838509	0.100000	0.738509	0.918367	0.836735	6.000000

Tabla 4.17. *Klosgen* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
6	"IF p12_7 = !{S} AND p17_115 = !{S} AND p8_5 = !{S} AND p12_-1 = !{S} AND p16_112 = !{S} AND p17_107 = !{S} AND p16_93 = !{S} AND p16_129 = !{S} AND p17_100 = !{S} AND p17_101 = !{S} AND p17_102 = !{S} AND p17_110 = !{S} AND p17_111 = !{S} AND p17_122 = !{S} AND p16_18 = !{S} THEN centro_hospitalario = {H1}"	0.832298	0.050000	0.782298	0.957143	0.914286	5.000000
7	"IF p12_7 = !{S} AND p17_115 = !{S} AND p9_6 = !{S} AND p16_38 = !{S} AND p12_8 = !{S} AND p17_132 = !{S} AND p17_38 = !{S} AND p17_117 = !{S} AND p17_126 = !{S} AND p17_119 = !{S} THEN centro_hospitalario = {H1}"	0.782609	0.058333	0.724275	0.947368	0.894737	6.000000
8	"IF p4_1 = !{S} AND p17_115 = !{S} AND p16_38 = !{S} AND p16_112 = !{S} AND p16_120 = !{S} AND p16_99 = !{S} AND p13_15 = !{S} AND p17_103 = !{S} AND p16_108 = !{S} AND p17_118 = !{S} AND p17_124 = !{S} AND p17_132 = !{S} THEN centro_hospitalario = {H1}"	0.776398	0.075000	0.701398	0.932836	0.865672	7.000000
9	"IF p17_115 = !{S} AND p9_6 = !{S} AND p2_4 = !{S} AND p17_38 = !{S} AND p16_93 = !{S} AND p16_38 = !{S} AND p17_125 = !{S} AND p13_9 = !{S} AND p16_129 = !{S} THEN centro_hospitalario = {H1}"	0.838509	0.141667	0.696843	0.888158	0.776316	6.000000
10	"IF p17_10 = !{S} AND p6_5 = !{S} AND p16_17 = !{S} AND p16_33 = !{S} AND p7_13 = !{S} AND p13_8 = !{S} AND p17_115 = !{S} AND p3_3 = !{S} AND p3_2 = !{S} AND p16_9 = !{S} AND p17_2 = !{S} AND p17_25 = !{S} AND p17_3 = !{S} AND p16_27 = !{S} AND p3_4 = !{S} AND p6_7 = !{S}	0.828125	0.027650	0.800475	0.898305	0.796610	167.000000

Tabla 4.17. *Klosgen* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
	AND p17_15 = !{S} AND p17_20 = !{S} AND p3_5 = !{S} AND p11_7 = !{S} AND p17_31 = !{S} AND p17_22 = !{S} AND p17_49 = !{S} AND p13_16 = !{S} AND p13_18 = !{S} AND p17_17 = !{S} AND p16_1 = !{S} AND p17_8 = !{S} AND p17_82 = !{S} AND p17_84 = !{S} AND p16_95 = !{S} AND p17_123 = !{S} AND p17_130 = !{S} AND p16_131 = !{S} AND p17_131 = !{S} THEN centro_hospitalario = {H11}"						
11	"IF p10_4 = {S} AND p6_4 = !{S} AND p17_38 = !{S} AND p5_4 = !{S} AND p7_9 = !{S} AND p10_5 = !{S} AND p4_3 = !{S} AND p16_11 = !{S} AND p13_4 = !{S} AND p16_16 = !{S} AND p16_86 = !{S} AND p17_13 = !{S} AND p17_128 = !{S} AND p17_132 = !{S} AND p17_133 = !{S} THEN centro_hospitalario = {H11}"	0.796875	0.069124	0.727751	0.772727	0.545455	174.000000
12	"IF p17_10 = !{S} AND p16_8 = !{S} AND p3_4 = !{S} AND p11_6 = !{S} AND p9_-1 = !{S} AND p16_129 = !{S} AND p16_27 = !{S} AND p7_18 = !{S} AND p17_18 = !{S} AND p16_47 = !{S} AND p7_12 = !{S} AND p7_14 = !{S} AND p16_35 = !{S} AND p17_21 = !{S} AND p17_43 = !{S} AND p17_29 = !{S} AND p16_30 = !{S} AND p16_50 = !{S} AND p8_7 = !{S} AND p16_41 = !{S} AND p13_10 = !{S} AND p17_85 = !{S} AND p16_88 = !{S} AND p16_89 = !{S} AND p17_92 = !{S} AND p16_96 = !{S} AND p17_117 = !{S} AND p17_118 = !{S} AND p17_122 = !{S} AND p17_124 = !{S} AND p16_125 = !{S} AND p17_129 = !{S} AND p17_125 =	0.906250	0.129032	0.777218	0.674419	0.348837	168.000000

Tabla 4.17. *Klosgen* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
	!{S} AND p16_115 = !{S} AND p7_13 = !{S} AND p4_3 = !{S} AND p13_8 = !{S} AND p16_131 = !{S} THEN centro_hospitalario = {H11}"						
13	"IF p17_115 = !{S} AND p3_4 = !{S} AND p17_10 = !{S} AND p13_1 = !{S} AND p6_6 = !{S} AND p16_41 = !{S} AND p12_5 = !{S} AND p13_6 = !{S} AND p16_26 = !{S} AND p4_5 = !{S} AND p13_13 = !{S} AND p7_15 = !{S} AND p17_42 = !{S} AND p17_51 = !{S} AND p17_1 = !{S} AND p17_88 = !{S} AND p17_27 = !{S} AND p17_71 = !{S} AND p16_121 = !{S} AND p16_124 = !{S} AND p17_126 = !{S} AND p16_9 = !{S} AND p17_2 = !{S} AND p7_9 = !{S} AND p10_5 = !{S} AND p16_1 = !{S} AND p17_8 = !{S} AND p17_13 = !{S} AND p16_129 = !{S} AND p17_125 = !{S} THEN centro_hospitalario = {H11}"	0.968750	0.207373	0.761377	0.579439	0.158879	166.000000
14	"IF p3_2 = !{S} AND p6_5 = !{S} AND p11_8 = !{S} AND p13_16 = !{S} AND p17_130 = !{S} AND p17_14 = !{S} AND p17_16 = !{S} AND p9_8 = !{S} AND p16_12 = !{S} AND p7_22 = !{S} AND p17_32 = !{S} AND p16_97 = !{S} AND p17_119 = !{S} AND p17_10 = !{S} AND p17_25 = !{S} AND p17_3 = !{S} AND p16_16 = !{S} AND p16_17 = !{S} AND p16_47 = !{S} AND p16_33 = !{S} AND p6_7 = !{S} AND p16_26 = !{S} AND p17_15 = !{S} AND p3_5 = !{S} AND p4_5 = !{S} AND p17_43 = !{S} AND p17_31 = !{S} AND p17_22 = !{S} AND p13_18 = !{S} AND p17_38 = !{S} AND p17_17 =	0.906250	0.101382	0.804868	0.725000	0.450000	165.000000

Tabla 4.17. *Klosgen* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
	!{S} AND p13_10 = !{S} AND p17_82 = !{S} AND p17_88 = !{S} AND p16_95 = !{S} AND p17_118 = !{S} AND p17_123 = !{S} AND p17_131 = !{S} AND p16_27 = !{S} AND p13_8 = !{S} AND p16_41 = !{S} AND p17_115 = !{S} THEN centro_hospitalario = {H11}"						
15	"IF p17_115 = {S} AND p9_6 = !{S} THEN centro_hospitalario = {H12}"	0.375000	0.000000	0.375000	1.000000	1.000000	251.000000
16	"IF p16_8 = !{S} AND p17_10 = !{S} AND p13_1 = !{S} AND p16_13 = !{S} AND p16_99 = !{S} AND p11_6 = !{S} AND p6_4 = !{S} AND p16_41 = !{S} AND p8_-1 = !{S} AND p16_88 = !{S} AND p16_9 = !{S} AND p16_11 = !{S} AND p7_7 = !{S} AND p17_16 = !{S} AND p16_35 = !{S} AND p16_26 = !{S} AND p16_12 = !{S} AND p13_15 = !{S} AND p12_8 = !{S} AND p17_13 = !{S} AND p16_71 = !{S} THEN centro_hospitalario = {H12}"	1.000000	0.133333	0.866667	0.651163	0.302326	229.000000

4.10.2 Reglas interpretadas

La Tabla 4.18, correspondiente a las reglas interpretadas generadas mediante *Klosgen*, presenta la traducción semántica de los dieciocho patrones obtenidos, permitiendo comprender su significado desde una perspectiva clínica e institucional.

Las reglas interpretadas del Hospital Regional Río Blanco describen perfiles relacionados con la formación académica, la experiencia clínica y las percepciones sobre la utilidad diagnóstica de las autopsias. En este bloque se identifican patrones

que reflejan prácticas institucionales claras, aunque también aparecen combinaciones más generales que comparten elementos con otros hospitales, lo cual es consistente con las precisiones moderadas observadas en algunas reglas.

En el Hospital General de San Juan Bautista Tuxtepec, las reglas interpretadas conforman perfiles altamente consistentes, caracterizados por posturas bien definidas respecto a los motivos para solicitar o rechazar autopsias, así como por trayectorias formativas similares. La elevada precisión de estas reglas respalda su validez clínica y su fuerte capacidad discriminativa, aunque su redundancia sugiere que dichas interpretaciones se repiten bajo diferentes combinaciones de atributos.

Por su parte, el Hospital General de Zona 53 presenta reglas interpretadas con estructuras más extensas, que integran múltiples dimensiones formativas y perceptuales. Estas reglas reflejan la diversidad de perfiles clínicos presentes en el hospital, aunque su menor precisión en algunos casos indica que los patrones descritos no son exclusivos de este centro y pueden compartirse parcialmente con otros hospitales.

En conjunto, la Tabla 4.18 permite vincular los resultados cuantitativos del modelo *Klosgen* con interpretaciones semánticas coherentes, facilitando la comprensión de los perfiles institucionales identificados.

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad del Valle de México, no realizaron su especialidad en la UNAM, no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo, no estudiaron su especialidad en Hospital Ángeles, no estudiaron su especialidad en IMSS, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron medicina general en Estudios de Posgrado de Excelencia, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no realizaron su formación de medicina general en la Escuela Superior de Medicina, no estudiaron medicina general en Universidad Juárez Autónoma, no estudiaron medicina general en Hospital Juárez de

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	México, y no estudiaron medicina general en La Raza IMSS; entonces son del Hospital Regional Río Blanco.
2	Si los médicos no consideran que no se realizan suficientes autopsias en el hospital por falta de recursos humanos, no estudiaron medicina general en Universidad del Valle de México, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron su especialidad en Hospital general, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no opinan que se debe incrementar el número de autopsias, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no opinan que se debe obligar a los residentes, no estudiaron su especialidad en Estudios de Posgrado de Excelencia, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, y no estudiaron medicina general en Hospital Hermanos Ameijeiras; entonces son del Hospital Regional Río Blanco.
3	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos financieros, no estudiaron su especialidad en Escuela Militar de Medicina, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no realizaron su especialidad en la UNAM, no piensan que una causa o motivo para no solicitar autopsia es que no exista el servicio, y no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado; entonces son del Hospital Regional Río Blanco.
4	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estiman como motivo para solicitar autopsia un error médico, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron su especialidad en Centro Médico Nacional La Raza, no estudiaron su especialidad en Hospital Juárez de México, y no estudiaron su especialidad en Universidad Autónoma de México; entonces son del Hospital Regional Río Blanco.
5	Si los médicos no estiman como motivo para solicitar autopsia la negligencia médica, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en Secretaría de Salud de Oaxaca, sí especifican dónde realizaron su especialidad, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos financieros, no estudiaron medicina general en una Universidad Autónoma, no estudiaron su especialidad en Hospital General Palenque, no estudiaron medicina general en Universidad Autónoma de México, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, y no realizaron

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza; entonces son del Hospital Regional Río Blanco.
6	Si los médicos no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no creen que la causa de rechazo por parte del familiar es por inconformidad de los familiares con la atención prestada al paciente, no estudiaron medicina general en Universidad del Valle de México, sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en Hospital Juárez de México, no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, no estudiaron su especialidad en Hospital Ángeles, no estudiaron su especialidad en Escuela Militar de Medicina, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron su especialidad en Hospital general, y no estudiaron medicina general en Universidad Juárez Autónoma; entonces son del Hospital Regional Río Blanco.
7	Si los médicos no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través del área de patología, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica, no estudiaron medicina general en UNAM, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron su especialidad en Centro Médico Nacional La Raza, no realizaron su especialidad en la UNAM, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, y no estudiaron su especialidad en Hospital General La Villa Urgencias; entonces son del Hospital Regional Río Blanco.
8	Si los médicos no están totalmente de acuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estudiaron medicina general en UNAM, no estudiaron medicina general en Universidad del Valle de México, no estudiaron medicina general en Escuela Superior de Medicina, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no opinan que se debe obligar a los residentes, no estudiaron su especialidad en IMSS, no estudiaron medicina general en Instituto Politécnico Nacional, no estudiaron su especialidad en Hospital General Palenque, no estudiaron su especialidad en Hospital General, y no estudiaron su especialidad en Centro Médico Nacional La Raza; entonces son del Hospital Regional Río Blanco.
9	Si los médicos no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica,

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	no han participado entre 11 y 20 casos de autopsias, no realizaron su especialidad en la UNAM, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en UNAM, no estudiaron su especialidad en Universidad Autónoma de México, no opinan que el estudio aclara las causas de muerte, y no estudiaron medicina general en Hospital Juárez de México; entonces son del Hospital Regional Río Blanco.
10	Si los médicos no estudiaron su especialidad en Facultad de Medicina, no plantean como motivo para solicitar autopsia cuando existe falta de correlación entre los datos clínicos y los de laboratorio, no estudiaron medicina general en Universidad Benito Juárez, no estudiaron medicina general en Regional de Occidente, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no consideran incierto que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron su especialidad en Nachón/CEM, no estudiaron su especialidad en Ignacio Chávez, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no están en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no plantean como motivo para solicitar autopsia cuando es necesario determinar si fue correcta o no la aplicación del tratamiento empleado, no estudiaron su especialidad en Instituto Nacional de Pediatría, no estudiaron su especialidad en Instituto Nacional de Nutrición, no están totalmente en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estiman como motivo para solicitar autopsia la educación médica continua, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Secretaría de Salud, no estudiaron su especialidad en Universidad de Colima, no opinan que nunca le han realizado autopsias a sus pacientes, no opinan agregar patología a la rotación de los médicos internos, no estudiaron su especialidad en Universidad Benito Juárez, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Alta Especialidad en IMSS, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Universidad de Guadalajara, no estudiaron medicina general en Universidad Autónoma Benito Juárez de Oaxaca, sí especifican dónde realizaron su especialidad, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, y no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de San Juan Bautista Tuxtepec.

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
11	Si los médicos consideran que quien debe solicitar la autopsia es una institución, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no realizaron su especialidad en la UNAM, no están en desacuerdo en que los hallazgos de las autopsias dan origen a casos de demanda, no piensan que una causa o motivo para no solicitar autopsia es por cuestiones religiosas, no consideran que quien debe solicitar la autopsia es un familiar, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron medicina general en Instituto Nacional de Cancerología, no opinan que la encuesta es útil, no estudiaron medicina general en Universidad Popular Autónoma, no estudiaron medicina general en Universidad Michoacana de San Nicolás de Hidalgo, no estudiaron su especialidad en Universidad de San Carlos, no estudiaron su especialidad en Centro Médico Nacional La Raza, y no estudiaron su especialidad en Hospital Juárez de México; entonces son del Hospital General de San Juan Bautista Tuxtepec.
12	Si los médicos no estudiaron su especialidad en Facultad de Medicina, no estudiaron medicina general en Alta Especialidad en IMSS, no están en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estiman como motivo para solicitar autopsia la orientación diagnóstica, no estudiaron medicina general en Hospital Juárez de México, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no piensan que una causa o motivo para no solicitar autopsia es por factores sociales, no estudiaron su especialidad en Universidad Juárez Autónoma, no estudiaron medicina general en Hospital General Regional IMSS, no piensan que una causa o motivo para no solicitar autopsia es cuando el cuerpo se encuentra en descomposición, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos humanos, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron su especialidad en IMSS Adolfo Ruiz Cortines, no estudiaron su especialidad en Instituto de Salud Pública, no estudiaron su especialidad en Centro de Especialidades Médicas, no estudiaron medicina general en San Luis Potosí, no realizaron su formación de medicina general en la Facultad de Medicina, no creen que la causa de rechazo por parte del familiar es circunstancial, no estudiaron medicina general en Universidad Cristóbal Colón, no opinan que para este estudio deberían participar todos los servicios, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no estudiaron medicina general en Hospital General Dr. Manuel Gea González, no estudiaron medicina general en Universidad Regional del Sureste, no estudiaron su especialidad en Universidad Autónoma de Guadalajara, no estudiaron medicina general en Hospital Regional De Alta Especialidad de Veracruz, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital General Palenque, no estudiaron su especialidad en Hospital general, no estudiaron medicina general en Universidad Autónoma de México, no estudiaron su especialidad en Hospital Juárez

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas - Continuación

N.°	Regla Interpretada
	de México, no estudiaron su especialidad en Universidad Autónoma de México, no estudiaron medicina general en Hospital Hermanos Ameijeiras, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no consideran incierto que los hallazgos de las autopsias dan origen a casos de arbitraje, no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general, y no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de San Juan Bautista Tuxtepec.
13	Si los médicos no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no están en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron su especialidad en Facultad de Medicina, sí comentan, no plantean como motivo para solicitar autopsia cuando es necesario proporcionar información esencial a familiares en casos de enfermedades infecciosas o congénitas, no estudiaron medicina general en Universidad Cristóbal Colón, no están de acuerdo en que el método eficiente para solicitar autopsia es realizarlo a través de otra área, no opinan sobre observaciones a la encuesta, no estudiaron medicina general en Hospital General, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no opinan que se debe incrementar el número de autopsias, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos materiales, no estudiaron su especialidad en Centro Médico de Occidente, no estudiaron su especialidad en Hospital Clínico Quirúrgico Hermanos Ameijeiras, no estudiaron su especialidad en Universidad Veracruzana, no estudiaron su especialidad en Hospital General Dr. Manuel Gea González, no estudiaron su especialidad en UMAC 14 Adolfo Ruiz Cortines, no estudiaron su especialidad en Hospital general Miguel Enríquez, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no realizaron su formación de medicina general en la Escuela Superior de Medicina, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron su especialidad en Nachón/CEM, no piensan que una causa o motivo para no solicitar autopsia es por cuestiones religiosas, no consideran que quien debe solicitar la autopsia es un familiar, no estudiaron medicina general en Universidad Veracruzana, no estudiaron su especialidad en Alta Especialidad en IMSS, no estudiaron su especialidad en Universidad de San Carlos, no estudiaron medicina general en Hospital Juárez de México, y no estudiaron su especialidad en Universidad Autónoma de México; entonces son del Hospital General de San Juan Bautista Tuxtepec.
14	Si los médicos no están de acuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no plantean como motivo para solicitar autopsia cuando existe falta de correlación entre los datos clínicos y los de laboratorio, no estiman como

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
	motivo para solicitar autopsia un problema médico legal, no opinan que nunca le han realizado autopsias a sus pacientes, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no estudiaron su especialidad en Universidad La Salle, no estudiaron su especialidad en Universidad Popular Autónoma, no consideran que no se realizan suficientes autopsias en el hospital debido a la negación de los familiares, no estudiaron medicina general en Universidad Villa Rica, no piensan que una causa o motivo para no solicitar autopsia es el olvido en la práctica de este estudio, no estudiaron su especialidad en Hospital de la Mujer, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron su especialidad en Hospital General La Villa Urgencias, no estudiaron su especialidad en Facultad de Medicina, no estudiaron su especialidad en Ignacio Chávez, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron medicina general en Universidad Popular Autónoma, no estudiaron medicina general en Universidad Benito Juárez, no estudiaron medicina general en Hospital General Regional IMSS, no estudiaron medicina general en Regional de Occidente, no plantean como motivo para solicitar autopsia cuando es necesario determinar si fue correcta o no la aplicación del tratamiento empleado, no estudiaron medicina general en Hospital General, no estudiaron su especialidad en Instituto Nacional de Pediatría, no están totalmente en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no están totalmente en desacuerdo en que los hallazgos de las autopsias dan origen a casos de arbitraje, no estudiaron su especialidad en Instituto de Salud Pública, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Secretaría de Salud, no opinan agregar patología a la rotación de los médicos internos, no realizaron su especialidad en la UNAM, no estudiaron su especialidad en Universidad Benito Juárez, no opinan que para este estudio deberían participar todos los servicios, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Hospital General Dr. Manuel Gea González, no estudiaron medicina general en Universidad Autónoma Benito Juárez de Oaxaca, no estudiaron su especialidad en Hospital General Palenque, sí especifican dónde realizaron su especialidad, no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no opinan sobre fortalecer la relación entre anatomía patológica y cirugía general, no estudiaron medicina general en Universidad Cristóbal Colón, y no estudiaron su especialidad en Hospital Hermanos Ameijeiras; entonces son del Hospital General de San Juan Bautista Tuxtepec.
15	Si los médicos estudiaron su especialidad en Hospital Hermanos Ameijeiras y no consideran que no se realizan suficientes autopsias en el hospital porque se considera que los nuevos procedimientos diagnósticos y otros estudios especializados eliminan la necesidad de su práctica; entonces son del Hospital General de Zona 53.

Tabla 4.18. *Klosgen* – Matriz Binaria: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
16	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, no estudiaron su especialidad en Facultad de Medicina, sí comentan, no estudiaron medicina general en Universidad de San Carlos, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estiman como motivo para solicitar autopsia la orientación diagnóstica, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron medicina general en Hospital General Dr. Manuel Gea González, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron medicina general en Instituto Nacional de Cancerología, no piensan que una causa o motivo para no solicitar autopsia es por los trámites administrativos, no estudiaron su especialidad en Universidad Popular Autónoma, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron medicina general en Hospital General, no estudiaron medicina general en Universidad Villa Rica, no opinan que se debe obligar a los residentes, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron su especialidad en Universidad de San Carlos, y no estudiaron medicina general en la Facultad de Ciencias Médicas; entonces son del Hospital General de Zona 53.

4.10.3 Análisis Visual

El análisis visual permite evaluar de manera gráfica el comportamiento de las métricas asociadas a las reglas generadas mediante la función de calidad *Klosgen*, facilitando la identificación de patrones dominantes, variabilidad entre reglas y diferencias estructurales entre hospitales. A continuación, se presentan las gráficas correspondientes a soporte, soporte negativo, diferencia de soporte, precisión, diferencia de precisión y redundancia, las cuales complementan la información numérica presentada en las tablas previas.

La Figura 4.64 muestra la distribución de los valores de soporte para las reglas generadas mediante la función de calidad *Klosgen*. En la gráfica se observa una amplia variabilidad en la cobertura de los patrones, con valores que van desde niveles bajos hasta valores máximos cercanos a 1.0. Este comportamiento indica la coexistencia de reglas altamente representativas, que describen grandes subconjuntos del hospital

objetivo, junto con reglas más específicas que capturan perfiles particulares dentro de cada institución.

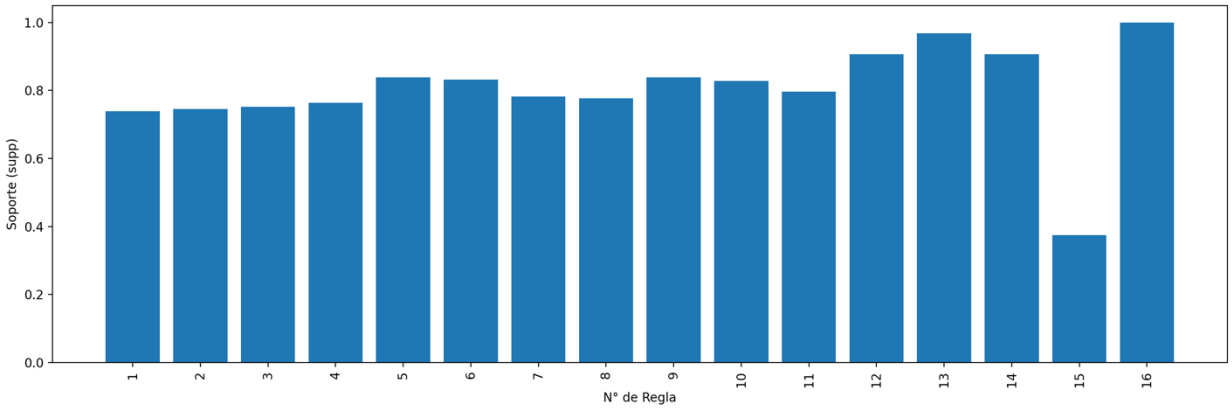


Figura 4.64 *Klosgen* – Matriz Binaria: Soporte por Regla

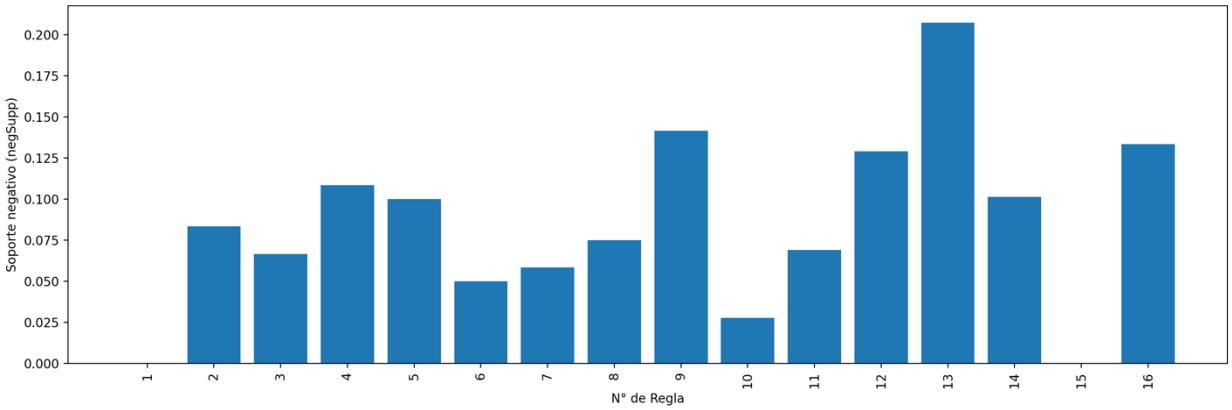


Figura 4.65 *Klosgen* – Matriz Binaria: Soporte Negativo por Regla

La Figura 4.65 presenta el comportamiento del soporte negativo asociado a cada regla. Se observa que varias reglas presentan valores de soporte negativo distintos de cero, lo que indica la existencia de contraejemplos en hospitales distintos al hospital objetivo. Este comportamiento explica la variabilidad observada posteriormente en la precisión de las reglas y confirma que *Klosgen* no restringe el descubrimiento de patrones exclusivamente a combinaciones totalmente exclusivas.

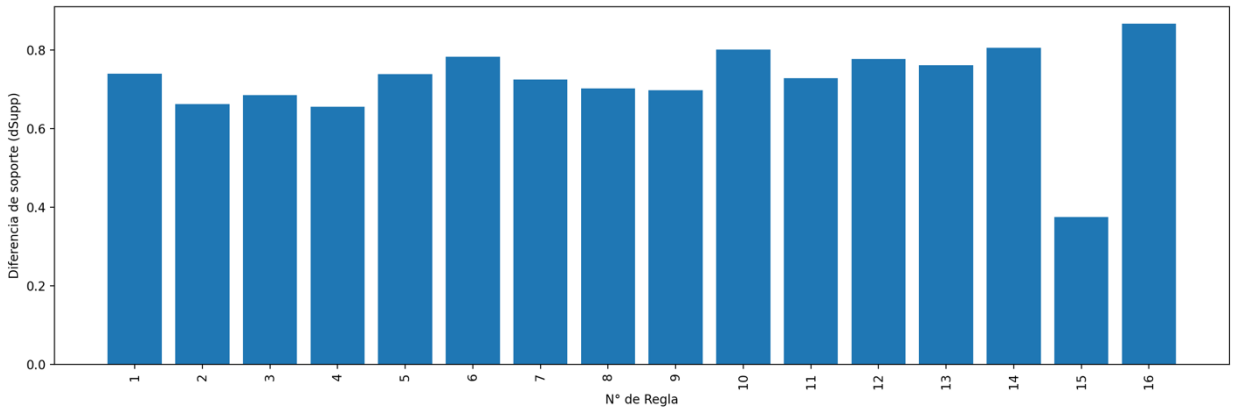


Figura 4.66 *Klosgen* – Matriz Binaria: Diferencia de Soporte por Regla

La Figura 4.66 muestra la diferencia de soporte para las reglas generadas. En general, los valores se mantienen mayoritariamente positivos, lo que indica que los patrones identificados aparecen con mayor frecuencia en el hospital objetivo que en el resto de la población. No obstante, se aprecia una dispersión mayor en comparación con otras métricas, lo que refleja diferencias en el grado de contrastividad de los patrones descubiertos.

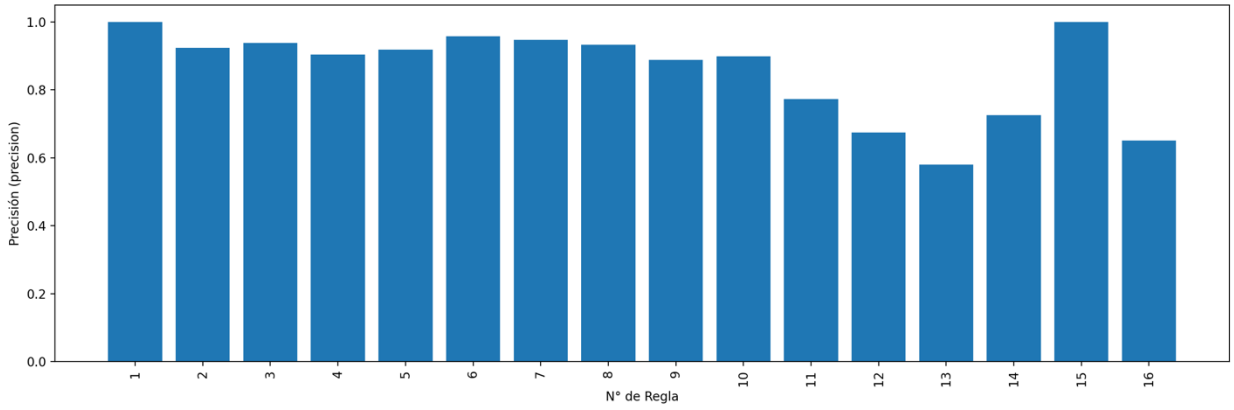


Figura 4.67 *Klosgen* – Matriz Binaria: Precisión por Regla

La Figura 4.67 presenta los valores de precisión correspondientes a cada regla. En la gráfica se observa un comportamiento heterogéneo, con reglas que alcanzan valores de precisión elevados y otras que muestran descensos moderados. Este comportamiento

confirma que, si bien *Klosgen* identifica patrones frecuentes, no todos los patrones mantienen un nivel uniforme de capacidad discriminativa entre hospitales.

La Figura 4.68 muestra la diferencia de precisión asociada a las reglas generadas. Se identifican reglas con ganancias claras de precisión respecto al comportamiento global del conjunto de datos, así como reglas con valores menores a 0.17. Este resultado evidencia que *Klosgen* prioriza el equilibrio entre frecuencia y discriminación, lo que puede derivar en patrones representativos, pero con una ganancia discriminativa limitada.

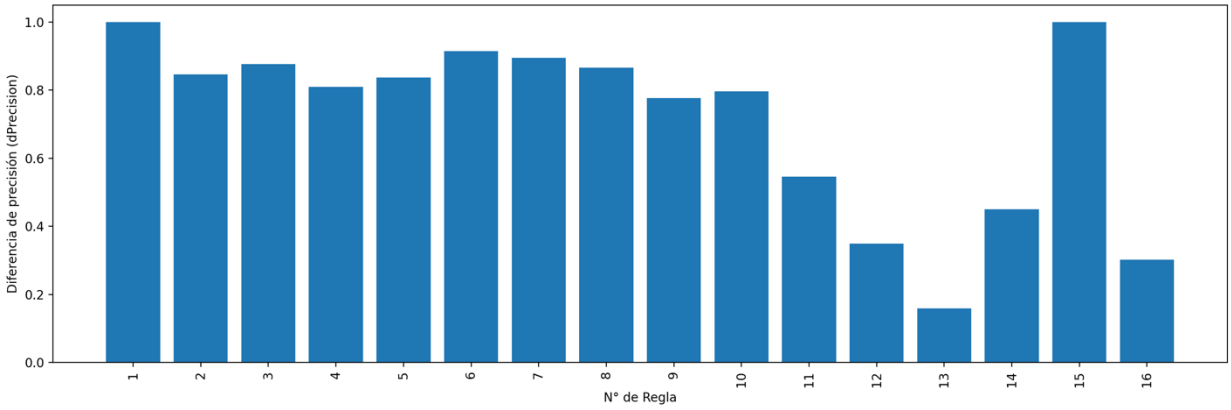


Figura 4.68 *Klosgen* – Matriz Binaria: Diferencia de Precisión por Regla

La Figura 4.69 presenta la redundancia de las reglas descubiertas mediante *Klosgen*. En la gráfica se observa la presencia de valores elevados de redundancia en un subconjunto importante de reglas, particularmente asociadas a ciertos hospitales. Este comportamiento indica la existencia de regiones densas del espacio de búsqueda, donde múltiples reglas describen combinaciones similares de atributos, reforzando la necesidad de un proceso posterior de selección de patrones para reducir el solapamiento informativo.

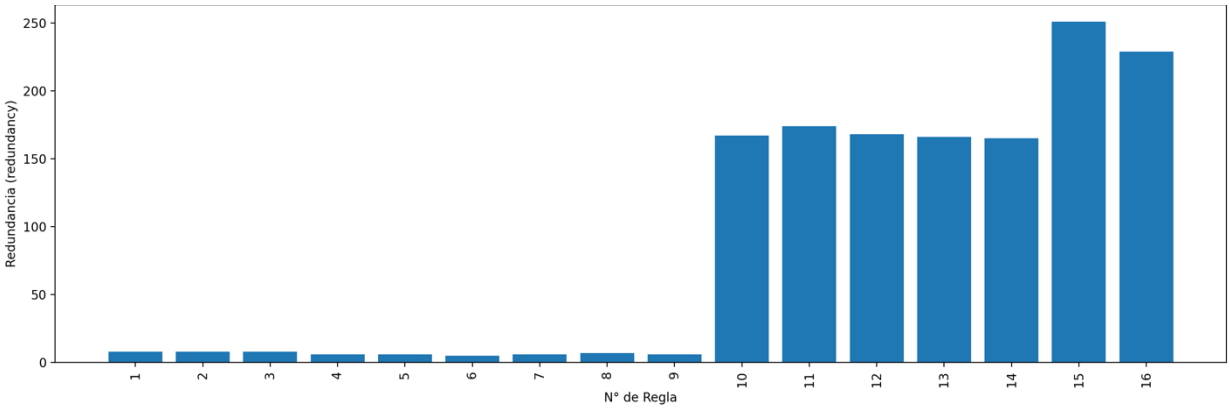


Figura 4.69 *Klosgen* – Matriz Binaria: Redundancia por Regla

4.10.4 Evaluación objetiva

Desde una perspectiva objetiva, el modelo generado mediante *Klosgen* muestra un comportamiento consistente pero heterogéneo. Si bien la mayoría de las reglas presentan diferencias de soporte positivas, la variabilidad en precisión y diferencia de precisión indica que no todas las reglas aportan el mismo nivel de capacidad discriminativa. *Klosgen* favorece reglas con alta cobertura, pero esto implica que algunos patrones, aun siendo frecuentes, no maximizan la discriminación entre hospitales.

La redundancia elevada en ciertos bloques constituye el principal elemento crítico del modelo, ya que refleja un solapamiento estructural importante. No obstante, esta característica también confirma la estabilidad de ciertos perfiles institucionales, especialmente en hospitales con respuestas homogéneas.

4.10.5 Evaluación subjetiva

La evaluación subjetiva de las reglas generadas mediante *Klosgen* fue completamente favorable. Desde la perspectiva del Dr. Palet, todos los patrones identificados presentan coherencia clínica y describen perfiles plausibles del personal médico que labora en el Hospital Regional Río Blanco, el Hospital General de San Juan Bautista Tuxtepec y el Hospital General de Zona 53.

Las reglas integran de manera equilibrada elementos de formación médica en escuelas públicas, trayectoria profesional y percepciones institucionales, reflejando tanto perfiles homogéneos como la diversidad real presente en algunos hospitales. Aunque ciertas reglas muestran una amplitud mayor en su estructura, el especialista considera que esta característica no debilita su valor interpretativo, sino que refleja adecuadamente la complejidad del contexto clínico. En este sentido, *Klosgen* destaca por generar patrones estadísticamente válidos con un alto nivel de aceptación clínica.

4.11 Matriz Binaria: Resultados aplicando la función de calidad *RelativeRisk*

La función de calidad *RelativeRisk* evalúa la fuerza de asociación entre la presencia de un patrón y la pertenencia a un grupo objetivo, comparando la probabilidad de ocurrencia del antecedente dentro del grupo de interés frente a su ocurrencia fuera de él. A diferencia de métricas centradas en la ganancia absoluta, *RelativeRisk* enfatiza el riesgo relativo, permitiendo identificar patrones cuya presencia incrementa significativamente la probabilidad de pertenecer a un hospital específico.

En este apartado se presentan los resultados del análisis de Minería de Conjuntos de Contraste aplicado al conjunto de datos Matriz Binaria, en nivel específico, utilizando la función de calidad *RelativeRisk*. La estructura del análisis mantiene el mismo estándar metodológico seguido en los apartados previos, incluyendo la presentación de reglas sin interpretar, su interpretación semántica, el análisis visual de métricas y la evaluación objetiva y subjetiva del modelo generado.

4.11.1 Reglas sin interpretar

La Tabla 4.19 presenta el conjunto de reglas sin interpretar generadas mediante la función de calidad *RelativeRisk*, conformado por veintitrés patrones distribuidos entre los tres hospitales analizados. En términos generales, las reglas muestran soportes elevados y diferencias de soporte positivas, acompañadas de valores de precisión notablemente altos, lo que evidencia un modelo con fuerte capacidad discriminativa.

La regla asociada al Hospital Regional Río Blanco se caracteriza por un valor de soporte que alcanza el nivel máximo, lo que indica que el patrón identificado cubre prácticamente la totalidad del hospital objetivo. La diferencia de soporte se mantiene positiva y elevada, mientras que la precisión muestra valores altos. La redundancia es relativamente baja en comparación con otros hospitales, lo que sugiere un patrón más compacto y menos solapado.
La regla vinculada al Hospital General de San Juan Bautista Tuxtepec presenta 0.7, con diferencia de soporte pronunciada y valores de precisión cercanos a la 0.7. La diferencia de precisión es marcadamente positiva, reflejando una ganancia clara frente al comportamiento global. Sin embargo, esta regla exhibe un nivel alto de redundancia, lo que indica que múltiples combinaciones de atributos producen patrones equivalentes, evidenciando una estructura institucional altamente homogénea.

Finalmente, la regla correspondiente al Hospital General de Zona 53 tiene soporte elevado y diferenciade soporte consistentemente positivas. La precisión es más alta que en los otros hospitales. La redundancia vuelve a ser elevada, lo que confirma la presencia de regiones densas del espacio de búsqueda donde múltiples reglas describen perfiles similares.

En conjunto, la Tabla 4.19 evidencia que *RelativeRisk* privilegia patrones altamente representativos y con fuerte asociación al hospital objetivo, aun cuando esto implique un incremento significativo en la redundancia.

Tabla 4.19. *RelativeRisk* – Matriz Binaria: Reglas Sin Interpretar

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
1	"IF p17_107 = !{S} AND p16_112 = !{S} AND p16_131 = !{S} AND p17_103 = !{S} AND p16_129 = !{S} AND p17_106 = !{S} AND p17_105 = !{S} AND p17_111 = !{S} AND p16_121 = !{S} AND p17_100 = !{S} AND p17_101 = !{S} AND p17_102 = !{S} AND p17_104 = !{S} AND p16_108 = !{S} AND p17_109 = !{S} AND p17_110 = !{S} AND p17_118 = !{S} AND p17_123 = !{S} AND p16_125 = !{S} AND p17_130 = !{S} AND p17_131 = !{S} AND p16_115 = !{S} AND p17_114 = !{S} AND p17_119 = !{S} AND p17_124 = !{S} AND p17_126 = !{S} THEN centro_hospitalario = {H1}"	1.000000	0.241667	0.758333	0.847368	0.694737	5.000000
2	"IF p17_10 = !{S} AND p17_115 = !{S} AND p11_6 = !{S} AND p16_27 = !{S} AND p7_9 = !{S} AND p10_5 = !{S} AND p17_13 = !{S} AND p6_4 = !{S} AND p17_8 = !{S} AND p16_129 = !{S} AND p16_9 = !{S} AND p17_2 = !{S} AND p6_5 = !{S} AND p16_41 = !{S} AND p17_125 = !{S} AND p16_16 = !{S} AND p16_17 = !{S} AND p16_47 = !{S} AND p16_33 = !{S} AND p16_35 = !{S} AND p6_8 = !{S} AND p16_1 = !{S} AND p16_131 = !{S} AND p16_11 = !{S} AND p13_4 = !{S} AND p17_3 = !{S} AND p17_16 = !{S} AND p7_13 = !{S} AND p7_14 = !{S} AND p3_5 = !{S} AND p16_12 = !{S} AND p17_43 = !{S} AND p17_22 = !{S} AND p13_13 = !{S} AND p13_16 = !{S} AND p13_18 = !{S} AND p17_42 = !{S} AND p17_82 = !{S} AND p16_95 = !{S} AND p16_97 = !{S} AND p17_117 = !{S} AND p17_118 = !{S} AND p17_123	1.000000	0.156682	0.843318	0.653061	0.306122	165.000000

Tabla 4.19. *RelativeRisk* – Matriz Binaria: Reglas Sin Interpretar - Continuación

N.°	Regla	supp	negSupp	dSupp	precision	dPrecision	redundancy
	= !{S} AND p17_128 = !{S} AND p17_129 = !{S} AND p17_130 = !{S} AND p17_133 = !{S} AND p17_28 = !{S} AND p17_29 = !{S} AND p17_31 = !{S} AND p17_7 = !{S} AND p17_17 = !{S} AND p17_32 = !{S} AND p17_51 = !{S} AND p17_85 = !{S} AND p17_88 = !{S} AND p17_71 = !{S} AND p17_119 = !{S} AND p17_122 = !{S} AND p17_124 = !{S} AND p17_126 = !{S} AND p17_131 = !{S} THEN centro_hospitalario = {H11}"						
3	"IF p16_8 = !{S} AND p13_1 = !{S} AND p12_8 = !{S} AND p17_10 = !{S} AND p16_13 = !{S} AND p16_99 = !{S} AND p16_93 = !{S} AND p16_9 = !{S} AND p7_7 = !{S} AND p16_33 = !{S} AND p7_17 = !{S} AND p17_105 = !{S} AND p17_106 = !{S} AND p16_11 = !{S} AND p16_27 = !{S} AND p17_16 = !{S} AND p16_35 = !{S} AND p16_12 = !{S} AND p17_22 = !{S} AND p13_15 = !{S} AND p16_1 = !{S} AND p16_41 = !{S} AND p17_13 = !{S} AND p16_89 = !{S} AND p16_95 = !{S} AND p17_109 = !{S} AND p16_16 = !{S} AND p17_18 = !{S} AND p17_21 = !{S} AND p13_17 = !{S} AND p17_1 = !{S} AND p17_82 = !{S} AND p17_84 = !{S} AND p17_85 = !{S} AND p17_88 = !{S} AND p16_97 = !{S} AND p17_71 = !{S} AND p17_114 = !{S} THEN centro_hospitalario = {H12}"	1.000000	0.017778	0.982222	0.933333	0.866667	229.000000

4.11.2 Reglas interpretadas

La Tabla 4.20 presenta la interpretación semántica de las reglas generadas mediante *RelativeRisk*, permitiendo traducir los patrones numéricos a descripciones comprensibles desde el punto de vista clínico e institucional.

En el Hospital Regional Río Blanco, las reglas interpretadas describen perfiles clínicos claramente definidos, relacionados con la formación académica, la experiencia profesional y percepciones consistentes sobre la utilidad y los resultados de las autopsias. La elevada cobertura de estas reglas refuerza su carácter representativo del hospital.

Las reglas interpretadas correspondientes al Hospital General de San Juan Bautista Tuxtepec muestran perfiles altamente homogéneos, asociados a posturas institucionales bien establecidas respecto a los motivos para solicitar o rechazar autopsias. La consistencia semántica de estas reglas es coherente con los valores elevados de precisión y diferencia de precisión observados.

Por su parte, el Hospital General de Zona 53 presenta reglas interpretadas con estructuras más amplias, que integran múltiples dimensiones formativas y perceptuales. Aunque algunas reglas muestran menor exclusividad, su cobertura elevada indica que describen perfiles dominantes dentro del hospital.

En conjunto, la Tabla 4.20 permite vincular los resultados cuantitativos de *RelativeRisk* con interpretaciones semánticas coherentes, facilitando la comprensión de los perfiles institucionales identificados.

Tabla 4.20. *Klosgen* – Matriz Binaria: Reglas Interpretadas

N.°	Regla Interpretada
1	Si los médicos sí especifican dónde realizaron su especialidad, no estudiaron medicina general en Universidad del Valle de México, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, no estudiaron su especialidad en IMSS, no estudiaron medicina

Tabla 4.20. *Klosgen* – Matriz Binaria: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
	general en Hospital Juárez de México, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron su especialidad en Hospital General de Zona 53, no estudiaron medicina general en Centro Interdisciplinario de Ciencias de la Salud, no estudiaron su especialidad en Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, no estudiaron su especialidad en Hospital Ángeles, no estudiaron su especialidad en Universidad de Guanajuato, no estudiaron medicina general en Instituto Politécnico Nacional, no estudiaron su especialidad en Universidad de Nuevo León, no estudiaron su especialidad en Escuela Militar de Medicina, no estudiaron su especialidad en Hospital General Palenque, no estudiaron medicina general en Universidad Autónoma de México, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza, no estudiaron medicina general en Hospital Hermanos Ameijeiras, no estudiaron su especialidad en Estudios de Posgrado de Excelencia, no estudiaron su especialidad en Hospital General La Villa Urgencias, y no estudiaron su especialidad en Hospital General; entonces son del Hospital Regional Río Blanco.
2	Si los médicos no estudiaron su especialidad en Facultad de Medicina, no estudiaron su especialidad en Hospital Hermanos Ameijeiras, no estiman como motivo para solicitar autopsia la orientación diagnóstica, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no piensan que una causa o motivo para no solicitar autopsia es por cuestiones religiosas, no consideran que quien debe solicitar la autopsia es un familiar, no estudiaron su especialidad en Universidad de San Carlos, no plantean como motivo para solicitar autopsia la aplicabilidad del estudio, no estudiaron su especialidad en Alta Especialidad en IMSS, no estudiaron medicina general en Hospital Juárez de México, no estudiaron medicina general en Hospital del Departamento del D.F., no estudiaron su especialidad en Nachón/CEM, no plantean como motivo para solicitar autopsia cuando existe falta de correlación entre los datos clínicos y los de laboratorio, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron su especialidad en Universidad Autónoma de México, no estudiaron medicina general en Universidad Popular Autónoma, no estudiaron medicina general en Universidad Benito Juárez, no estudiaron medicina general en Hospital General Regional IMSS, no estudiaron medicina general en Regional de Occidente, no estudiaron medicina general en San Nicolás de Hidalgo, no plantean como motivo para solicitar autopsia problemas de atención médica, no estudiaron medicina general en Universidad Veracruzana, no realizaron sus estudios de medicina general en la Facultad de Estudios Superiores Zaragoza, no estudiaron medicina general en Instituto Nacional de Cancerología, no opinan que la encuesta es útil, no estudiaron su especialidad en Hospital del Departamento del D.F., no estudiaron su especialidad en Universidad Popular

Tabla 4.20. *Klosgen* – Matriz Binaria: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
	Autónoma, no piensan que una causa o motivo para no solicitar autopsia es la condición del finado, no piensan que una causa o motivo para no solicitar autopsia es la falta de recursos humanos, no están totalmente en desacuerdo en que los hallazgos de las autopsias pueden ser discrepantes con los diagnósticos clínicos, no estudiaron medicina general en Universidad Villa Rica, no estudiaron su especialidad en Instituto de Salud Pública, no estudiaron su especialidad en Secretaría de Salud, no opinan que se debe incrementar el número de autopsias, no opinan que nunca le han realizado autopsias a sus pacientes, no opinan agregar patología a la rotación de los médicos internos, no estudiaron su especialidad en Centro Médico de Occidente, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron medicina general en Universidad Autónoma Benito Juárez de Oaxaca, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron su especialidad en Universidad Autónoma Metropolitana, no estudiaron su especialidad en Hospital General Palenque, sí especifican dónde realizaron su especialidad, no estudiaron su especialidad en Centro Médico Nacional La Raza, no estudiaron su especialidad en Hospital Juárez de México, no estudiaron su especialidad en Hospital General Dr. Miguel Silva, no estudiaron su especialidad en Universidad Autónoma, no estudiaron su especialidad en Centro de Especialidades Médicas, no estudiaron su especialidad en Servicios de Salud, no estudiaron su especialidad en Universidad Benito Juárez, no estudiaron su especialidad en Hospital de la Mujer, no estudiaron su especialidad en Hospital Clínico Quirúrgico Hermanos Ameijeiras, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no estudiaron su especialidad en Hospital General Dr. Manuel Gea González, no estudiaron su especialidad en Hospital general Miguel Enríquez, no estudiaron su especialidad en Hospital General La Villa Urgencias, no estudiaron su especialidad en Hospital general, y no estudiaron su especialidad en Facultad de Estudios Superiores Zaragoza; entonces son del Hospital General de San Juan Bautista Tuxtepec.
3	Si los médicos no estudiaron medicina general en Alta Especialidad en IMSS, sí comentan, no están de acuerdo en que el método eficiente para solicitar autopsia es a través del área de psicología, no estudiaron su especialidad en Facultad de Medicina, no estudiaron medicina general en Universidad de San Carlos, no estudiaron medicina general en Secretaría de Salud de Oaxaca, no estudiaron medicina general en Universidad Anáhuac, no estudiaron medicina general en Hospital del Departamento del D.F., no piensan que una causa o motivo para no solicitar autopsia es por los trámites administrativos, no estudiaron medicina general en Regional de Occidente, no piensan que una causa o motivo para no solicitar autopsia es que se haga sin fines de enseñanza, no estudiaron su especialidad en Escuela de Salud Pública de México, no estudiaron su especialidad en Hospital General Balbuena, no estudiaron medicina general en

Tabla 4.20. *Klosgen* – Matriz Binaria: Reglas Interpretadas – Continuación

N.°	Regla Interpretada
	Instituto Nacional de Cancerología, no estudiaron medicina general en UMAC 14 Adolfo Ruiz Cortines, no estudiaron su especialidad en Universidad Popular Autónoma, no estudiaron medicina general en San Nicolás de Hidalgo, no estudiaron medicina general en Universidad Villa Rica, no estudiaron su especialidad en Secretaría de Salud, no opinan que se debe obligar a los residentes, no estudiaron medicina general en Universidad Veracruzana, no estudiaron medicina general en Universidad Cristóbal Colón, no estudiaron su especialidad en Universidad de San Carlos, no estudiaron medicina general en Universidad Regional del Sureste, no estudiaron medicina general en Universidad Autónoma Benito Juárez de Oaxaca, no estudiaron su especialidad en Universidad de Nuevo León, no estudiaron medicina general en Universidad Popular Autónoma, no estudiaron su especialidad en Universidad Juárez Autónoma, no estudiaron su especialidad en IMSS Adolfo Ruiz Cortines, no opinan que se debe solicitar el consentimiento, no estudiaron su especialidad en Universidad Veracruzana, no estudiaron su especialidad en Benemérita Universidad Autónoma de Puebla, no estudiaron su especialidad en Universidad de Guadalajara, no estudiaron su especialidad en Universidad Autónoma de Tlaxcala, no estudiaron su especialidad en Hospital General Dr. Manuel Gea González, no realizaron su formación de medicina general en Instituto de Ciencias y Estudios Superiores de Tamaulipas, no estudiaron su especialidad en Hospital general Miguel Enríquez, y no estudiaron su especialidad en Estudios de Posgrado de Excelencia; entonces son del Hospital General de Zona 53.

4.11.3 Análisis visual

Dado que la función de calidad *RelativeRisk* generó únicamente tres reglas en el análisis específico sobre la Matriz Binaria, el análisis visual se centra en la comparación directa del comportamiento de cada regla individual. Esta representación permite evaluar con claridad la estabilidad, cobertura y capacidad discriminativa de cada patrón, sin la necesidad de agruparlos en bloques más amplios como en los casos anteriores.

La Figura 4.70 muestra los valores de soporte correspondientes a las tres reglas generadas. Se observa que la Regla 1, la Regla 2 y la Regla 3 presentan valores máximos de soporte, lo que indica que cada uno de los patrones cubre la totalidad o casi la totalidad del hospital objetivo. Este comportamiento confirma que *RelativeRisk* prioriza reglas altamente representativas del grupo analizado.

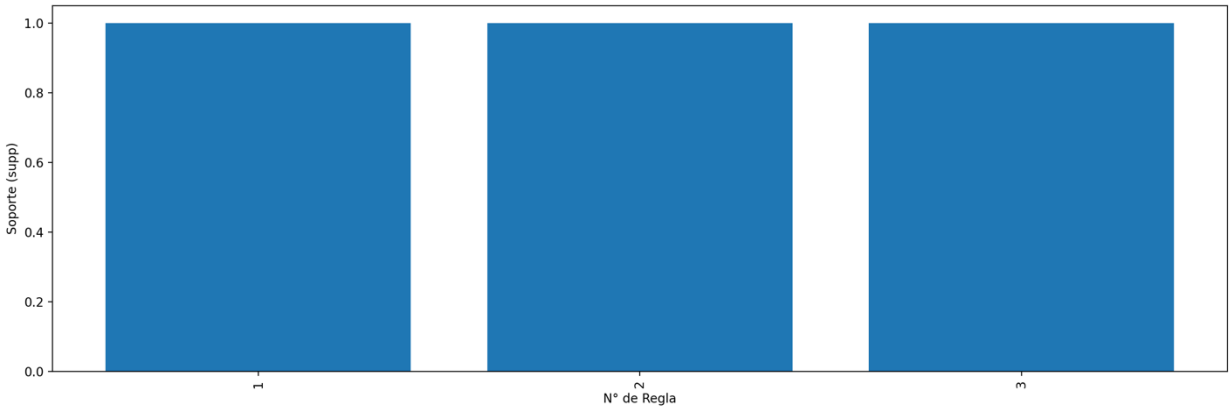


Figura 4.70 *RelativeRisk* – Matriz Binaria: Soporte por Regla

En la Figura 4.71 se presenta el soporte negativo de las tres reglas. La Regla 1 muestra un valor moderado de soporte negativo, lo que indica la presencia de contraejemplos en otros hospitales. En contraste, la Regla 2 y la Regla 3 presentan valores notablemente menores, evidenciando un mayor grado de exclusividad respecto al hospital objetivo. Esta diferencia explica variaciones posteriores en los valores de precisión.

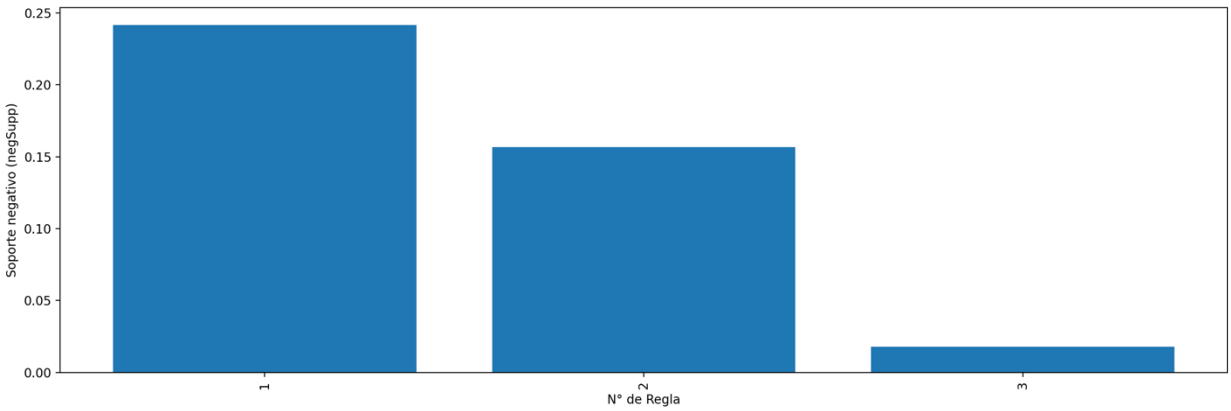


Figura 4.71 *RelativeRisk* – Matriz Binaria: Soporte Negativo por Regla

La Figura 4.72 evidencia que las tres reglas presentan una diferencia de soporte claramente positiva, siendo la Regla 2 y la Regla 3 las que alcanzan los valores más elevados. Este comportamiento confirma que los patrones identificados aparecen con

una frecuencia considerablemente mayor en el hospital objetivo que en el resto de la población.

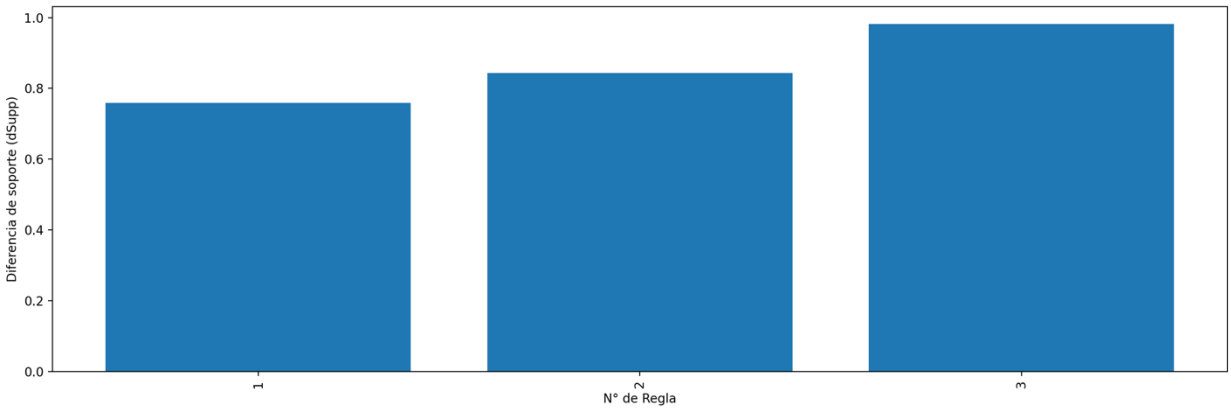


Figura 4.72 *RelativeRisk* – Matriz Binaria: Diferencia de Soporte por Regla

La Figura 4.73 muestra los valores de precisión asociados a cada regla. La Regla 1 y la Regla 3 alcanzan valores de precisión cercanos a la unidad, reflejando una alta pureza de los patrones. La Regla 2, aunque mantiene una precisión elevada, presenta un valor ligeramente inferior, consistente con su alto soporte negativo observado previamente.

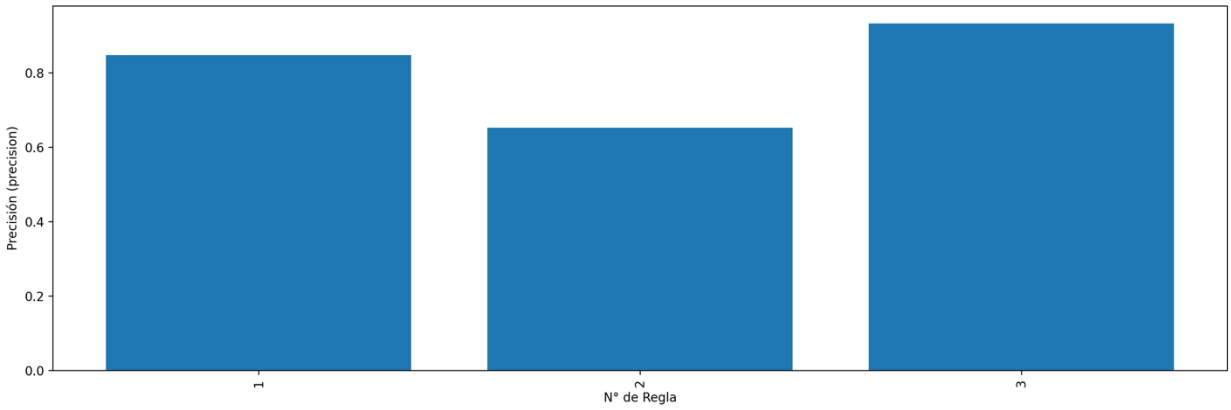


Figura 4.73 *RelativeRisk* – Matriz Binaria: Precisión por Regla

En la Figura 4.74 se observa que las tres reglas presentan diferencias de precisión positivas, destacando nuevamente la Regla 1 y la Regla 3 con las mayores ganancias. Este resultado confirma que todas las reglas aportan una mejora sustancial respecto al comportamiento global del conjunto de datos, aunque con distintos niveles de intensidad.

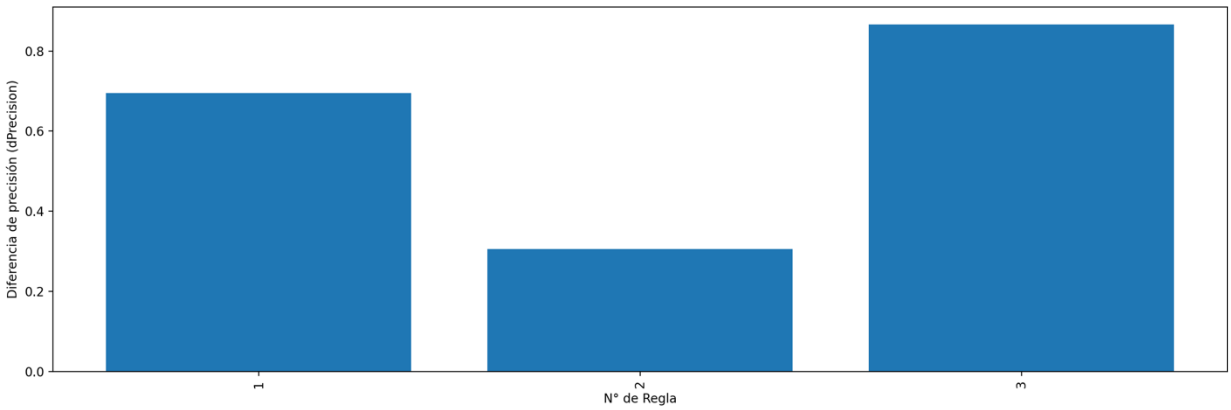


Figura 4.74 *RelativeRisk* – Matriz Binaria: Diferencia de Precisión por Regla

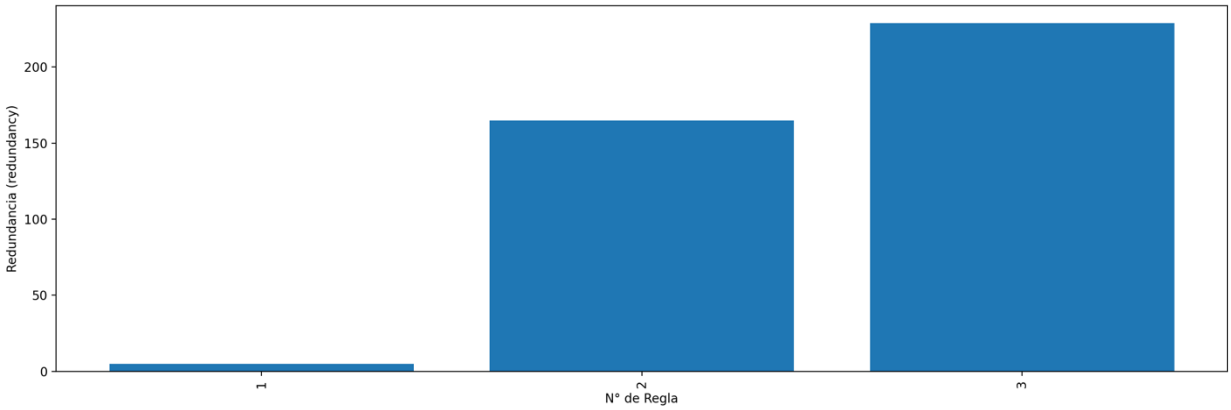


Figura 4.75 *RelativeRisk* – Matriz Binaria: Redundancia por Regla

La Figura 4.75 muestra la redundancia asociada a cada regla. Se observa que las tres reglas presentan valores elevados de redundancia, particularmente la Regla 2 y la Regla 3, lo que indica que cada patrón pertenece a regiones densas del espacio de búsqueda. Este comportamiento sugiere que, aunque las reglas son altamente representativas, describen perfiles institucionales que podrían expresarse mediante múltiples combinaciones equivalentes de atributos.

4.11.4 Evaluación objetiva

Desde una perspectiva objetiva, *RelativeRisk* produce un modelo altamente discriminativo, caracterizado por valores elevados de soporte, diferencia de soporte y

precisión. Todas las reglas muestran una ganancia clara respecto al comportamiento global, lo que confirma la estabilidad y consistencia estadística del modelo.

No obstante, la elevada redundancia observada constituye el principal elemento crítico del análisis, ya que implica un solapamiento informativo considerable. Este comportamiento sugiere que, aunque *RelativeRisk* identifica patrones muy representativos, resulta necesario aplicar criterios adicionales de selección para reducir la redundancia y priorizar reglas con mayor valor informativo.

4.11.5 Evaluación subjetiva

En el análisis realizado mediante *RelativeRisk*, la evaluación subjetiva fue completamente positiva. Las reglas generadas describen perfiles clínicos altamente representativos y coherentes con la práctica observada en los hospitales analizados, incorporando de forma clara la formación médica en escuelas públicas, la experiencia profesional y las percepciones institucionales sobre la práctica de la autopsia.

Aunque el número de reglas obtenidas es reducido, el Dr. Palet considera que cada patrón posee una alta claridad interpretativa y un fuerte anclaje en la realidad clínica, lo que compensa ampliamente la menor diversidad de reglas. Desde esta perspectiva, *RelativeRisk* se distingue por producir patrones de fácil interpretación y alto valor institucional.

4.12 Análisis de Resultados

La evaluación del desempeño de las funciones de calidad implementadas en el módulo de minería de conjuntos de contraste *Lift, WRAcc, Correlation, Klosgen* y *RelativeRisk* permite establecer un juicio crítico sobre su idoneidad para la identificación de patrones discriminativos en el contexto biomédico y hospitalario. Si bien todas las métricas lograron generar reglas que satisfacen los requisitos mínimos de consistencia estadística (diferencia de soporte positiva), el análisis comparativo revela discrepancias

significativas en cuanto a la precisión, la interpretabilidad clínica y, fundamentalmente, la validez subjetiva otorgada por el experto.

4.12.1 Comparativa Técnica y Estadística

Desde una perspectiva estrictamente cuantitativa, la comparación entre las funciones de calidad aplicadas en los experimentos específicos evidencia diferencias sustanciales en el comportamiento estadístico de los modelos generados, las cuales están directamente relacionadas tanto con la naturaleza de cada métrica como con el esquema de representación de los datos empleado. La Tabla 4.21 sintetiza de manera comparativa los principales indicadores estadísticos obtenidos para cada función de calidad, permitiendo una evaluación objetiva de su capacidad discriminativa, estabilidad y complejidad estructural.

En el experimento basado en la Vista Minable, la función de calidad *Lift* mostró un desempeño destacado en términos de discriminación, al generar un conjunto de dieciocho reglas con valores de precisión elevados y diferencias de precisión consistentemente positivas. Este comportamiento refleja su capacidad para identificar asociaciones fuertes y exclusivas entre los antecedentes y los hospitales objetivo, privilegiando subgrupos con identidad institucional bien definida. En contraste, *RelativeRisk* produjo un modelo extremadamente compacto, limitado a tres reglas con soporte máximo, pero acompañado de soportes negativos elevados y diferencias de precisión negativas en dos de los tres hospitales, lo que evidencia una penalización excesiva en contextos donde los perfiles institucionales comparten características comunes.

Por su parte, *WRAcc* y *Correlation* generaron modelos con mayor cobertura global, reflejada en valores de soporte elevados, pero también incorporaron reglas con soporte negativo significativo. Esta característica, si bien garantiza estabilidad estadística, deriva en patrones más amplios y menos específicos, lo que reduce su capacidad para caracterizar con nitidez las diferencias operativas y formativas entre instituciones como el Hospital Regional Río Blanco y el Hospital General de Zona 53. De manera similar,

Klosgen, aunque produjo reglas con altos valores de soporte y diferencias de soporte positivas, presentó índices de redundancia superiores, lo que sugiere una mayor complejidad combinatoria sin una ganancia proporcional en claridad interpretativa.

Al incorporar el experimento específico basado en la Matriz Binaria, se observa un cambio sustancial en el perfil estadístico de las funciones de calidad. La binarización de los atributos incrementó la homogeneidad estructural de los patrones, reduciendo de forma significativa el soporte negativo en la mayoría de las métricas y favoreciendo la generación de reglas con mayor pureza estadística. Este efecto es particularmente evidente en *Lift*, donde todas las reglas alcanzan precisión y diferencia de precisión máximas, acompañadas de un soporte negativo nulo. No obstante, este comportamiento se presenta junto con valores de soporte promedio más bajos, lo que indica la identificación de subgrupos altamente específicos con cobertura limitada.

En el caso de *WRAcc* y *Correlation*, la Matriz Binaria permitió conservar un equilibrio notable entre cobertura y capacidad discriminativa, con diferencias de soporte superiores a 0.60 y precisiones cercanas a 0.95, acompañadas de diferencias de precisión ampliamente positivas. Estos resultados sugieren que ambas métricas se benefician de la regularización introducida por la binarización, logrando patrones estadísticamente estables sin una pérdida significativa de poder discriminativo. *Klosgen* mantuvo un comportamiento robusto, con altos valores de soporte y diferencias de soporte positivas, aunque con un número reducido de reglas y una tendencia a concentrarse en patrones dominantes del espacio de búsqueda.

Finalmente, *RelativeRisk* mostró una mejora cuantitativa relevante bajo la representación binaria. A diferencia de lo observado en la Vista Minable, la reducción de la redundancia y el incremento en la precisión promedio indican que la Matriz Binaria atenúa la penalización excesiva característica de esta métrica, permitiendo la identificación de patrones más equilibrados sin perder su naturaleza compacta.

En conjunto, los resultados presentados en la Tabla 4.21 confirman que el desempeño de las funciones de calidad no es independiente del esquema de representación de los datos. Mientras la Vista Minable maximiza la riqueza semántica y la interpretabilidad clínica de los patrones, la Matriz Binaria refuerza la estabilidad estadística y la pureza estructural de las reglas generadas. Esta comparativa técnica y estadística valida la solidez metodológica del módulo propuesto, al demostrar que los patrones contrastivos identificados se mantienen consistentes bajo diferentes representaciones, y que las diferencias observadas responden a estructuras reales del dominio analizado y no a artefactos del proceso de codificación.

Tabla 4.21. Vista Minable: Tabla Comparativa de Funciones de Calidad

Representación	Función de Calidad	Nº de Reglas	supp	negSupp	dSupp	precision	dPrecision	redundancy
Vista Minable	Lift	18	0.45507	0.06013	0.39494	0.86144	0.72289	17,422.72
Vista Minable	WRAcc	18	0.68054	0.06543	0.61001	0.84161	0.68969	21,145.17
Vista Minable	Correlation	18	0.68054	0.06543	0.61001	0.84161	0.68969	21,145.17
Vista Minable	Klosgen	20	0.78652	0.10990	0.68700	0.77403	0.48779	22,336.40
Vista Minable	RelativeRisk	3	1.00000	0.42879	0.57121	0.48386	-0.03227	17,187.00
Matriz Binaria	Lift	53	0.15151	0.00000	0.15151	1.00000	1.00000	90.38
Matriz Binaria	WRAcc	22	0.64536	0.03500	0.61036	0.94860	0.89721	145.14
Matriz Binaria	Correlation	22	0.64536	0.03500	0.61036	0.94860	0.89721	145.14
Matriz Binaria	Klosgen	16	0.80310	0.08445	0.71865	0.85690	0.71380	86.25
Matriz Binaria	RelativeRisk	3	1.00000	0.13871	0.86129	0.81125	0.62251	133.00

4.12.2 Evaluación Subjetiva y Criterio Experto

La evaluación subjetiva, realizada por el especialista clínico, constituye un componente fundamental para complementar el análisis técnico y estadístico presentado en el apartado anterior. Mientras que las métricas cuantitativas permiten medir la estabilidad, cobertura y capacidad discriminativa de los modelos generados, el criterio experto aporta una validación semántica y clínica indispensable para determinar la pertinencia real de los patrones descubiertos. En este sentido, la evaluación se llevó a cabo considerando de manera conjunta los resultados obtenidos tanto en la Vista Minable como en la Matriz Binaria, con el objetivo de analizar la consistencia interpretativa de las reglas bajo distintos esquemas de representación.

En el caso de la Vista Minable, el especialista identificó diferencias claras entre la solidez estadística de las reglas y su valor clínico. Algunas reglas, particularmente aquellas generadas por métricas que privilegian el soporte absoluto, fueron consideradas poco representativas desde el punto de vista clínico, aun cuando cumplieran con los criterios formales de la función de calidad. Este fenómeno fue especialmente evidente en patrones con soporte máximo, pero con precisiones moderadas o bajas, los cuales, si bien reflejan características ampliamente compartidas entre hospitales, no describen perfiles institucionales clínicamente diferenciables. Estas reglas fueron descartadas por el experto al no aportar conocimiento accionable ni coherente con la experiencia observada en los centros hospitalarios analizados.

En contraste, las reglas generadas por funciones de calidad con mayor énfasis en la discriminación, como *Lift* y *WRAcc*, fueron evaluadas de manera más favorable cuando sus antecedentes integraban atributos clínicos y formativos con coherencia semántica. En estos casos, el especialista reconoció patrones que reflejan prácticas reales, percepciones institucionales consistentes y dinámicas organizacionales plausibles, particularmente en hospitales con perfiles más homogéneos. La coincidencia entre altos valores de precisión y una interpretación clínica sólida reforzó la validez de estos patrones dentro del contexto del estudio.

Al analizar los resultados obtenidos a partir de la Matriz Binaria, el criterio experto identificó un comportamiento diferenciado. La binarización de los atributos redujo la complejidad semántica de los antecedentes, pero permitió aislar combinaciones estructurales más puras y consistentes. En este escenario, las reglas generadas presentaron una mayor estabilidad estadística y una menor ambigüedad interpretativa, lo que facilitó su validación clínica. El especialista destacó que, aunque estas reglas sacrifican detalle contextual, capturan de forma clara la presencia o ausencia de factores clave asociados a la solicitud o rechazo de autopsias, lo que las hace útiles como descriptores institucionales generales.

Particularmente, las reglas derivadas de *Lift* bajo la Matriz Binaria fueron valoradas positivamente por su alta coherencia interna y su ausencia de contraejemplos, al describir subgrupos clínicamente consistentes, aunque de cobertura limitada. De manera similar, los patrones generados por *WRAcc* y *Correlation* fueron considerados válidos al mantener un equilibrio entre generalidad y diferenciación, permitiendo identificar tendencias institucionales sin incurrir en combinaciones arbitrarias de atributos. En el caso de *RelativeRisk*, el especialista reconoció una mejora interpretativa respecto a la Vista Minable, al observar patrones más equilibrados y menos afectados por características compartidas entre hospitales.

De forma transversal, el criterio experto coincidió en que ninguna de las representaciones, por sí sola, es suficiente para capturar la totalidad del fenómeno analizado. La Vista Minable proporciona riqueza semántica y profundidad clínica, permitiendo comprender el contexto y la lógica subyacente a las decisiones médicas. En contraste, la Matriz Binaria ofrece una visión estructural depurada, útil para validar la estabilidad de los patrones y reducir la influencia de ruido semántico. La convergencia de patrones clínicamente plausibles en ambos conjuntos de datos fue interpretada como un indicador sólido de la robustez del módulo propuesto.

En síntesis, la evaluación subjetiva y el criterio experto confirman que los patrones contrastivos identificados no solo cumplen con criterios estadísticos formales, sino que

también poseen significado clínico y coherencia institucional. La consistencia interpretativa observada entre la Vista Minable y la Matriz Binaria refuerza la validez del enfoque metodológico adoptado y demuestra que el módulo de minería de conjuntos de contraste es capaz de generar conocimiento útil, interpretable y alineado con la realidad clínica de los hospitales analizados.

4.12.3 Selección de la Función de Calidad

La selección de la función de calidad para la operación del módulo de minería de conjuntos de contraste no se realizó bajo un criterio único y absoluto, sino a partir de una evaluación diferenciada y contextual, considerando de manera conjunta los resultados obtenidos en la Vista Minable, los resultados derivados de la Matriz Binaria y la evaluación subjetiva del experto clínico. Este enfoque permitió reconocer que el desempeño de las funciones de calidad depende directamente del esquema de representación de los datos, así como del tipo de conocimiento que se desea priorizar.

En el caso de la Vista Minable, donde los atributos conservan su riqueza semántica y permiten interpretar de manera directa aspectos formativos, organizacionales y clínicos, la función de calidad *Lift* demostró un desempeño superior. Esta métrica mostró una capacidad consistente para identificar patrones contrastivos con alta precisión y diferencias de soporte positivas, generando reglas que describen perfiles institucionales claramente diferenciados y clínicamente plausibles. La evaluación experta confirmó que las reglas obtenidas mediante *Lift* en este conjunto de datos reflejan de forma adecuada la realidad observada en los hospitales analizados, lo que la posiciona como la función de calidad más adecuada para este esquema de representación.

Por otro lado, en la Matriz Binaria, donde los atributos se representan de forma estructural mediante la presencia o ausencia de características, la evaluación experta reveló un comportamiento distinto. En este contexto, la función de calidad *Klosgen* alcanzó una aceptación clínica total, validándose la totalidad de las reglas generadas. Este resultado indica que, bajo una representación binaria, *Klosgen* produce patrones con alta coherencia clínica y estabilidad estructural, evitando combinaciones ambiguas o

inconsistentes desde el punto de vista médico. Aunque su comportamiento tiende a privilegiar patrones dominantes y puede presentar redundancia, esta característica resulta menos problemática en un entorno donde se prioriza la claridad estructural sobre la diversidad semántica.

Otras funciones de calidad, como *WRAcc* y *Correlation*, mostraron tasas de aceptación clínica elevadas en la Matriz Binaria; sin embargo, su comportamiento global evidenció una menor capacidad discriminativa en la Vista Minable, así como una tendencia a generar reglas amplias que limitan su utilidad para el análisis comparativo profundo entre hospitales. *RelativeRisk*, si bien obtuvo aceptación clínica completa en la Matriz Binaria, generó modelos excesivamente compactos que no capturan adecuadamente la complejidad institucional requerida por los objetivos del estudio.

La Tabla 4.22 sintetiza de manera integrada el desempeño estadístico y clínico de las funciones de calidad considerando ambos conjuntos de datos, y fundamenta la decisión metodológica adoptada en este trabajo.

En consecuencia, se establece un criterio dual de selección de la función de calidad: Lift se define como la función de referencia para el análisis basado en la Vista Minable, al maximizar el contraste semántico e institucional, mientras que *Klosgen* se selecciona como la función óptima para la Matriz Binaria, al garantizar coherencia clínica total y estabilidad estructural. Este criterio refuerza la robustez del módulo propuesto y demuestra su capacidad de adaptación a distintos esquemas de representación, evitando decisiones arbitrarias y fortaleciendo la validez metodológica del análisis.

Tabla 4.22. Matriz Comparativa de Desempeño

Función de Calidad	Desempeño Estadístico – Vista Minable	Desempeño Estadístico – Matriz Binaria	Evaluación del Experto – Vista Minable	Evaluación del Experto – Matriz Binaria	Observaciones Críticas Integradas	Decisión Metodológica
Lift	Alta capacidad discriminativa; reglas precisas y contrastivas	Muy alta pureza estructural; reglas altamente específicas	18 / 18 reglas aceptadas (100%)	42 / 53 reglas aceptadas (79.2%)	Excelente para contraste fino; algunas reglas binarias excesivamente específicas	**Óptima para Vista Minable**
WRAcc	Cobertura elevada; tendencia a generalizar	Buen equilibrio estadístico	13 / 18 reglas acetapadas (72.22%)	19 / 22 reglas aceptadas (86.4%)	Menor poder discriminativo institucional	No seleccionada
Correlation	Similar a WRAcc; estabilidad media	Alta estabilidad estadística	17/18 reglas aceptadas (94.44 %)	20 / 22 reglas aceptadas (90.9%)	Redundancia conceptual; bajo contraste	No seleccionada
Klosgen	Robusta cuantitativamente; alta redundancia	Patrones dominantes y estructuralmente claros	13 / 20 reglas aceptadas (65 %)	16 / 16 reglas aceptadas (100%)	Alta coherencia clínica en binarización; baja diversidad contrastiva	**Óptima para Matriz Binaria**
RelativeRisk	Modelo excesivamente compacto	Mejora con binarización	2 / 3 reglas aceptadas (66.67 %)	3 / 3 reglas aceptadas (100%)	Demasiado limitado para análisis comparativo	No seleccionada

Capítulo 5. Conclusiones y Recomendaciones

La investigación abordó la problemática de la disminución de autopsias clínicas desde una perspectiva tecnológica, identificando la carencia de herramientas que permitieran validar si las causales de este fenómeno en el Hospital Regional Río Blanco respondían a dinámicas exclusivamente locales o a patrones sistémicos compartidos con otras instituciones. Tras el desarrollo del módulo de minería de conjuntos de contraste y el análisis comparativo de los resultados obtenidos en dicho nosocomio frente al Hospital General de Zona 53 y el Hospital General San Juan Bautista Tuxtepec, se exponen a continuación las conclusiones derivadas de la validación del sistema, seguidas de las recomendaciones estratégicas orientadas a la continuidad de la investigación y la aplicación operativa del conocimiento adquirido.

5.1 Conclusiones

A continuación, se presentan las conclusiones categóricas derivadas de esta investigación, las cuales sintetizan el alcance técnico del módulo desarrollado y la validez clínica de los hallazgos. Estas aseveraciones son el resultado de la triangulación entre la fundamentación teórica, la implementación algorítmica y la validación experta, respondiendo directamente a las interrogantes planteadas al inicio del estudio.

- **Sobre el cumplimiento del objetivo general:** Se cumplió satisfactoriamente con el objetivo general al desarrollar e implementar un módulo de software capaz de comparar los resultados de encuestas médicas entre distintos hospitales mexicanos (Hospital Regional Río Blanco, Hospital General de Zona 53, Hospital General San Juan Bautista Tuxtepec). La incorporación de técnicas de Minería de Conjuntos de Contraste permitió, por primera vez, validar matemáticamente las similitudes y diferencias entre las opiniones médicas, trascendiendo el análisis descriptivo tradicional hacia un modelo de descubrimiento de conocimiento discriminativo.

- **Sobre la selección de la función de calidad (*Lift y Klosgen*):** Uno de los aportes científicos más relevantes de esta tesis radica en la evaluación crítica y contextual de las funciones de calidad, demostrando que su idoneidad no es absoluta, sino dependiente del esquema de representación de los datos y del tipo de conocimiento que se busca extraer en el dominio médico. A partir de la triangulación entre métricas estadísticas objetivas y la evaluación experta en anatomía patológica, se concluye que, en la Vista Minable, donde los atributos conservan su riqueza semántica y permiten una interpretación clínica directa, la métrica *Lift* es superior a *WRAcc*, *Correlation*, *Klosgen* y *RelativeRisk*, al lograr un equilibrio óptimo entre cobertura y precisión y generar reglas contrastivas, clínicamente plausibles y alineadas con los perfiles institucionales reales. En contraste, en la Matriz Binaria, donde se prioriza la estabilidad estructural y la reducción de ambigüedad semántica, la función de calidad *Klosgen* alcanza una aceptación clínica total por parte del experto, evidenciando su capacidad para producir patrones coherentes y libres de artefactos bajo representaciones binarias. En conjunto, estos resultados confirman que la utilidad de la minería de datos en medicina no depende únicamente del rendimiento estadístico, sino de la capacidad del algoritmo para aislar patrones clínicamente plausibles, coherentes con el contexto institucional y robustos frente a cambios en la representación de los datos, lo que refuerza la validez metodológica y el aporte científico del módulo propuesto.
- **Sobre la validación del conocimiento extraído:** El módulo permitió confirmar que las barreras para la realización de autopsias no son homogéneas entre las instituciones. Se identificaron patrones discriminativos claros (contrastes) relacionados con la formación académica, la percepción del riesgo legal y los procedimientos administrativos diferenciados entre los hospitales de Veracruz, Estado de México y Oaxaca. Esto valida la hipótesis de que las estrategias para revertir la tendencia deben ser adaptadas al contexto específico de cada centro médico, en lugar de aplicar políticas genéricas.
- **Impacto tecnológico y social:** Desde una perspectiva operativa, el sistema reduce significativamente la complejidad del análisis de datos para los patólogos.

La automatización del descubrimiento de patrones transforma datos brutos en conocimiento accionable, dotando a las instituciones de una herramienta de soporte a la decisión fundamentada para diseñar estrategias educativas y administrativas que fomenten la práctica de la autopsia.

5.2 Recomendaciones

Con base en la experiencia adquirida durante el desarrollo de la tesis y los resultados obtenidos, se proponen las siguientes líneas de acción:

Recomendaciones Operativas para las Instituciones de Salud

- **Adopción institucional:** Se recomienda la integración formal del módulo desarrollado en los procesos de gestión de calidad y enseñanza de los hospitales participantes. Su uso periódico permitiría monitorear la evolución de las opiniones médicas tras la implementación de cursos o cambios en la normativa hospitalaria.
- **Capacitación focalizada:** Dado que la función de calidad *Lift* identificó patrones específicos de "temor legal" y "desconocimiento del procedimiento" en subgrupos concretos de médicos, se sugiere diseñar programas de capacitación dirigidos específicamente a mitigar estas barreras, en lugar de cursos generalistas.

Recomendaciones Técnicas y Trabajo Futuro

- **Expansión del *dataset*:** Aunque los resultados son robustos para los tres hospitales estudiados, se recomienda ampliar la recolección de datos a instituciones del norte del país para enriquecer la base de conocimiento y evaluar si los patrones descubiertos mediante *Lift* o *Klosgen* se mantienen en contextos socioculturales distintos.
- **Hibridación de algoritmos:** Se sugiere explorar en trabajos futuros la implementación de enfoques híbridos que combinen la capacidad discriminativa de *Lift* con la cobertura de *WRAcc* en una etapa de post-procesamiento, con el fin de recuperar reglas de soporte bajo que podrían contener "casos raros" de alto interés clínico.

- **Interoperabilidad:** Para escalar la solución a nivel nacional, sería pertinente desarrollar un API (*Application Programming Interfaces)* que permitan la interconexión del módulo con los Sistemas de Información Hospitalaria y Expedientes Clínicos Electrónicos, facilitando la alimentación de datos en tiempo real sin depender de encuestas manuales.

Productos académicos



Missael Hernández-Rosado, Lisbeth Rodríguez-Mazahua, Jose Antonio Palet-Guzmán, Araceli Olmos-Vallejo, Ulises Juárez-Martínez, Ignacio López-Martínez y Giner Alor-Hernández.

“Usage of Contrast Set Mining and Contrast Pattern Mining Methods: A Scoping Review”

INDAUTOR
INSTITUTO NACIONAL DEL DERECHO DE AUTOR

Missael Hernández-Rosado, Lisbeth Rodríguez-Mazahua, Jose Antonio Palet-Guzmán, Araceli Olmos-Vallejo, Giner Alor-Hernández y Asdrubal López Chau.

“MÓDULO DE MINERÍA DE CONJUNTOS DE CONTRASTE PARA LA COMPARACIÓN DE OPINIONES MÉDICAS SOBRE LA DISMINUCIÓN DE AUTOPSIAS EN HOSPITAL MEXICANOS”

Referencias

[1] O. B. Morales, “Comité de ética en investigación,” Sitio Web del Comité de ética en investigación. Accessed: Oct. 09, 2024. [Online]. Available: https://www.incmnsz.mx/opencms/contenido/investigacion/

[2] “Verbal autopsy standards: ascertaining and attributing causes of death tool.” Accessed: Oct. 09, 2024. [Online]. Available: https://www.who.int/standards/classifications/other-classifications/verbal-autopsy-standards-ascertaining-and-attributing-causes-of-death-tool

[3] G. C. Alfsen et al., “Code of practice for medical autopsies: a minimum standard position paper for pathology departments performing medical (hospital) autopsies in adults,” Virchows Arch., vol. 480, no. 3, pp. 509–517, Mar. 2022, doi: 10.1007/s00428-021-03242-y.

[4] R. Menezes and F. Monteiro, “Forensic Autopsy,” StatPearls, Apr. 2023, Accessed: Oct. 09, 2024. [Online]. Available: https://www.statpearls.com/point-of-care/18038

[5] “Cause & Manner of Death | Snohomish County, WA - Official Website.” Accessed: Oct. 09, 2024. [Online]. Available: https://snohomishcountywa.gov/806/Cause-Manner-of-Death

[6] A. Olmos Vallejo, L. Rodríguez-Mazahua, and I. Machorro-Cano, “Comparación de resultados de encuestas sobre la disminución de autopsias en hospitales mexicanos usando técnicas de Minería de Datos,” Instituto Tecnológico de Orizaba, 2023. [Online]. Available: http://repositorios.orizaba.tecnm.mx:8080/xmlui/handle/123456789/770

[7] A. Serra Marrugat, “Comparación de algoritmos de clasificación supervisada,” Master thesis, Universitat Politècnica de Catalunya, 2020. Accessed: Oct. 09, 2024. [Online]. Available: https://upcommons.upc.edu/handle/2117/330482

[8] “¿Qué es el aprendizaje no supervisado? | IBM.” Accessed: Feb. 08, 2025. [Online]. Available: https://www.ibm.com/es-es/topics/unsupervised-learning

[9] A.-L. Pérez-Suasnavas, K. Cela, and W. Hasperué, “Beneficios del uso de técnicas de minería de datos para extraer y analizar datos de Twitter aplicados en

la educación superior : una revisión sistemática de la literatura," pp. 181–218, 2020, doi: 10.14201/teri.22171.

[10]Á. M. García Vico, Modelos descriptivos basados en aprendizaje supervisado para el tratamiento de big data y flujos continuos de datos. Jaén : Universidad de Jaén, 2020. Accessed: Oct. 09, 2024. [Online]. Available: http://ruja.ujaen.es/jspui/handle/10953/1164

[11]E. Rubio Delgado et al., "Analysis of Medical Opinions about the Nonrealization of Autopsies in a Mexican Hospital Using Association Rules and Bayesian Networks," Sci. Program., vol. 2018, no. 1, p. 4304017, 2018, doi: 10.1155/2018/4304017.

[12]I. A. R. Méndez, "Análisis de encuestas sobre la disminución de autopsias en hospitales mexicanos usando técnicas de minería de datos," de Maestría, Tecnológico Nacional de México Campus Orizaba, 2020.

[13]T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised Machine Learning: A Brief Primer," Behav. Ther., vol. 51, no. 5, pp. 675–687, Sep. 2020, doi: 10.1016/j.beth.2020.05.002.

[14]H. Jayne, A. Lukovnikova, H. H. Nhung, and Z. Liang, "Mining Contrast Rules in a Sleep Apnea Dataset," in 2023 IEEE 12th Global Conference on Consumer Electronics (GCCE), Nara, Japan: IEEE, Oct. 2023, pp. 1057–1060. doi: 10.1109/GCCE59613.2023.10315646.

[15]U. Ahmed, J. C.-W. Lin, and G. Srivastava, "Fuzzy Contrast Set Based Deep Attention Network for Lexical Analysis and Mental Health Treatment," ACM Trans. Asian Low-Resour. Lang. Inf. Process., vol. 21, no. 5, pp. 1–16, Sep. 2022, doi: 10.1145/3506701.

[16]A. M. Trasierras, J. M. Luna, and S. Ventura, "A contrast set mining based approach for cancer subtype analysis," Artif. Intell. Med., vol. 143, p. 102590, Sep. 2023, doi: 10.1016/j.artmed.2023.102590.

[17]R. Qian, Y. Yu, W. Park, V. Murali, S. Fink, and S. Chandra, "Debugging crashes using continuous contrast set mining," in Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering in

Practice, Seoul South Korea: ACM, Jun. 2020, pp. 61–70. doi: 10.1145/3377813.3381369.

[18]S. Priyadarshini et al., “Application of STUCCO Algorithm for Finding Contrast Sets for Agricultural Datasets,” J. Indian Soc. Agric. Stat., 2022.

[19]I. Varlamis, “Affinity analysis for studying physicians’ prescription behavior.: The case of hypertension and hyperlipidemia,” Data Min. Knowl. Discov., vol. 35, no. 4, pp. 1739–1759, Jul. 2021, doi: 10.1007/s10618-021-00758-4.

[20]M. Diouf, M. Thiam, and M. Roche, “New approach to discover meaningful terms to specify cause of death from narratives verbal autopsy using TF-IDF and the LDA topic model,” in IEEE EUROCON 2023 - 20th International Conference on Smart Technologies, Torino, Italy: IEEE, Jul. 2023, pp. 502–507. doi: 10.1109/EUROCON56442.2023.10198945.

[21]I. Machorro-Cano et al., “Medical Opinions Analysis about the Decrease of Autopsies Using Emerging Pattern Mining,” Data, vol. 9, no. 1, p. 2, Dec. 2023, doi: 10.3390/data9010002.

[22]A. Huidobro, R. Monroy, M. A. Godoy, and B. Cervantes, “A Contrast-Pattern Characterization of Web Site Visitors in Terms of Conversions,” in Technology-Enabled Innovations in Education, S. Hosseini, D. H. Peluffo, J. Nganji, and A. Arrona-Palacios, Eds., in Transactions on Computer Systems and Networks. , Singapore: Springer Nature Singapore, 2022, pp. 31–51. doi: 10.1007/978-981-19-3383-7_3.

[23]M. T. Mapundu, C. W. Kabudula, E. Musenge, V. Olago, and T. Celik, “Explainable Stacked Ensemble Deep Learning (SEDL) Framework to Determine Cause of Death from Verbal Autopsies,” Mach. Learn. Knowl. Extr., vol. 5, no. 4, pp. 1570–1588, Oct. 2023, doi: 10.3390/make5040079.

[24]M. R. Shahid, A. Munir, L. Ifraheem, H. Aldabbas, A. Wadood, and T. Alwada’n, “Machine learning for Autopsy reports Forensic using Text Classification Techniques,” in 2022 2nd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia: IEEE, Jan. 2022, pp. 148–153. doi: 10.1109/ICCIT52419.2022.9711658.

[25]S.-K. Chang, D. Liu, J. Mitchem, C. Papageorgiou, J. Kaifi, and C.-R. Shyu, “Understanding common key indicators of successful and unsuccessful cancer drug trials using a contrast mining framework on ClinicalTrials.gov,” J. Biomed. Inform., vol. 139, p. 104321, Mar. 2023, doi: 10.1016/j.jbi.2023.104321.

[26]X. Chen et al., “Hierarchical and nested associations of suicide with marriage, social support, quality of life, and depression among the elderly in rural China: Machine learning of psychological autopsy data,” Front. Psychiatry, vol. 13, p. 1000026, Sep. 2022, doi: 10.3389/fpsyt.2022.1000026.

[27]T. Manaka, T. Van Zyl, and D. Kar, “Improving Cause-of-Death Classification from Verbal Autopsy Reports,” in Artificial Intelligence Research, vol. 1734, A. Pillay, E. Jembere, and A. Gerber, Eds., in Communications in Computer and Information Science, vol. 1734. , Cham: Springer Nature Switzerland, 2022, pp. 46–59. doi: 10.1007/978-3-031-22321-1_4.

[28]S. Fan et al., “From Narratives to Numbers: Valid Inference Using Language Model Predictions from Verbal Autopsy Narratives,” Apr. 03, 2024, arXiv: arXiv:2404.02438. Accessed: Oct. 09, 2024. [Online]. Available: http://arxiv.org/abs/2404.02438

[29]A. Gudyś, C. Maszczyk, J. Badura, A. Grzelak, M. Sikora, and Ł. Wróbel, “RuleKit 2: Faster and simpler rule learning,” arXiv preprint, arXiv:2504.20650, 2025.

[30]A. Gudyś, M. Sikora, and Ł. Wróbel, “Separate and conquer heuristic allows robust mining of contrast sets in classification, regression, and survival data,” Expert Systems with Applications, vol. 248, p. 123376, 2024.

[31]A. Gudyś, M. Sikora, and Ł. Wróbel, “RuleKit: A comprehensive suite for rule-based learning,” Knowledge-Based Systems, vol. 194, p. 105480, 2020.

[32]M. Sikora, Ł. Wróbel, and A. Gudyś, “GuideR: A guided separate-and-conquer rule learning in classification, regression, and survival settings,” Knowledge-Based Systems, vol. 173, pp. 1–14, 2019.

[33]Ł. Wróbel, A. Gudyś, and M. Sikora, “Learning rule sets from survival data,” BMC Bioinformatics, vol. 18, no. 1, p. 285, May 2017, doi: 10.1186/s12859-017-1693-x.

[34]Arksey, H., O'Malley, L.: Scoping studies: Towards a methodological framework. Int. J. Soc. Res. Methodol. 8(1), 19-32 (2005)

[35]Levac, D., Colquhoun, H., O'Brien, K.K.: Scoping studies: Advancing the methodology. Implement. Sci. 5, 1-9 (2010)

[36]Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., PRISMA Group.: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 6, e1000097 (2009)

[37]Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D.J., Horsley, T., Weeks, L., et al.: PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Ann. Intern. Med. 169, 467-473 (2018)

[38]Haddaway, N.R., Page, M.J., Pritchard, C.C., McGuinness, L.A.: PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. Campbell Systematic Reviews 18(2), e1230 (2022)

[39]Loyola-González, O., Medina-Pérez, M.A., Coronilla Valdez, R.A., Choo, K.-K.R.: A Contrast Pattern-Based Scientometric Study of the QS World University Ranking. IEEE Access 8, 206088–206104 (2020). https://doi.org/10.1109/ACCESS.2020.3037665

[40]Lisi, F.A., Sterlicchio, G.: A Declarative Approach to Contrast Pattern Mining. In: Dovier, A. et al. (eds.) AIxIA 2022 - Advances in Artificial Intelligence. Lecture Notes in Artificial Intelligence, vol. 13796, pp. 17–30. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-27181-6_2

[41]Carmona, C.J., del Jesus, M.J., Herrera, F.: A Unifying Analysis for the Supervised Descriptive Rule Discovery via the Weighted Relative Accuracy. Knowledge-Based Systems 139, 89–100 (2018). https://doi.org/10.1016/j.knosys.2017.10.015

[42]Nguyen, D., Luo, W., Vo, B., Nguyen, L.T.T., Pedrycz, W.: Con2Vec: Learning Embedding Representations for Contrast Sets. Knowledge-Based Systems 229, 107382 (2021). https://doi.org/10.1016/j.knosys.2021.107382

[43]Li, Q., Zhao, L., Lee, Y.-C., Lin, J.: Contrast Pattern Mining in Paired Multivariate Time Series of Controlled Driving Behavior Experiment. ACM Transactions on

Spatial Algorithms and Systems 9(4), Article 39 (2018). https://doi.org/10.1145/0000001.0000001

[44]Loyola-González, O., Monroy, R., Rodríguez, J., López-Cuevas, A., Mata-Sánchez, J.I.: Contrast Pattern-Based Classification for Bot Detection on Twitter. IEEE Access 7, 45800–45817 (2019). https://doi.org/10.1109/ACCESS.2019.2904220

[45]Ahmed, U., Lin, J.C.-W., Srivastava, G.: Deep Fuzzy Contrast-Set Deviation Point Representation and Trajectory Detection. IEEE Transactions on Fuzzy Systems 31(2), 571–581 (2023). https://doi.org/10.1109/TFUZZ.2022.3197876

[46] Saadat, S., Colmenares, N., Sukthankar, G.: Do Bots Modify the Workflow of GitHub Teams? In: Proceedings of the Third International Workshop on Bots in Software Engineering (BotSE 2021), pp. 8–17. IEEE (2021). https://doi.org/10.1109/BotSE52550.2021.00008

[47]Saadat, S., Colmenares, N., Sukthankar, G.: Do Bots Modify the Workflow of GitHub Teams? In: Proceedings of the Third International Workshop on Bots in Software Engineering (BotSE 2021), pp. 8–17. IEEE (2021). https://doi.org/10.1109/BotSE52550.2021.00008

[48]Fu, M., Duan, L., Yu, Z.: Effective Mining of Contrast Hybrid Patterns from Nominal-Numerical Mixed Data. In: Chen, W. et al. (eds.) Advances in Data Mining and Applications (ADMA 2022), Lecture Notes in Artificial Intelligence, vol. 13725, pp. 352–367. Springer (2022). https://doi.org/10.1007/978-3-031-22064-7_26

[49]Cito, J., Dillig, I., Kim, S., Murali, V., Chandra, S.: Explaining Mispredictions of Machine Learning Models using Rule Induction. In: Proceedings of the 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2021), pp. 1–12. ACM (2021). https://doi.org/10.1145/3468264.3468614

[50]Liu, D., Baskett, W., Beversdorf, D., Shyu, C.-R.: Exploratory Data Mining for Subgroup Cohort Discoveries and Prioritization. IEEE Journal of Biomedical and Health Informatics 24(5), 1456–1468 (2020). https://doi.org/10.1109/JBHI.2019.2939149

[51]Ozaki, T.: Extraction of Characteristic Subgraph Patterns with Support Threshold from Databases of Floor Plans. In: Proceedings of the Seventh International Symposium on Computing and Networking (CANDAR 2019), pp. 197–203. IEEE (2019). https://doi.org/10.1109/CANDAR.2019.00033

[52]Khade, R., Lin, J., Patel, N.: Finding Meaningful Contrast Patterns for Quantitative Data. In: Proceedings of the 22nd International Conference on Extending Database Technology (EDBT 2019), pp. 444–455. OpenProceedings (2019). https://doi.org/10.5441/002/edbt.2019.41

[53]Spencer, M., Takahashi, N., Chakraborty, S., Miles, J., Shyu, C.-R.: Heritable Genotype Contrast Mining Reveals Novel Gene Associations Specific to Autism Subgroups. Journal of Biomedical Informatics 77, 50–61 (2018). https://doi.org/10.1016/j.jbi.2017.11.016

[54]Manchanayaka, I., Zaidi, Z., Karunasekera, S., Leckie, C.: Identifying Coordinated Activities on Online Social Networks Using Contrast Pattern Mining. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN 2024). IEEE (2024). https://doi.org/10.1109/IJCNN60899.2024.10651552

[55]Jayne, H., Nhung, H.H., Lukovnikova, A., Liang, Z.: Mining Contrast Rules in a Sleep Apnea Dataset. In: Proceedings of the 12th IEEE Global Conference on Consumer Electronics (GCCE 2023), pp. 1057–1060. IEEE (2023). https://doi.org/10.1109/GCCE59613.2023.10315646

[56]Wu, R., Li, Q., Chen, X.: Mining Contrast Sequential Pattern Based on Subsequence Time Distribution Variation with Discreteness Constraints. Applied Intelligence 49(6), 3247–3264 (2019). https://doi.org/10.1007/s10489-019-01492-7

[57]Alipourchavary, E., Erfani, S.M., Leckie, C.: Mining Rare Recurring Events in Network Traffic using Second Order Contrast Patterns. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN 2021), pp. 1–8. IEEE (2021). https://doi.org/10.1109/IJCNN52387.2021.9533918

[58]Neubarth, K., Conklin, D.: Modelling Pattern Interestingness in Comparative Music Corpus Analysis. Journal of Mathematics and Music 15(2), 154–167 (2021). https://doi.org/10.1080/17459737.2021.1900436

[59]Cervantes, B., Gómez, F., Monroy, R., Loyola-González, O., Medina-Pérez, M.A., Ramírez-Márquez, J.: Pattern-Based and Visual Analytics for Visitor Analysis on Websites. Applied Sciences 9(18), 3840 (2019). https://doi.org/10.3390/app9183840

[60]Alipourchavary, E., Erfani, S.M., Leckie, C.: Scalable Contrast Pattern Mining over Data Streams. In: Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM 2021), pp. 2842–2846. ACM (2021). https://doi.org/10.1145/3459637.3482174

[61]Gudyś, A., Sikora, M., Wróbel, Ł.: Separate and Conquer Heuristic Allows Robust Mining of Contrast Sets in Classification, Regression, and Survival Data. arXiv preprint arXiv:2204.00497v3 (2023)

[62]Nguyen, D., Luo, W., Vo, B., Pedrycz, W.: Succinct Contrast Sets via False Positive Controlling with an Application in Clinical Process Redesign. Expert Systems with Applications 161, 113670 (2020). https://doi.org/10.1016/j.eswa.2020.113670

[63]Orabi, M., Al Aghbari, Z.: Temporal Contrast Sets Mining. In: Proceedings of the International Conference on Soft Computing and Pattern Recognition (icSoftComp 2023). Communications in Computer and Information Science, vol. 2030, pp. 94–103. Springer (2024). https://doi.org/10.1007/978-3-031-53731-8_8

[64]Chang, S.-K., Liu, D., Mitchem, J., Papageorgiou, C., Kaifi, J., Shyu, C.-R.: Understanding Common Key Indicators of Successful and Unsuccessful Cancer Drug Trials Using a Contrast Mining Framework on ClinicalTrials.gov. Journal of Biomedical Informatics 139, 104321 (2023). https://doi.org/10.1016/j.jbi.2023.104321

[65]Aryabarzan, N., Minaei-Bidgoli, B.: A Novel Pruning Strategy for Mining Discriminative Patterns. Iranian Journal of Science and Technology, Transactions of Electrical Engineering 45, 505–527 (2021). https://doi.org/10.1007/s40998-020-00397-3

[66]Ahmed, U., Lin, J.C.-W., Tomasiello, S., Srivastava, G.: An Explainable Mental Health Fuzzy Deep Active Learning Technique. In: 2022 IEEE International

Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–8. IEEE (2022). https://doi.org/10.1109/FUZZ-IEEE55066.2022.9882622

[67]An, D.W., Kim, S., Kim, H.K., Kim, C.O.: Commonality Analysis for Detecting Failures Caused by Inspection Tools in Semiconductor Manufacturing Processes. IEEE Transactions on Semiconductor Manufacturing 35(4), 596–604 (2022). https://doi.org/10.1109/TSM.2022.3201654

[68]Li, L., Zhang, X., He, S., Kang, Y., Zhang, H., Ma, M., Dang, Y., Xu, Z., Rajmohan, S., Lin, Q., Zhang, D.: CONAN: Diagnosing Batch Failures for Cloud Systems. In: Proceedings of the 45th IEEE/ACM International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP 2023), pp. 97–110. IEEE (2023). https://doi.org/10.1109/ICSE-SEIP58684.2023.00018

[69]Saadat, S., Sukthankar, G.: Contrast Motif Discovery in Minecraft. In: Proceedings of the Sixteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-20). AAAI (2020)

[70]Wu, Y., Meng, Y., Li, Y., Guo, L., Zhu, X., Fournier-Viger, P., Wu, X.: COPP-Miner: Top-k Contrast Order-Preserving Pattern Mining for Time Series Classification. IEEE Transactions on Knowledge and Data Engineering 36(6), 2372–2387 (2024). https://doi.org/10.1109/TKDE.2023.3321749

[71]Loyola-González, O., Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., García-Borroto, M.: Cost-Sensitive Pattern-Based Classification for Class Imbalance Problems. IEEE Access 7, 60411–60427 (2019). https://doi.org/10.1109/ACCESS.2019.2913982

[72]Khade, R., Lin, J., Patel, N.: Finding Contrast Patterns for Mixed Streaming Data. In: Proceedings of the 21st International Conference on Extending Database Technology (EDBT 2018), pp. 632–641. ACM (2018). https://doi.org/10.5441/002/edbt.2018.73

[73]Rawassizadeh, R., Dobbins, C., Akbari, M., Pazzani, M.: Indexing Multivariate Mobile Data through Spatio-Temporal Event Detection and Clustering. Sensors 19(3), 448 (2019). https://doi.org/10.3390/s19030448

[74]D. Nguyen, W. Luo, D. Phung, and S. Venkatesh, “Exceptional Contrast Set Mining: Moving Beyond the Deluge of the Obvious,” in AI 2016: Advances in

Artificial Intelligence, B. H. Kang and Q. Bai, Eds., Cham: Springer International Publishing, 2016, pp. 455–468. doi: 10.1007/978-3-319-50127-7_39.