



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México

**Centro Nacional de Investigación
y Desarrollo Tecnológico**

Tesis de Maestría

**Detección Automática de Ansiedad a Través del
Monitoreo de Sensores Vestibles**

presentada por

Lic. Patricia Alejandra Cuevas Chávez

como requisito para la obtención del grado de
Maestra en Ciencias de la Computación

Director de tesis

Dra. Alicia Martínez Rebollar

Codirector de tesis

M.C. Angélica Fabiola Sánchez Gutiérrez

Cuernavaca, Morelos, México. Noviembre de 2020.



“2020, Año de Leona Vicario, Benemérita Madre de la Patria”

Cuernavaca, Mor., 15/julio/2020

OFICIO No. DCC/106/2020

Asunto: Aceptación de documento de tesis
CENIDET-AC-004-M14-OFICIO

C. DR. GERARDO VICENTE GUERRERO RAMÍREZ
SUBDIRECTOR ACADÉMICO
PRESENTE

Por este conducto, los integrantes de Comité Tutorial de la **C. Ing. Patricia Alejandra Cuevas Chávez**, con número de control M18CE074, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado **“Detección automática de ansiedad a través del monitoreo de sensores vestibles”** y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

Dra. Alicia Martínez Rebollar
Doctora en Informática
7399055
Directora de tesis

M.C. Angélica Fabiola Sánchez Gutiérrez
Maestra en Ciencias y Humanidades
Co-directora de tesis

Dr. Javier Ortiz Hernández
Doctor en Informática
10138166
Revisor 1

Dr. José Jasson Flores Prieto
Doctor en Ingeniería
4355408
Revisor 2

C.c.p. Depto. Servicios Escolares
Expediente / Estudiante
JGGS/lmz



“2020, Año de Leona Vicario, Benemérita Madre de la Patria”

Cuernavaca, Morelos **22/octubre/2020**

OFICIO No. SAC/ 256/2020

Asunto: Autorización de impresión de tesis

PATRICIA ALEJANDRA CUEVAS CHÁVEZ
CANDIDATA AL GRADO DE MAESTRA EN CIENCIAS
DE LA COMPUTACIÓN
P R E S E N T E

Por este conducto tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado *“Detección automática de ansiedad a través del monitoreo de sensores vestibles”*, ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

A T E N T A M E N T E

Excelencia en Educación Tecnológica®

“Conocimiento y tecnología al servicio de México”

DR. CARLOS MANUEL ASTORGA ZARAGOZA
SUBDIRECTOR ACADÉMICO



**CENTRO NACIONAL
DE INVESTIGACIÓN
Y DESARROLLO
TECNOLÓGICO
SUBDIRECCIÓN
ACADÉMICA**

C.c.p. M.E. Guadalupe Garrido Rivera. Jefa del Departamento de Servicios Escolares

Expediente

CMAZ/CHG

Dedicatoria

«桃栗三年柿八年»

«*Peaches and chestnuts take three good years, persimmons take eight*»

Bien decía Victor Hugo, el éxito no se logra sólo con cualidades especiales. Es sobre todo, un trabajo de constancia, de método y de organización. Así como el éxito no se verá a corto plazo, ni sus frutos serán cosechados precipitadamente. El éxito no es como el crecimiento presuroso de los árboles de durazno y los castaños; es más como el de los árboles de persimon, con un crecimiento pausado; pero con una cosecha exuberante. El éxito para muchos, puede ser sólo personal; pero el éxito del trabajo realizado en esta tesis se debe en gran medida a diversos factores, personas y organizaciones que contribuyeron en gran medida a impulsarme en la vida de la investigación, a querer generar un cambio, a querer ser parte del cambio.

Este trabajo de investigación, va dedicado a todas esas personas que han estado siempre para mí, que me han apoyado incondicionalmente, que me han sabido aconsejar, y, que han influido en mí de tantas maneras: a mis padres *Patricia* y *Manuel*, a mis hermanos *Oscar*, *Sócrates* y *Cynthia*. A los pequeños *Rodrigo* y *Sebastián*. A mi mejor amigo y pareja sentimental, *Francisco*. Sobre todo, está dedicada a aquellas personas que lamentablemente ya no están para compartir el logro; pero que se siguen recordando con mucho cariño. Sin ustedes, el éxito obtenido y la meta de culminar mis estudios de maestría no se hubieran logrado.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada para la realización de mis estudios de maestría. Así como a la beca de movilidad para realizar una estancia de investigación en la Universidad de Hokkaido, Japón.

Al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), que forma parte del TecNM, por brindarme la oportunidad de pertenecer a su comunidad estudiantil.

Agradezco también a mi directora de tesis la Dra. Alicia Martínez y a los miembros del comité tutorial por su tiempo, dedicación, consejos, y observaciones para mejorar esta investigación: Dr. Javier Ortíz y Dr. Jassón Flores. A mi co-directora, M. en C. Fabiola Sánchez por brindarme su infinito apoyo para la comprensión de los temas psicológicos relacionados a esta investigación.

Al doctor Mineichi Kudo y a los miembros pertenecientes del laboratorio de Reconocimiento de Patrones y Aprendizaje Automático de la Universidad de Hokkaido, por haberme permitido realizar una estancia en sus instalaciones, así como brindarme apoyo y consejos para desarrollar este trabajo de investigación. A la Agencia de Cooperación Internacional del Japón (JICA) por haberme dado la oportunidad de descubrir el mundo de la investigación y buscar siempre el altruismo a la sociedad.

A mi familia por siempre impulsarme. A mis padres por haberme dado la oportunidad de estudiar la licenciatura en una de las mejores universidades (ITESM), y por darme su amor incondicional. A mis hermanos por ponerme siempre metas nuevas que cumplir a través de su ejemplo y por su infinito apoyo. ¡Gracias, me siento muy afortunada de tenerlos como familia!

Al M. en A. Francisco A. Brito Román, por exigirme siempre ser mejor y a entregar resultados de calidad, por sus consejos, su infinita paciencia y todo el apoyo inmenso que me ha brindado no solo en el desarrollo de esta investigación. ¡Gracias! Sin ti no hubiera podido superar todos los retos involucrados en esta investigación.

Resumen

Los trastornos de ansiedad son los trastornos mentales más comunes; además de ser la sexta causa de discapacidad laboral en el mundo. Las personas con trastornos de ansiedad generalizada tienden a sufrir una preocupación excesiva y persistente acerca de varios eventos rutinarios que está presente en al menos seis meses afectando su vida diaria.

En esta investigación, se aborda la detección automática del trastorno de ansiedad generalizada a través del monitoreo de sensores vestibles. Por lo que, el objetivo general de la investigación, se enfoca en desarrollar un modelo de comportamiento para adultos jóvenes, monitoreando respuestas de déficit cognitivo conductual relacionadas al trastorno de ansiedad generalizada a través de dispositivos vestibles.

En la revisión del estado del arte se encontró que, para dar solución al problema, se hace uso de un dispositivo capaz de monitorear signos vitales; tales como la frecuencia cardíaca, la respiración, o la temperatura corporal. Además de obtener otros datos fisiológicos como la actividad física o el movimiento. En contraste con el estado del arte, se aplicaron instrumentos psicológicos para detectar el trastorno de ansiedad generalizada; además de seleccionar un dispositivo comercial para monitorear a los participantes en su vida diaria.

Para evaluar la propuesta de solución, se realizaron pruebas de monitoreo donde 5 participantes fueron reclutados para la obtención de un modelo de comportamiento. A cada participante se le otorgó un dispositivo Fitbit y se les pidió que lo usaran día y noche por un total de 15 días. A través de este dispositivo se recolectaron datos fisiológicos característicos del trastorno de ansiedad generalizada (frecuencia cardíaca, actividad física y calidad del sueño). Para obtener este modelo de comportamiento como una estructura de diagrama de flujo, se implementó el algoritmo clasificador árbol de decisión. Este algoritmo eliminó 17 atributos de 22, seleccionando sólo 5 como más importantes: media de la frecuencia cardíaca, eficiencia del ciclo del sueño, número de pasos dados, minutos dormidos en etapa de *sueño REM* y minutos en actividad *ligeramente activa*. La evaluación con métricas del rendimiento del modelo obtenido, en

términos de precisión, exactitud y cobertura (sensibilidad); mostró un valor del 100% y un valor F-1 de 1.0, al detectar dos niveles de ansiedad correspondiente al trastorno de ansiedad generalizada. El modelo de comportamiento obtenido detectó dos niveles de ansiedad: bajo y medio. Así como sólo 2 participantes reportaron 8 niveles de ansiedad medio; mientras el resto reportó un nivel de ansiedad bajo. Sin embargo, ninguno de los participantes reportó un nivel alto de ansiedad.

Para la validación de esta propuesta de solución, se realizaron otras pruebas de monitoreo donde se recolectaron los datos pertenecientes a los 5 atributos seleccionados como más relevantes por el algoritmo árbol de decisión. 3 participantes fueron reclutados para la evaluación del modelo. A cada participante se le otorgó un dispositivo Fitbit y se les pidió usarlo día y noche por un total de 16 días. En total, sólo un participante obtuvo dos veces un nivel medio de ansiedad. El resto de los participantes obtuvieron un nivel bajo de ansiedad. Adicionalmente, los datos recolectados de la calidad del sueño muestran que los participantes sufren de alguna alteración del sueño al obtener entre 22% y 60% en la eficiencia del sueño. No obstante, su frecuencia cardíaca no mostró alguna alteración, manteniéndose en una frecuencia cardíaca ideal.

Los resultados obtenidos en esta investigación muestran que el trastorno de ansiedad generalizada puede ser detectado con dispositivos vestibles pero es complejo el detectar un nivel de ansiedad alto en un periodo de monitoreo de 16 días.

Keywords

Ansiedad, Trastorno de ansiedad generalizada, Detección, Dispositivos vestibles, Fitbit

Abstract

Anxiety disorders are the most common type of mental disorder, as the sixth work disability in the world. Generalized anxiety disorder, for example, is a condition where people tend to suffer from excessive worrying persisting with other symptoms at least for six months about everyday events that can hinder their daily functions.

In this research, the automatic detection of generalized anxiety disorder is addressed through wearable sensor monitoring. Therefore, the general objective is to focus on developing a behavior model for young adults, monitoring cognitive-behavioral deficit responses related to generalized anxiety disorder through wearable devices.

In the review of the state of the art, the solution found was to use a wearable device able to monitor vital signs, such as heart rate, breathing, or body temperature. Besides, other physiological data, such as physical activity, or movement, was detected. In contrast to the state of the art, psychological instruments were applied to detect generalized anxiety disorder. Also, a commercial device to monitor the participants in their daily life was selected.

The evaluation of the proposed solution was conducted by monitoring test, in which 5 participants were recruited to obtain a generalized anxiety disorder behavioral model. Each one was given the Fitbit wearable device and asked to wear day and night for 15 days. The following features were collected as data by a Fitbit wearable device: heart rate, physical activity, and sleep quality. In order to build the behavioral model as a flowchart-like structure, the Decision Tree algorithm was applied. This algorithm removed 17 features from 22 selecting only 5: median heart rate, sleep efficiency, steps, minutes asleep in the REM stage, and light activity minutes. The evaluation metrics of the model, in terms of accuracy, specificity, and sensitivity, achieved a 100% value and an F1 score 1.0 value in detecting two intensity levels of generalized anxiety disorder. The behavioral model of generalized anxiety disorder built only detected two anxiety intensity levels: low and medium. In total, 2 participants reported 8 medium anxiety levels. The rest of the data reported a low anxiety level, but none of them registered a high anxiety intensity level.

Instead, the validation of the proposed solution was conducted by monitor testing, but, in contrast to the evaluation, this monitoring test only collected the 5 most relevant features selected by the decision tree algorithm. 3 subjects were recruited and participated in the evaluation study. Each one was given the Fitbit wearable device and asked to wear it day and night for 16 days. In total, only one participant reported two medium anxiety levels. The rest of the participants reported a low-intensity level. Also, sleep data collected show the participants have sleep disorders. They reported between 22% and 60% in sleep efficiency, but their heart rate was not being affected. On the other hand, some participants reported a heart rate of over 85 beats per minute, but their sleep efficiency for that day was good.

The results obtained in this research show that generalized anxiety disorder can be detected, but it is complex to detect a high-intensity level in 16 days monitoring their daily life.

Keywords

Anxiety, Generalized Anxiety Disorder, Detection, Wearable Devices, Fitbit

Contenido

Índice de figuras	X
Índice de tablas	XIII
Índice de fórmulas	XIV
Glosario de variables	XV
Capítulo 1. Introducción	1
1.1 Introducción/Motivación	1
1.2 Planteamiento del problema	2
1.3 Objetivos	4
1.3.1 Objetivo general	4
1.3.2 Objetivos específicos	5
1.4 Estado del arte	5
1.4.1 Revisión sistemática	5
1.4.2 Criterios de clasificación	13
1.4.3 Conclusiones	17
1.5 Metodología de solución	18
1.6 Estructura de la tesis	21
Capítulo 2. Marco conceptual	22
2.1 Ansiedad	22
2.1.1 Trastorno de ansiedad generalizada	22
2.1.2 ¿Cómo se produce la ansiedad?	23
2.1.3 Manifestaciones sintomatológicas de la ansiedad	26
2.2 Marcos regulatorios	27
2.2.1 Manual diagnóstico y estadístico de trastornos mentales (DSM-V)	28
2.2.2 Clasificación estadística internacional de enfermedades y problemas relacionados con la salud (CIE-10)	28
2.3 Instrumentos de valoración para determinar el trastorno de ansiedad generalizada	28
2.4 Dispositivos vestibles	29
2.4.1 Pulseras de actividad: Fitbit	30
2.5 Herramientas de software utilizadas en la investigación	30

2.5.1 IBM SPSS Statistics	30
2.5.2 Anaconda	31
2.5.3 Jupyter	31
2.6 Librerías utilizadas en la investigación	31
2.6.1 Pandas	31
2.6.2 Scikit-learn	31
2.7 Aprendizaje automático o machine learning	32
2.7.1 Algoritmos de aprendizaje automático o machine learning	32
Capítulo 3. Validación de instrumentos psicológicos para la detección de ansiedad generalizada	33
3.1 Validación de instrumentos para la detección del trastorno de ansiedad generalizada	33
3.2 Definición de los cuestionarios	34
3.2.1 Definición del perfil del grupo de estudio	34
3.2.2 Instrumentos psicológicos seleccionados para la investigación	35
3.3 Aplicación de los instrumentos	38
3.3.1 Equilibrio de datos con técnicas estadísticas	39
3.4 Validación de los instrumentos	41
3.4.1 Análisis factorial	41
3.4.2 Análisis de fiabilidad	50
3.4.3 Resultados	52
Capítulo 4. Modelo de comportamiento del trastorno de ansiedad generalizada	53
4.1 Formación de grupos de conveniencia	55
4.1.1 Definición del perfil del grupo de estudio	55
4.1.2 Grupos de conveniencia	56
4.2 Modelo de comportamiento	57
4.2.1 Pruebas piloto de monitoreo	57
4.2.2 Extracción automática de los datos del dispositivo vestible	60
4.2.3 Análisis exploratorio de los datos	63
4.2.4 Preparación de los datos	73
4.2.5 Obtención de atributos relevantes	88
4.2.6 Obtención del modelo de comportamiento	96
4.3 Evaluación del modelo obtenido con métricas	107

Capítulo 5. Evaluación del modelo y resultados	109
5.1 Descripción de las pruebas de monitoreo para la evaluación del modelo de comportamiento del trastorno de ansiedad generalizada obtenido	110
5.1.1 Perfil de los participantes	110
5.1.2 Selección de los participantes	110
5.1.3 Pruebas de monitoreo	111
5.2 Evaluación del modelo de comportamiento del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada	111
5.2.1 Implementación de reglas obtenidas del árbol de decisión correspondiente al trastorno de ansiedad generalizada	112
5.2.2 Obtención del nivel de ansiedad con las reglas obtenidas del árbol de decisión correspondiente al trastorno de ansiedad generalizada	114
5.3 Comparación de resultados con escalas de valoración	118
Capítulo 6. Conclusiones y trabajos futuros	124
6.1 Conclusiones	124
6.2 Logros obtenidos	126
6.3 Trabajos futuros	127
Referencias	129
Anexo I	139
Cuestionario de preocupación del estado de Pensilvania	
Anexo II	140
Escala de detección del trastorno de ansiedad generalizada	
Anexo III	141
Cuestionario de preocupación del estado de Pensilvania (16 ítems)	
Anexo IV	142
Escala de valoración del trastorno de ansiedad generalizada (6 ítems)	
Anexo V	143
Consentimiento informado	

Índice de figuras

Figura 1.1 Diagrama de flujo: revisión sistemática	9
Figura 1.2 Metodología de solución	19
Figura 2.1 Sistema Nervioso: Sistema Nervioso Parasimpático y Sistema Nervioso Simpático	25
Figura 3.1 Fase 1 de la metodología de solución	33
Figura 3.2 Datos estadísticos pertenecientes al cuestionario de preocupación del estado de Pensilvania	40
Figura 3.3 Datos estadísticos pertenecientes a la escala de detección del trastorno de ansiedad generalizada	40
Figura 3.4 Prueba de Kaiser-Meyer-Olkin y esfericidad de Barlett	42
Figura 3.5 Matriz anti-imagen perteneciente al cuestionario de preocupación del estado de Pensilvania	44
Figura 3.6 Definición de factores y matriz de patrones pertenecientes al cuestionario de preocupación del estado de Pensilvania	45
Figura 3.6(a) Número de componentes	45
Figura 3.6(b) Matriz de patrones	45
Figura 3.7 Porcentajes de varianza perteneciente al cuestionario de preocupación del estado de Pensilvania	46
Figura 3.8 Prueba de Kaiser-Meyer-Olkin y esfericidad de Barlett	46
Figura 3.9 Matriz anti-imagen perteneciente a la escala de detección del trastorno de ansiedad generalizada	48
Figura 3.10 Definición de factores y matriz de patrones pertenecientes a la escala de detección del trastorno de ansiedad generalizada	49
Figura 3.10(a) Número de componentes	49
Figura 3.10(b) Matriz de patrones	49
Figura 3.11 Porcentajes de varianza perteneciente a la escala de detección del trastorno de ansiedad generalizada	50
Figura 3.12 Coeficiente alfa de Cronbach perteneciente al cuestionario de preocupación del estado de Pensilvania	51
Figura 3.13 Coeficiente alfa de Cronbach perteneciente a la escala de detección del trastorno de ansiedad generalizada	51
Figura 4.1 Fase 2 de la metodología de solución	53
Figura 4.2 Datos atípicos correspondientes a la frecuencia cardíaca	66
Figura 4.3 Niveles de actividad por día de la semana	67
Figura 4.4 Obtención de medidas centrales	69

Figura 4.5 Eficiencia de los ciclos del sueño	70
Figura 4.6 Porcentaje de las etapas de sueño	72
Figura 4.7 Valores pertenecientes a la frecuencia cardíaca	78
Figura 4.7(a) Valores pertenecientes a la frecuencia cardíaca como valores perdidos etiquetados como NAN	78
Figura 4.7(b) Valores perdidos reemplazados por la media de la frecuencia cardíaca: 73.5	78
Figura 4.8 Comando correspondiente al reemplazo de los valores etiquetados como NAN	79
Figura 4.9 Estadística descriptiva perteneciente a los registros de la frecuencia cardíaca	80
Figura 4.10 Distribución de los datos correspondientes a los registros de la frecuencia cardíaca	81
Figura 4.11 Tabla correspondiente a los registros de la frecuencia cardíaca	81
Figura 4.11(a) Transformación en el formato de la fecha	81
Figura 4.11(b) Agrupación de la media por día perteneciente a la frecuencia cardíaca	81
Figura 4.11(c) Agrupación de la desviación estándar por día perteneciente a la frecuencia cardíaca	81
Figura 4.12 Proceso de discretización de los datos: antes y después	82
Figura 4.12(a) Antes del proceso de discretización de los datos	82
Figura 4.12(b) Después del proceso de discretización de los datos	82
Figura 4.13 Datos perdidos etiquetados automáticamente como NAN	83
Figura 4.14 Reemplazo de los datos perdidos por la media	83
Figura 4.15 Valores correspondientes al primer, segundo y tercer cuartil, rango intercuartílico, y valores: inferior y superior	85
Figura 4.16 Comparación de diagramas de cajas y bigotes correspondientes a la frecuencia cardíaca	86
Figura 4.17 Normalización de los datos	88
Figura 4.18 Algoritmo para obtener el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada	93
Figura 4.19 Comando utilizado para la obtención de los atributos relevantes	95
Figura 4.20 Selección de atributos relevantes para el trastorno generalizado	95
Figura 4.21 Modelo de comportamiento obtenido para el trastorno de ansiedad generalizada	98
Figura 4.22 Pseudocódigo para la construcción del árbol de decisión	105

usando el valor gini	
Figura 4.23 Diagrama de flujo perteneciente al modelo de comportamiento	106
Figura 4.24 Resultados obtenidos con las métricas para el trastorno de ansiedad generalizada	108
Figura 5.1 Fase 3 perteneciente a la metodología de solución	109
Figura 5.2 Reglas obtenidas del modelo correspondiente al trastorno de ansiedad generalizada	112
Figura 5.3 Pseudocódigo propuesto para la obtención del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada	114

Índice de tablas

Tabla 1.1 Palabras clave utilizadas en la revisión sistemática	8
Tabla 1.2 Cadena de búsqueda genérica	8
Tabla 1.3 Resultados de artículos obtenidos por base de datos	9
Tabla 1.4 Resultados de artículos seleccionados para la revisión sistemática	10
Tabla 1.5 Resultados clasificados por tecnología utilizada	14
Tabla 1.6 Resultados de artículos clasificados por atributos monitoreados	15
Tabla 1.7 Resultados de artículos clasificados como revisiones sistemáticas	16
Tabla 3.1 Rangos de ansiedad generalizada según el cuestionario de preocupación del estado de Pensilvania	37
Tabla 3.2 Rangos de ansiedad generalizada según la escala de detección del trastorno de ansiedad generalizada	38
Tabla 5.1 Resultados del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada obtenidos con las reglas proporcionadas por el algoritmo árbol de decisión	116
Tabla 5.2 Resultados obtenidos de la comparación por día entre el cuestionario de preocupación del estado de Pensilvania y los reportados por el modelo de comportamiento pertenecientes a los 3 participantes	118

Índice de fórmulas

Fórmula 1 para la obtención del valor gini para la división de atributos en un árbol de decisión	99
Fórmula 2 para obtener la división de nodos/ramas perteneciente al modelo de comportamiento del trastorno de ansiedad generalizada	100
Fórmula 3 para obtener la división del nodo raíz perteneciente al atributo frecuencia cardíaca	100
Fórmula 4 para obtener la división del subnodo perteneciente al atributo ciclo del sueño (<i>efficiency</i>)	100
Fórmula 5 para obtener la división del subnodo perteneciente al número de pasos dados (<i>steps</i>)	101
Fórmula 6 para obtener la división del subnodo perteneciente al atributo minutos dormidos en etapa de sueño REM (<i>sleep_REM</i>)	101
Fórmula 7 para obtener la división del subnodo perteneciente al atributo minutos en actividad ligeramente activa (<i>light_min</i>)	101

Glosario de variables

Variable identificador

ID 11: participante para la obtención del modelo de comportamiento

ID 10: participante para la obtención del modelo de comportamiento

ID 100: participante para la obtención del modelo de comportamiento

ID 101: participante para la obtención del modelo de comportamiento

ID 110: participante para la obtención del modelo de comportamiento

ID 1000: participante para la evaluación del modelo de comportamiento

ID 1001: participante para la evaluación del modelo de comportamiento

ID 1010: participante para la evaluación del modelo de comportamiento

Variabes sociodemográficas

Sex: sexo (femenino=0, masculino=1)

Grade: grado (maestría=0)

Edad: 23-30 (23-25=0, 26-30=1)

Variabes frecuencia cardíaca

HR_Mean: media de la frecuencia cardíaca

HR_SD: desviación estándar de la frecuencia cardíaca

Variabes calidad del sueño

Duration: duración del ciclo del sueño (dada en milisegundos)

Efficiency: eficiencia del ciclo del sueño (1%-100%)

Minutes Asleep (M_Asleep): Minutos dormidos

Minutes Awake (M_Awake): minutos despierto

Sleep_Deep: minutos dormidos en etapa sueño profundo

Sleep_Light: minutos dormidos en etapa sueño ligero

Sleep_REM: minutos dormidos en etapa sueño REM

Sleep_stage_Wake: minutos despiertos durante el ciclo del sueño

Sleep_stage_Wake_Count: número de veces que hubo una interrupción en el ciclo del sueño

Time in Bed (T_Bed): tiempo transcurrido en cama (dada en minutos)

Variabes actividad física

Calories_total: total de calorías quemadas

Calories_bmr: calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal)

Steps: número de pasos dados
Floors: pisos escalados
Distance_mile: distancia total recorrida (dada en millas)
Elevation_foot: elevación de los pasos
Sedendary_min: minutos en actividad sedentaria
Light_min: minutos en actividad ligeramente activa
Moderate_min: minutos en actividad moderadamente activa
Very_min: minutos en actividad muy activa

Variables predictoras (X)

Sex: sexo (femenino=0, masculino=1)
Grade: grado (maestría=0, doctorado=1)
Edad: 23-30 (23-25=0, 26-30=1)
HR_Mean: media de la frecuencia cardíaca
HR_SD: desviación estándar de la frecuencia cardíaca
Duration: duración del ciclo del sueño (dada en milisegundos)
Efficiency: eficiencia del ciclo del sueño (1%-100%)
Minutes Asleep (M_Asleep): Minutos dormidos
Minutes Awake (M_Awake): minutos despierto
Sleep_Deep: minutos dormidos en etapa sueño profundo
Sleep_Light: minutos dormidos en etapa sueño ligero
Sleep_REM: minutos dormidos en etapa sueño REM
Sleep_stage_Wake: minutos despiertos durante el ciclo del sueño
Sleep_stage_Wake_Count: número de veces que hubo una interrupción en el ciclo del sueño
Time in Bed (T_Bed): tiempo transcurrido en cama (dada en minutos)
Calories_total: total de calorías quemadas
Calories_bmr: calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal)
Steps: número de pasos dados
Floors: pisos escalados
Distance_mile: distancia total recorrida (dada en millas)
Elevation_foot: elevación de los pasos
Sedendary_min: minutos en actividad sedentaria
Light_min: minutos en actividad ligeramente activa
Moderate_min: minutos en actividad moderadamente activa
Very_min: minutos en actividad muy activa

Variable objetivo (y)

Anxiety_Level: nivel de ansiedad (bajo=1, medio=2, alto=3)

Capítulo 1

Introducción

1.1 Introducción/Motivación

En la actualidad, los trastornos mentales comunes están en aumento. Un ejemplo es la ansiedad, sexta causa de discapacidad laboral en el mundo (Rivera, 2017). De acuerdo con datos de la Organización Mundial de la Salud entre 1990 y 2013, el número de personas que padecen ansiedad o depresión ha aumentado cerca de un 50% (Organización Mundial de la Salud (OMS), 2016), afectando así a 264 millones de personas en el mundo (Elgendi & Menon, 2019). De acuerdo con el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V), los desórdenes de ansiedad incluyen trastornos que comparten un miedo excesivo y una alteración en el comportamiento (American Psychiatric Association, 2013).

El trastorno de ansiedad generalizada forma parte de los trastornos de ansiedad (Pintelas, Kotsilieris & Livieris, 2018); además de ser considerado como el trastorno emocional más básico; porque se compone de características propias de una preocupación crónica y del afecto negativo¹ que están presentes en diversos grados en todos los trastornos emocionales (González, Ibáñez & García, 2018). Este tipo de trastorno debe producirse más días de los que ha estado ausente durante un mínimo de 6 meses en relación con diversos sucesos, o actividades (laboral, académica o escolar) (American Psychiatric Association, 2013), afectando así la calidad de vida del individuo.

Los trastornos de ansiedad se caracterizan por inducir cambios fisiológicos en el cuerpo como la frecuencia cardíaca, la respiración, la temperatura corporal, la sudoración, los escalofríos, los temblores, las náuseas, la tensión muscular y el mareo. No obstante, es difícil percibir cuando un evento es estresante para el individuo en términos fisiológicos y psicológicos (Zheng, Wong, Leung & Poon, 2016).

Con la evolución de la tecnología y de las ciencias computacionales, es posible monitorear estos cambios a través de dispositivos, dispositivos vestibles y sensores. Los dispositivos como tecnologías médicas prometen ayudar a las

1. La dimensión de afecto negativo hace referencia a un malestar subjetivo que representa el factor común entre varios estados emocionales aversivos, como el miedo, la tristeza, la ira, el disgusto y la culpa (Watson, Clark & Stasik, 2011).

personas a buscar un estilo de vida más saludable proporcionando datos médicos continuos para el seguimiento activo del estado metabólico (Yetisen, Martinez-Hurtado & Ünal, 2018). Este tipo de tecnologías ofrecen la ventaja de habilitar un monitoreo en pacientes en riesgo, interviniendo oportunamente a detectar enfermedades en etapas tempranas, y reducir gastos sanitarios mediante la predicción y prevención de enfermedades. Por otro lado, las tecnologías computacionales ubicuas contienen microprocesadores, sensores y aplicaciones móviles que incluyen comunicaciones *wireless* para almacenar datos médicos en tiempo real, y mandar información a otros dispositivos o bases de datos centralizadas. Gracias a la combinación de estas tecnologías es posible facilitar la administración de trastornos de ansiedad permitiendo a los pacientes y a sus doctores evaluar los cambios en las respuestas de déficit cognitivo conductual en tiempo real (Yetisen et al, 2018).

1.2 Planteamiento del problema

Los trastornos por ansiedad son problemas habituales de salud mental que tienen un efecto negativo en la vida de una persona; ya que se encuentra abrumada por un ciclo de preocupaciones que afectan sus actividades diarias y su bienestar. Estas preocupaciones pueden deberse a cuestiones de dinero, salud, familia, o trabajo (Oficina de Información Científica y Tecnológica para el Congreso de la Unión (INCyTU), 2018).

En México, (Medina-Mora, Borges & Lara Muñoz, 2003) la Encuesta Nacional de Epidemiología Psiquiátrica reportó que el 28.6% de la población mexicana presentó alguno de los 23 trastornos mentales de la Clasificación Internacional de Enfermedades alguna vez en su vida, entre estos 23 trastornos mentales se encuentra incluido el trastorno de ansiedad generalizada. Por otro lado, se ha estimado que el trastorno de ansiedad generalizada está presente alguna vez en la vida de la persona entre un 0.7% de incidencia; mientras que se reporta un 1.2% de incidencia en los últimos 12 meses (Padros-Blazquez, Gonzalez-Betanzos & Martinez-Medina, 2018). Los trastornos de ansiedad se han convertido en uno de los desafíos más importantes en México; ya que la tasa de incidencia es del 14.3%; es decir, el 14.3% de la población en general ha llegado a presentar trastornos de ansiedad alguna vez en su vida. La región conformada

por las tres áreas metropolitanas (Ciudad de México, Guadalajara y Monterrey) muestra un 3.4% de incidencia en trastornos de ansiedad, siendo el área con más prevalencia en cuanto a trastornos de ansiedad (Medina-Mora et al, 2003).

En México, un primer problema es que no hay concientización sobre los trastornos mentales, y desafortunadamente, muchas personas que sufren de estos trastornos no buscan ayuda, lo que favorece a la cronificación del trastorno. También, hay ocasiones en que las personas no saben que sufren de algún trastorno de ansiedad, o por el contrario, hay personas que saben que sufren de estos trastornos; pero no buscan ayuda por temor a ser estigmatizados por su propia familia, amigos, y la sociedad en general (American Psychiatric Association, 2005).

En este aspecto, las nuevas tecnologías, así como las herramientas computacionales, pueden aportar una gran ayuda al detectar y predecir, gracias a dispositivos tecnológicos comerciales y de investigación, problemas muy diversos que monitorean la salud de las personas (Ruiz, Aufran & Balderas, 2016).

La predicción de enfermedades y trastornos mentales son áreas que se han estado investigando cada vez con más frecuencia. En este sentido, temas como el análisis de datos, han tomado gran relevancia en el campo de la investigación; ya que ofrecen una amplia gama de soluciones a través de tecnologías para detectar, y monitorear, estados de salud relacionados a enfermedades y trastornos mentales en tiempo real.

Detectar la ansiedad con tecnología y técnicas recomendadas en la literatura; tales como técnicas de minería de datos, sensores vestibles, adquisición de señales a través de electroencefalogramas y electrocardiogramas; representa un gran desafío; por lo que se necesita tener cuidado al estar trabajando con temas relacionados a la salud monitoreando signos vitales y síntomas fisiológicos. En este aspecto, siempre se debe buscar un coeficiente alto en precisión en las herramientas tecnológicas con las que se está trabajando, y de preferencia, elegir aquellos dispositivos que se encuentran validados oficialmente por organismos gubernamentales y privados, así como agencias federales.

El aprendizaje automático, comúnmente llamado en inglés machine learning; junto con las técnicas de minería de datos; pueden ser utilizadas como una herramienta de apoyo para la toma de decisiones en el complejo entorno del análisis de datos en temas relacionados a la salud. Algunos de los algoritmos utilizados que destacan en la literatura y que han otorgado buenos resultados son: redes bayesianas, redes neuronales artificiales, máquinas de soporte vectorial, árboles de decisión, regresiones lineales y sistemas neuro-difusos (Pintelas et al, 2018).

De esta forma, se puede otorgar una solución para ayudar a las personas con trastornos de ansiedad; en específico, con el trastorno de ansiedad generalizada; a detectar cuándo están sufriendo un cuadro ansioso a través del monitoreo de su estado de salud con dispositivos vestibles de una manera no intrusiva, evitando así, que se sientan atemorizados por ser estigmatizados al presentar un trastorno mental.

Esta investigación propone un modelo para la detección de ansiedad generalizada a través de un monitoreo de la frecuencia cardíaca, las horas de sueño dormidas y la actividad física con dispositivos vestibles comerciales. Los datos recolectados de estos dispositivos se registran en una base de datos donde se procesan, se analizan, y se genera el modelo de comportamiento correspondiente al trastorno de ansiedad generalizada con el algoritmo de clasificación árbol de decisión. De esta manera se obtienen los atributos relevantes y el modelo de comportamiento que determina el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.

1.3 Objetivos

En esta sección se presentan el objetivo general y los objetivos específicos que forman parte de esta investigación.

1.3.1 Objetivo general

Desarrollar un modelo de comportamiento para adultos jóvenes que englobe respuestas de déficit cognitivo conductual relacionadas al trastorno de ansiedad

generalizada a través del monitoreo de estas respuestas con sensores vestibulares para la detección automática de cuadros ansiosos.

1.3.2 Objetivos específicos

Para lograr el objetivo principal, se identifican los siguientes objetivos específicos:

1. Validar los instrumentos psicológicos a utilizar en la investigación (cuestionario de preocupación del estado de Pensilvania y la escala de detección del trastorno de ansiedad generalizada) para detectar el trastorno de ansiedad generalizada y para la formación de grupos de conveniencia.
2. Obtener un modelo de comportamiento que identifique los atributos más relevantes asociados a los niveles de ansiedad reportados por los instrumentos psicológicos correspondientes al trastorno de ansiedad generalizada.
3. Evaluar el modelo de comportamiento obtenido a través de las métricas de rendimiento: exactitud, sensibilidad, precisión, valor F-1 y matriz de confusión, pertenecientes al aprendizaje automático. Además de su evaluación con pruebas de monitoreo.
4. Evaluar el modelo de comportamiento con nuevos casos monitoreados y con los resultados obtenidos por el cuestionario de preocupación del estado de Pensilvania.

1.4 Estado del arte

El estado del arte se llevó a cabo realizando una revisión sistemática para localizar las investigaciones que abordan el tema de la detección de ansiedad a través de dispositivos, algoritmos de aprendizaje automático y minería de datos.

1.4.1 Revisión sistemática

Una revisión sistemática es una característica esencial de una investigación académica; es decir, es una metodología científica y avanzada que se utiliza para conocer, evaluar, resumir, analizar y sintetizar la información existente en un determinado campo de investigación. Idealmente las revisiones sistemáticas

deben ser conducidas antes de la investigación empírica y como un subconjunto de la literatura encontrada en la revisión sistemática (Xiao & Watson, 2017).

1.4.1.1 Formular el problema de investigación

La revisión sistemática que se desarrolló para esta investigación hace uso de la checklist PRISMA y del diseño de estudio formato PICOC por sus elementos en inglés (Patient/Population/Problem, Intervention, Comparison, Outcomes, Context: Paciente/población, Intervención, Comparación, Resultados, Contexto). El diseño de estudio o mejor conocido como formato PICOC favorece el desarrollo de las revisiones sistemáticas de forma clara y ordenada; además de centrar las revisiones en preguntas precisas. Para el caso de esta investigación el elemento P del formato PICOC se definió como problema, refiriéndose así a los artículos o investigaciones pertenecientes a la detección o predicción de la ansiedad.

El formato de estudio PICOC para esta revisión sistemática se define de la siguiente manera:

- **P:** artículos que hablen de la detección o predicción de la ansiedad con ayuda de herramientas tecnológicas;
- **I:** técnicas o métodos implementados para detectar o predecir la ansiedad;
- **C:** comparación de técnicas o métodos implementados para detectar o predecir la ansiedad;
- **O:** evaluación de los métodos propuestos;
- **C:** métodos y técnicas utilizando mecanismos computacionales para predecir o detectar la ansiedad.

La revisión sistemática de esta investigación se centra en dar respuesta a la siguiente pregunta de investigación, de acuerdo con el formato de estudio PICOC definido anteriormente:

¿Qué tipos de dispositivos vestibles existen que ayuden a detectar computacionalmente las respuestas de déficit cognitivo conductual asociadas al trastorno de ansiedad generalizada?

1.4.1.2 Desarrollo del protocolo de la revisión sistemática

El objetivo de esta revisión sistemática es buscar el estado actual de la literatura para la detección del trastorno de ansiedad generalizada a través de dispositivos y sensores vestibles. Para lograr esto se definieron los siguientes elementos:

1. Criterios de inclusión

Los criterios de inclusión que se tomaron en cuenta para esta revisión fueron artículos de revista y artículos de conferencia publicados entre los años 2013 y 2019 en el idioma inglés. Además, solamente se tomaron en cuenta artículos pertenecientes al área de computación, o de ciencias de la información que detectan, predicen o monitorean respuestas de déficit cognitivo conductual relacionadas al trastorno de ansiedad generalizada a través del uso de la tecnología o de sensores vestibles.

2. Criterios de exclusión

Los criterios de exclusión de esta revisión sistemática no tomaron en cuenta artículos publicados antes del 2013, o artículos en un idioma diferente al inglés. También se excluyeron artículos que pertenecen a otras áreas diferentes a las ciencias computacionales.

3. Fuentes de información

Las fuentes de información utilizadas para realizar esta revisión sistemática fueron las bases de datos ACM, IEEE y Springer Link.

4. Estrategias de búsqueda

La estrategia de búsqueda se inició con la identificación de las palabras clave en idioma inglés, así como posibles sinónimos que ayudaron a facilitar la búsqueda en la literatura (ver Tabla 1.1).

Para esta revisión, las palabras clave se clasificaron en tres grupos:

- i) palabras clave referente a los trastornos de ansiedad (ansiedad generalizada);
- ii) palabras clave referente a la detección, predicción o monitoreo del trastorno de ansiedad o del trastorno de ansiedad generalizada;

iii) palabras clave referentes a la tecnología utilizada para la detección, predicción o monitoreo del trastorno de ansiedad o del trastorno de ansiedad generalizada.

Tabla 1.1. Palabras clave utilizadas en la revisión sistemática.

Grupo	Palabras clave
Trastornos de ansiedad	Anxiety, Generalized anxiety, Generalized anxiety disorder, GAD
Detección/predicción/monitoreo	Detect, Detection, Monitoring, Predict, Prediction, Predicting
Tecnología utilizada	Sensors, Wearable devices, Devices

Con las palabras clave descritas anteriormente, se procedió a generar la cadena de búsqueda genérica separando cada concepto con operadores lógicos OR y AND. La cadena de búsqueda resultante es presentada en la Tabla 1.2.

Tabla 1.2. Cadena de búsqueda genérica.

(“anxiety” OR “generalized anxiety” OR “generalized anxiety disorder” OR “GAD”) AND (“detect” OR “detection” OR “monitoring” OR “predict” OR “prediction” OR “predicting”) AND (“sensors” OR “wearable devices” OR “devices”)

Al llevar a cabo la búsqueda en las bases de datos ACM, IEEE y Springer Link, se obtuvo un total de 3,642 artículos. De la base de datos AMC se obtuvieron 1,495 artículos. De la base de datos IEEE se obtuvieron 80 artículos; mientras que de la base de datos Springer Link se obtuvieron 2,067 artículos. En esta primera etapa se eliminaron 168 artículos duplicados, obteniendo 3,474 artículos. La Tabla 1.3 muestra los resultados obtenidos por base de datos: ACM, IEEE y Springer Link. La Figura 1.1 muestra el diagrama de flujo con los pasos realizados para esta revisión sistemática.

Tabla 1.3. Resultados de artículos obtenidos por base de datos.

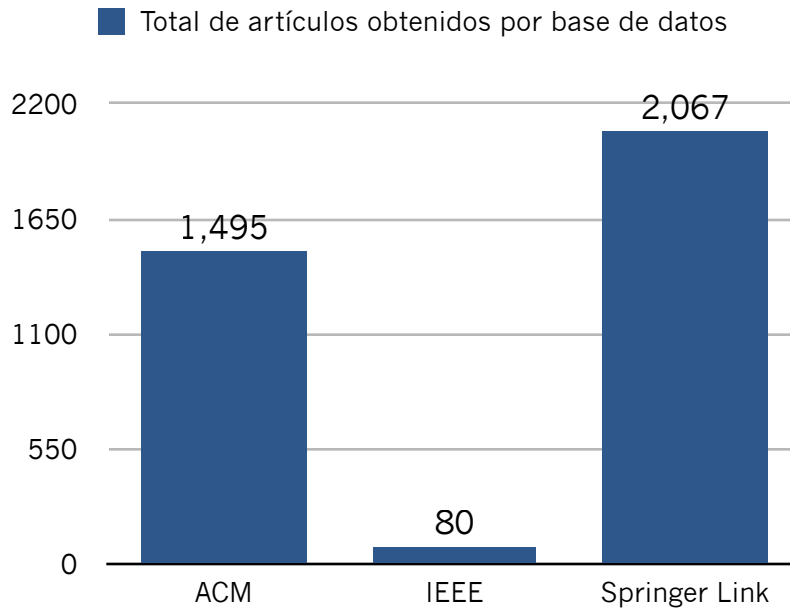


Diagrama de flujo: Revisión sistemática

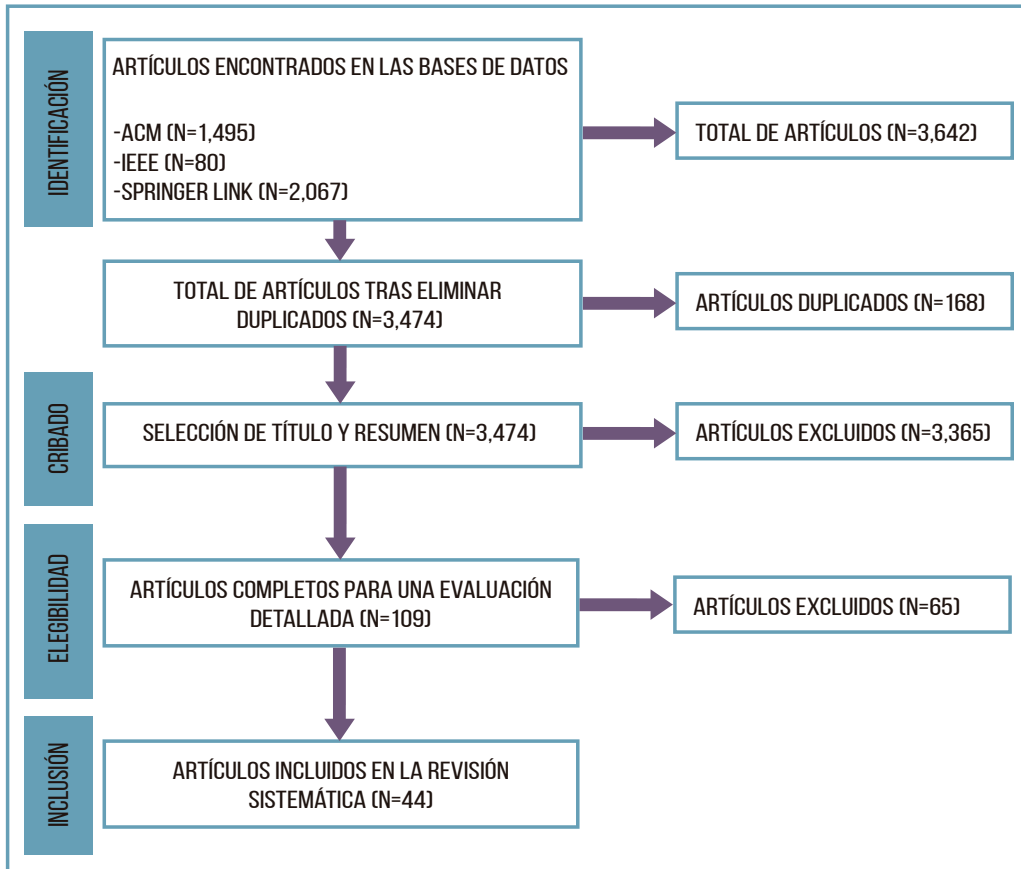


Figura 1.1. Diagrama de flujo con pasos realizados en la revisión sistemática.

La revisión sistemática se realizó a cabo a finales de diciembre del año 2019; por lo que, estudios posteriores a esta fecha pueden no haber sido considerados para la revisión. A partir de los 3,642 artículos obtenidos, los criterios de inclusión y exclusión fueron aplicados, así como la eliminación de artículos duplicados en las bases de datos.

El cribado se realizó leyendo los títulos, los resúmenes y las palabras clave de cada artículo obtenido en la búsqueda, dando como resultado 109 artículos que se les dio lectura con la finalidad de descartar aquellos que hablan de la detección de ansiedad; pero no muestran alguna tecnología vestible para monitorearla o predecirla. A partir de los 109 artículos seleccionados, se concluyó que solamente 44 artículos daban respuesta a la pregunta de investigación planteada en el punto 1.4.1.1 *Formular el problema de investigación*. Los resultados se presentan en la Tabla 1.4, donde los encabezados de la tabla denotan:

- ID
- Año
- Título
- Autores

Tabla 1.4. Resultados de artículos seleccionados para la revisión sistemática.

ID	Año	Título	Autores
1	2013	Mental health and the impact of ubiquitous technologies	Bert Arnrich, Venet Osmani, Jakob Bardram
2	2013	A mobile data collection platform for mental health research	Andrea Gaggioli, Giovanni Pioggia, Gennaro Tartarisco, Giovanni Baldus, Daniele Corda, Pietro Cipresso, Giuseppe Riva
3	2013	Virtual reality and mobile phones in the treatment of generalized anxiety disorders: a phase-2 clinical trial	Claudia Repetto, Andrea Gaggioli, Federica Pallavicini, Pietro Cipresso, Simona Raspelli, Giuseppe Riva
4	2014	Anxiety detection using wearable monitoring	Darién Miranda, Marco Calderón, Jesus Favela
5	2014	Short-term anxiety recognition from blood volume pulse signal	W. Handouzi, C. Maaoui, A. Pruski, A. Moussaoui
6	2014	Physiological measurement of anxiety to evaluate performance in simulation training	Jennifer G. Tichon, Guy Wallis, Stephan Riek, Timothy Mavin

ID	Año	Título	Autores
7	2014	Sense of Presence and Metacognition Enhancement in Virtual Reality Exposure Therapy in the Treatment of Social Phobias and the Fear of Flying	Ioannis Paliokas, Athanasios Tsakiris, Athanasios Vidalis, Dimitrios Tzovaras
8	2015	Towards a mobile and wearable system for predicting panic attacks	Jonathan Rubin, Hoda Eldardiry, Rui Abreu, Shane Ahern, Honglu Du, Ashish Pattekar, Daniel G. Bobrow
9	2015	A Social Robot to Mitigate Stress, Anxiety, and Pain in Hospital Pediatric Care	Sooyeon Jeong, Deirdre E. Logan, Matthew S. Goodwin, Suzanne Graca, Brianna O'Connell, Honey Goodenough, Laurel Anderson, Nicole Stenquist, Katie Fitzpatrick, Miriam Zisook, Luke Plummer, Cynthia Breazeal, Peter Weinstock
10	2015	A wearable and mobile intervention delivery system for individuals with panic disorder	Luis Cruz, Jonathan Rubin, Rui Abreu, Shane Ahern, Hoda Eldardiry, Daniel G. Bobrow
11	2015	Displayed Uncertainty Improves Driving Experience and Behavior: The Case of Range Anxiety in an Electric Car	Malte F. Jung, David Sirkin, Turgut M. Gür, Martin Steinert
12	2015	A Kalman Filtering Framework for Physiological Detection of Anxiety-Related Arousal in Children With Autism Spectrum Disorder	A. Kushki, A. Khan, J. Brian, E. Anagnostou
13	2015	Mobile phones as medical devices in mental disorder treatment: an overview	Franz Gravenhorst, Amir Muaremi, Jakob Bardram, Agnes Grünerbl, Oscar Mayora, Gabriel Wurzer, Mads Frost, Venet Osmani, Bert Arrnich, Paul Lukowicz, Gerhard Tröster
14	2015	A comparative study of game mechanics and control laws for an adaptive physiological game	Avinash Parnandi, Ricardo Gutierrez-Osuna
15	2015	Detecting State Anxiety When Caring for People with Dementia	Darien Miranda, Jesus Favela, Catalina Ibarra
16	2016	Good vibes: the impact of haptic patterns on stress levels	Chelsea Kelling, Daniella Pitaro, Jussi Rantala
17	2016	Unobtrusive and Multimodal Wearable Sensing to Quantify Anxiety	Y. Zheng, T. C. H. Wong, B. H. K. Leung, C. C. Y. Poon
18	2016	Investigation of Multimodal Mobile Applications for Improving Mental Health	Sushunova G. Martinez, Karla A. Badillo-Urquiola, Rebecca A. Leis, Jamie Chavez, Tiffany Green, Travis Clements
19	2016	Using Mobile-Based Games as a Means for the Self-treatment of Depression and Anxiety in Youth	Scott Cabot, Brett Wilkinson
20	2017	Movements Indicate Threat Response Phases in Children at Risk for Anxiety	E. W. McGinnis, R. S. McGinnis, M. Muzik, J. Hruschak, N. L. Lopez-Duran, N. C. Perkins, K. Fitzgerald, K. L. Rosenblum
21	2017	Wearable Social Sensing and Its Application in Anxiety Assesment	J. Gu, B. Gao, Y. Chen, L. Jiang, Z. Gao, X. Ma, Y. Ma, W. L. Woo
22	2017	A wearable, EEG-based massage headband for anxiety alleviation	M. Nair Chaitanya, S. Jayakkumar, E. Chong, C. H. Yeow

ID	Año	Título	Autores
23	2017	Multi-level assessment model for wellness service based on human mental stress level	Yuchae Jung, Yong Ik Yoon
24	2017	On-time Measurement of Subjective Anxiety of a Passenger in an Autonomous Vehicle: Gradually Changing Sounds Decreases Anxiety of Passenger	Akitoshi Tomita, Etsuko T. Harada, Satoshi Ando, Kozue Miyashiro, Maito Ohmori, Hiroaki Yano
25	2017	Virtual Sophrologist: A Virtual Reality Neurofeedback Relaxation Training System	Guoxin Gu, Claude Frasson
26	2017	Racing Heart and Sweaty Palms	Mathieu Chollet, Talie Massachi, Stefan Scherer
27	2018	A review of machine learning prediction methods for anxiety disorders	Emmanuel G. Pintelas, Theodore Kotsilieris, Ioannis E. Livieris, Panagiotis Pintelas
28	2018	IoT-based healthcare system for real-time maternal stress monitoring	Olugbenga Oti, Iman Azimi, Arman Anzanpour, Amir M. Rahmani, Anna Axelin, Pasi Liljeberg
29	2018	Wearable sensors and machine learning diagnose anxiety and depression in young children	R. S. McGinnis, E. W. McGinnis, J. Hruschak, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, M. Muzik
30	2018	Rapid Anxiety and Depression Diagnosis in Young Children Enabled by Wearable Sensors and Machine Learning	R. S. McGinnis, E. W. McGinnis, J. Hruschak, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, M. Muzik
31	2018	Analyzing the changes of health condition and social capital of elderly people using wearable devices	Siyu Zhou, Atsushi Ogihara, Shoji Nishimura, Qun Jin
32	2019	Detecting Temporal Phases of Anxiety in The Wild: Toward Continuously Adaptive Self-Regulation Technologies	Hashini Senaratne
33	2019	An Investigation of Gender Differences in Computer Science Using Physiological, Psychological and Behavioural Metrics	Keith Nolan, Aidan Mooney, Susan Bergin
34	2019	HCI and Affective Health: Taking stock of a decade of studies and charting future research directions	Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, Corina Sas
35	2019	Machine Learning and Therapeutic Strategies in VR	Ray Lc, Yuka Fukuoka
36	2019	Anxiety detection from Electrodermal Activity Sensor with movement & interaction during Virtual Reality Simulation	J. Kritikos, G. Tzannetos, C. Zoitaki, S. Pouloupoulou, D. Koutsouris
37	2019	A Flexible Wearable Device for Measurement of Cardiac, Electrodermal, and Motion Parameters in Mental Healthcare Applications	B. M. G. Rosa, G. Z. Yang
38	2019	Speak Up! Studying the Interplay of Individual and Contextual Factors to Physiological-Based Models of Public Speaking Anxiety	M. Yadav, A. Behzadan, T. Chaspari
39	2019	A Comparative Study of Stress and Anxiety Estimation in Ecological Settings Using a Smart-shirt and a Smart-bracelet	A. Tiwari, R. Cassani, S. Narayanan, T. H. Falk

ID	Año	Título	Autores
40	2019	A Sensor-based Framework for Real-time Detection and Alleviation of Public Speaking Anxiety	E. Kimani
41	2019	Internet of Things based pervasive sensing of psychological anxiety via wearable devices under naturalistic settings	R. Rother, Y. Sun, B. Lo
42	2019	Monitoring Mental Health Using Smart Devices with Text Analytical Tool	N. S. Alghamdi
43	2019	Analysis of Biofeedback Through Heartbeat Obtained by Exposure to Phobia Through Virtual Reality	Edvaldo de Vasconcelos, Amaro Neto, Lillian dos Santos, Paula Ribeiro
44	2019	Biofeedback as a stress management tool: a systematic review	Lauren Kennedy, Sarah Henrickson Parker

1.4.2 Criterios de clasificación

Los artículos que respondieron a la pregunta planteada para esta revisión sistemática se dividieron de la siguiente manera:

- i) tecnología utilizada;
- ii) atributos monitoreados;
- iii) revisiones sistemáticas.

Para el criterio de clasificación tecnología utilizada, se estableció la búsqueda de los diversos dispositivos que monitorean, registran y almacenan datos fisiológicos como la frecuencia cardíaca, la frecuencia respiratoria, la temperatura corporal, entre otros, relacionados a los trastornos de ansiedad. El objetivo de esta clasificación fue elegir el dispositivo vestible que se utilizó en la investigación para la adquisición de datos.

En cambio, para el criterio de clasificación atributos monitoreados, se estableció la búsqueda de aquellas variables fisiológicas (respuestas de déficit cognitivo conductual) asociadas al trastorno de ansiedad que son monitoreadas con más frecuencia por la amplia diversidad de los dispositivos vestibles existentes en el mercado. El objetivo de esta clasificación fue elegir los atributos relacionados al trastorno de ansiedad que pudieran ser monitoreados por los dispositivos vestibles existentes en el mercado, y reportados por los artículos pertenecientes a esta revisión sistemática.

Por último, para el criterio de clasificación revisiones sistemáticas, se estableció la búsqueda de aquellas revisiones que proporcionaban soluciones relacionadas a los trastornos de ansiedad tanto en dispositivos vestibles, como en herramientas de minería de datos; tales como los algoritmos de aprendizaje automático que permiten establecer métodos para detectar o predecir trastornos de ansiedad. El objetivo de esta clasificación fue obtener más opciones de los dispositivos vestibles existentes en el mercado, así como los diversos algoritmos utilizados para la detección o predicción de los trastornos de ansiedad.

1.4.2.1 Tecnología utilizada

Esta clasificación permite identificar cuáles fueron los dispositivos utilizados para detectar, monitorear o predecir las respuestas de déficit cognitivo conductual asociadas al trastorno de ansiedad generalizada. La Tabla 1.5 muestra las investigaciones clasificadas por el tipo de tecnología utilizada para monitorear, detectar, o predecir, aquellas respuestas de déficit cognitivo conductual asociadas al trastorno de ansiedad.

Tabla 1.5. Resultados clasificados por tecnología utilizada.

Tecnología utilizada	Investigaciones	Cantidad
1. Dispositivos vestibles	[4], [8], [10], [11], [15], [16], [17], [19], [26], [28], [31], [33], [38], [39], [42]	14
2. Sensores	[2], [5], [6], [11], [12], [14], [20], [21], [22], [23], [26], [29], [30], [32], [37], [41]	16
3. Combinada (simulación+sensores)	[6]	1
4. Combinada (video juegos+sensores)	[14]	1
5. Combinada (realidad virtual+dispositivos vestibles/sensores)	[3], [5], [7], [9], [24], [25], [26], [35], [36], [40], [43]	12

La tecnología más utilizada en los trabajos presentados en la Tabla 1.5 fueron los sensores, seguida por los dispositivos vestibles y por la combinación entre estas dos últimas tecnologías con la realidad virtual. En cambio, la tecnología menos utilizada fue la combinación de sensores con las simulaciones y los video juegos.

En los artículos [4], [8], [10], [11] y [26], se encontró que los dispositivos vestibles utilizados para monitorear respuestas fisiológicas, son los dispositivos Zephyr (BioPatch) y Empatica 4. Sólo en el artículo [39], se reportó haber trabajado con un dispositivo Fitbit (Fitbit Charge 2). Por otro lado, los sensores más utilizados reportados en los artículos [2], [12] y [37] son los que monitorean la frecuencia cardíaca o la variabilidad de la frecuencia cardíaca. Otros sensores reportados son los que monitorean la actividad electrodermal [11] y [14], el pulso sanguíneo [5], la actividad eléctrica de los músculos [6], y la actividad eléctrica del cerebro [22].

1.4.2.2 Atributos monitoreados

Esta clasificación permite identificar los atributos monitoreados (las respuestas de déficit cognitivo conductual) asociados al trastorno de ansiedad (ver Tabla 1.6), que fueron monitoreados en los artículos pertenecientes a esta revisión sistemática.

Tabla 1.6. Resultados de artículos clasificados por atributos monitoreados.

Atributos monitoreados	Investigaciones	Cantidad
1. Frecuencia cardíaca	[2], [3], [4], [7], [8], [10], [11], [12], [16], [19], [23], [24], [26], [28], [31], [39], [40], [42]	18
2. Variabilidad de frecuencia cardíaca	[2], [4], [8], [10], [12], [15], [28], [33], [39]	9
3. Frecuencia respiratoria	[8], [10], [11], [23], [31], [39]	6
4. Temperatura	[8], [10], [37]	3
5. Actividad electrodermal/Respuesta galvánica de la piel	[3], [9], [11], [14], [15], [26], [33], [35], [36], [37], [38], [40], [41]	13
6. Fotopletismografía	[17], [33], [38], [39]	4
7. Electroencefalograma	[7], [15], [17], [22], [23], [25]	6
8. Electrocardiograma	[2], [12], [23], [37], [39]	5
9. Electromiografía	[6]	1
10. Presión arterial	[23], [31]	2
11. Saturación de oxígeno en la piel	[23]	1

Atributos monitoreados	Investigaciones	Cantidad
12. Sueño	[23], [28], [39]	3
13. Actividad	[10], [23], [28]	3
14. Frecuencia de parpadeo	[4]	1
16. Movimiento	[11], [20], [29], [30], [37]	5
17. Aceleración	[20], [21], [29], [30], [32], [37], [38], [41]	8

El atributo más importante monitoreado en los artículos incluidos en esta revisión sistemática, fue la frecuencia cardíaca, seguido de la actividad electrodermal o la respuesta galvánica de la piel y la variabilidad de la frecuencia cardíaca. Los atributos con menor frecuencia de monitoreo fueron la electromiografía, la saturación de oxígeno en la piel y la frecuencia de parpadeo. La aceleración, el electroencefalograma, el movimiento y el electrocardiograma también fueron atributos que tomaron relevancia en los artículos incluidos en la revisión sistemática. No obstante, algunos de estos últimos atributos suelen ser monitoreados con dispositivos un poco más intrusivos; en el caso de la tecnología perteneciente al electrocardiograma y electroencefalograma. A su vez, la aceleración y el movimiento suelen ser atributos, que, algunas veces no se encuentran disponibles para su monitoreo en dispositivos vestibles no intrusivos como Fitbit o Zephyr.

1.4.2.3 Revisiones sistemáticas

Esta clasificación permite identificar las diferentes investigaciones reportadas como revisiones o revisiones sistemáticas relacionadas a la detección o predicción del trastorno de ansiedad. La Tabla 1.7 muestra las investigaciones clasificadas como revisiones o revisiones sistemáticas.

Tabla 1.7. Resultados de artículos clasificados como revisiones sistemáticas.

Revisiones sistemáticas	Investigaciones	Cantidad
1. Revisiones sistemáticas	[1], [13], [18], [27], [34], [44]	6

Los artículos [18] y [44] reportaron dentro de sus revisiones dispositivos vestibles, y en general, tecnología disponible en el mercado para el posible monitoreo de atributos relacionados a los trastornos de ansiedad. Por otro lado, los artículos [27] y [34], reportaron dentro de sus revisiones: metodologías, atributos monitoreados y la forma de conducir una investigación relacionada a trastornos mentales. Mientras que el artículo [27] reportó investigaciones que hacen uso de algoritmos de aprendizaje automático para dar solución a la predicción de trastornos de ansiedad. Por último, el artículo [1] reportó investigaciones existentes en el estado del arte que dan solución a diversos trastornos mentales. Estas revisiones ayudaron a elegir el dispositivo vestible para la adquisición de datos, así como conducir la investigación y a elegir el algoritmo con el que trabajar para detectar el trastorno de ansiedad generalizada.

1.4.3 Conclusiones

La principal razón de realizar una revisión sistemática para esta investigación, fue seguir un método preciso para elaborar el estado del arte. En la revisión sistemática realizada para esta investigación, se identificaron las tecnologías utilizadas en la literatura, los dispositivos vestibles y sensores más usados para monitorear respuestas fisiológicas, así como también se identificaron las variables que se tomaron en cuenta para realizar el monitoreo de las respuestas de déficit cognitivo conductual asociadas al trastorno de ansiedad. Esta revisión sistemática incluyó la revisión de 3,642 artículos obtenidos, los cuales fueron sometidos a los criterios de inclusión y exclusión definidos anteriormente, resultando 109 artículos relevantes. De estos artículos, se excluyeron 65, obteniendo así 44 artículos que dieron respuesta a la pregunta de investigación planteada anteriormente en la investigación.

Al realizar esta investigación sistemática se observó que, si bien en la literatura se habla sobre trastornos de ansiedad en general, existe una cantidad menor de artículos específicos de los diferentes trastornos de ansiedad, incluido el trastorno de ansiedad generalizada. También se observó que, dentro de las soluciones propuestas, se encuentran aquellas dirigidas a las terapias por exposición, que hacen uso de tecnologías como la realidad virtual o la realidad aumentada. No obstante, algunas de estas soluciones no proponen detectar las respuestas de déficit cognitivo conductual asociadas a los trastornos de ansiedad

con sensores o dispositivos vestibles, ofreciendo, en su gran mayoría, aplicaciones móviles dirigidas a técnicas de relajación.

En cuanto a los dispositivos vestibles, se encontró que los más utilizados son aquellos que no son intrusivos, tales como los relojes inteligentes, dispositivos con forma de pulsera o dispositivos con un tamaño considerable sin llegar a molestar al participante. Un ejemplo son los dispositivos Zephyr, Empatica y Fitbit reportados por los artículos [4], [8], [10], [11], [26] y [39]. Estos dispositivos ofrecen una amplia gama de sensores incorporados en un sólo dispositivo, como el caso del dispositivo Empatica 4 y Fitbit, siendo capaces de monitorear varias respuestas de déficit cognitivo conductual, contando con la gran ventaja de no ser una molestia para los participantes que forman parte de las investigaciones.

1.5 Metodología de solución

La metodología de solución perteneciente a la investigación está dividida en fases que han sido clasificadas en los distintos capítulos pertenecientes a esta tesis. Cada fase de la metodología de solución presenta los procesos realizados a lo largo de la investigación para obtener un resultado, que a su vez, es un elemento importante para iniciar la siguiente fase.

Para el caso de esta investigación, la metodología propuesta se divide en 3 fases:

- i) fase 1: validación de instrumentos psicológicos para la detección del trastorno de ansiedad generalizada;
- ii) fase 2: construcción del modelo de comportamiento del trastorno de ansiedad generalizada;
- iii) fase 3: evaluación del modelo para identificar el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.

Esta metodología de solución se clasificó en tres fases; porque, para desarrollar y evaluar un modelo de comportamiento relacionado al trastorno de ansiedad generalizada, se requiere de la aplicación y validación de instrumentos psicológicos para detectar la población que se ve afectada por este trastorno mental. También, se necesita de la recolección de datos, así como de su

preparación y selección de atributos relevantes para obtener el modelo de comportamiento. No obstante, este modelo de comportamiento se debe evaluar; además de comparar los resultados de esta evaluación con los obtenidos por los instrumentos psicológicos que detectan el trastorno de ansiedad generalizada.

La clasificación de estas tres fases correspondientes a la metodología de solución, se compone de la siguiente manera: la Fase 1 corresponde al Capítulo 3: *Validación de instrumentos psicológicos para la detección del trastorno de ansiedad generalizada*. La Fase 2 corresponde al Capítulo 4: *Modelo de comportamiento del trastorno de ansiedad generalizada*; mientras que la Fase 3 corresponde al Capítulo 5: *Evaluación del modelo y resultados*.

La Figura 1.2 presenta de manera ordenada y detallada cada fase que permite cumplir con el objetivo de esta investigación.

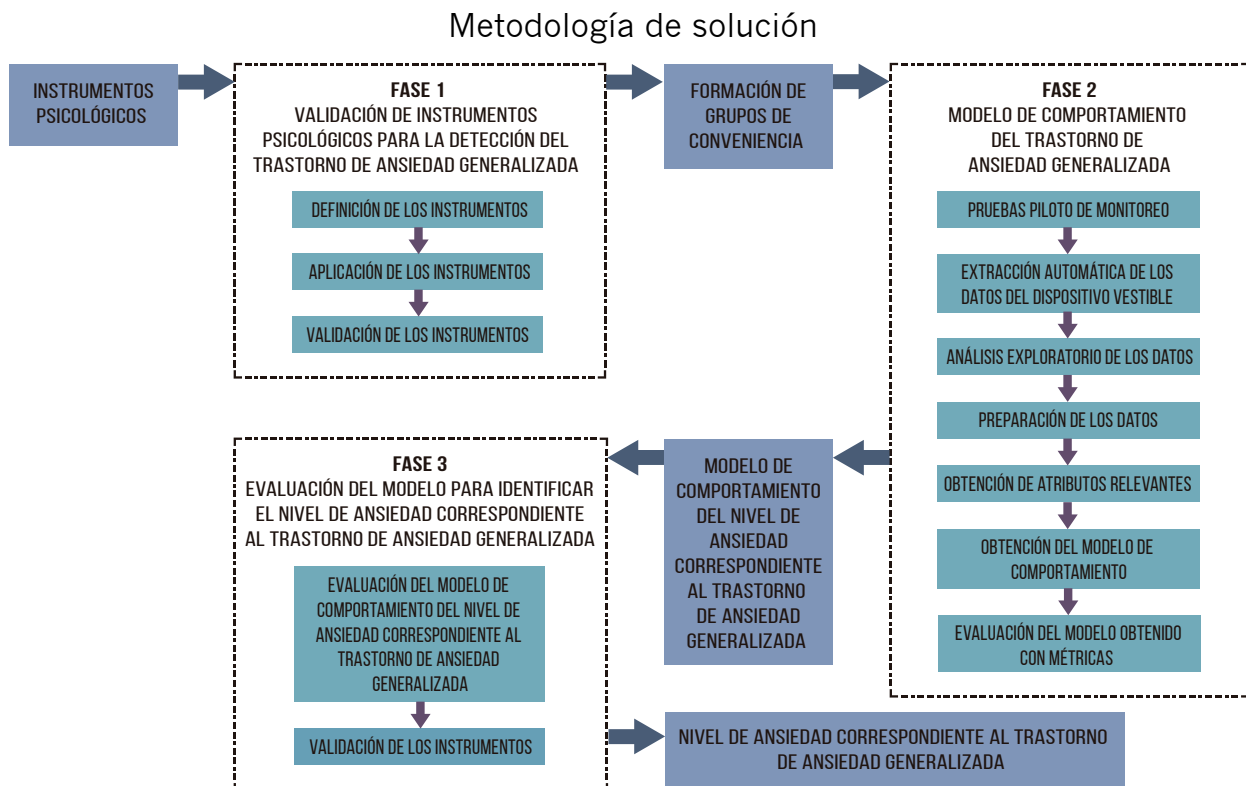


Figura 1.2. Metodología de solución.

A continuación, se describen las fases que componen los principales procesos de la metodología de solución utilizada a lo largo de esta investigación.

Fase 1. Validación de instrumentos psicológicos para la detección del trastorno de ansiedad generalizada

En esta fase se definen, aplican, y validan los instrumentos psicológicos que detectan el trastorno de ansiedad generalizada en una población de adultos jóvenes específica que reside en el estado de Morelos, con el objetivo de formar los grupos de conveniencia de acuerdo a los criterios de inclusión y exclusión definidos para la investigación. Esta fase engloba el capítulo 3 de esta investigación.

Fase 2. Modelo de comportamiento del trastorno de ansiedad generalizada

En esta fase se realizan las pruebas de monitoreo, proporcionando un dispositivo vestible a adultos jóvenes por un periodo establecido de 14 días con el objetivo de obtener un modelo de comportamiento para monitorear respuestas de déficit cognitivo conductual relacionadas al trastorno de ansiedad generalizada. En esta fase es donde se extraen los datos del dispositivo vestible, se realiza el análisis exploratorio de los datos, se preparan los datos, y se obtienen los atributos relevantes, generando así el modelo correspondiente al trastorno de ansiedad generalizada con el algoritmo árbol de decisión. También en esta fase se realiza la evaluación del modelo obtenido con las métricas de rendimiento: exactitud, sensibilidad, precisión, valor F-1 y matriz de confusión. Esta fase engloba el capítulo 4 de esta investigación.

Fase 3. Evaluación del modelo para identificar el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada

En esta fase se evalúa el modelo obtenido en la fase 2 perteneciente al trastorno de ansiedad generalizada. Esta evaluación se logra a través del monitoreo de los atributos relevantes obtenidos por el modelo de comportamiento en la fase 2, y que corresponden a las respuestas de déficit cognitivo conductual relacionadas al trastorno de ansiedad generalizada. En esta fase se comparan los resultados obtenidos por el modelo de comportamiento con los resultados obtenidos por el cuestionario de preocupación del estado de Pensilvania. Esta fase engloba el capítulo 5 de esta investigación.

1.6 Estructura de la tesis

La estructura, en la que está organizada esta investigación, es la siguiente:

Capítulo 2. Marco conceptual. Este capítulo define los conceptos básicos y esenciales que se utilizan a lo largo de la investigación.

Capítulo 3. Validación de instrumentos psicológicos para la detección del trastorno de ansiedad generalizada. Este capítulo presenta la validación de los instrumentos psicológicos utilizados en la investigación (cuestionario de preocupación del estado de Pensilvania y la escala de detección del trastorno de ansiedad generalizada) para la formación de grupos de conveniencia (Meyer, Metzger & Borkovec, 1990), (Carroll & Davidson, 2000).

Capítulo 4. Modelo de comportamiento del trastorno de ansiedad generalizada. Este capítulo describe el proceso para la obtención del modelo de comportamiento haciendo uso del algoritmo de clasificación árbol de decisión y que debe emplearse para determinar el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada. También se realiza la evaluación del modelo con métricas de rendimiento: exactitud, sensibilidad, precisión, valor F-1 y matriz de confusión.

Capítulo 5. Evaluación del modelo y resultados. Este capítulo describe el procedimiento para evaluar el modelo de comportamiento obtenido del trastorno de ansiedad generalizada y los resultados obtenidos, tanto del modelo, como el de un instrumento psicológico definido en la investigación (cuestionario de preocupación del estado de Pensilvania). Así como la comparación de resultados entre este instrumento psicológico de valoración de ansiedad y los obtenidos por el modelo de comportamiento.

Capítulo 6. Conclusiones y trabajos futuros. Este capítulo presenta las conclusiones de la investigación, así como los trabajos futuros .

Capítulo 2

Marco conceptual

En este capítulo se definen los conceptos más relevantes y necesarios para el desarrollo de esta investigación.

2.1 Ansiedad

La ansiedad; además de ser una respuesta emocional provocada por el miedo, es un sistema complejo de respuesta conductual, fisiológica, afectiva y cognitiva a modo de amenaza (Clark & Aaron T., 2012). Biológicamente es una respuesta adaptativa de un sistema de alarma que prepara a un organismo a contender contra un peligro potencial (Pérez de la Mora, 2003).

De acuerdo con la organización Panamericana de la Salud (Organización Panamericana de la Salud, 2015), la ansiedad es provocada exclusiva o predominantemente por ciertas situaciones que no son peligrosas pero que típicamente son evitadas, o soportadas con pavor. No obstante, la ansiedad puede incrementarse con un sentimiento de vergüenza, produciendo confusión y distorsiones de la percepción, no sólo en cuanto al tiempo y el espacio, sino respecto a la gente y al sentido de los diferentes sucesos. Estas distorsiones pueden interferir con el aprendizaje, la concentración, la memoria y la capacidad de hacer asociaciones. Cabe aclarar que la ansiedad por sí misma no es considerada como una enfermedad, si no que se encuentra contemplada en la amplia gama de respuestas emocionales que presenta cualquier persona (Montelongo, Zaragoza & Bonilla, 2005).

2.1.1 Trastorno de ansiedad generalizada

De acuerdo con el Manual diagnóstico y estadístico de los trastornos mentales (DSM-V), el trastorno de ansiedad generalizada se caracteriza por ser una ansiedad y preocupación excesiva que se produce durante más días de los que ha estado ausente durante un mínimo de seis meses, en relación con diversos sucesos o actividades (actividad laboral, escolar o social). Al individuo le es difícil controlar la preocupación y ésta puede estar asociada a tres (o más) de los seis síntomas siguientes:

1. inquietud o sensación de estar atrapado o con los nervios de punta;

2. fácilmente fatigado;
3. dificultad para concentrarse o quedarse con la mente en blanco;
4. irritabilidad;
5. tensión muscular;
6. problemas de sueño (dificultad para dormirse o para continuar durmiendo, o sueño inquieto e insatisfactorio).

Por otro lado, la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE-10), menciona que la característica fundamental del trastorno de ansiedad generalizada es una ansiedad persistente, que no se restringe ni siquiera en términos de algún fuerte predominio, a ninguna circunstancia del entorno en particular (es decir, la angustia flota libremente). Los síntomas principales son variables, pero incluyen quejas de permanente nerviosidad, temblor, tensiones musculares, sudoración, atolondramiento, palpitaciones, vértigo y malestar epigástrico. A menudo los pacientes manifiestan temores de una próxima enfermedad o de un accidente, que sufrirán en breve ellos mismos o algunos de sus familiares (Organización Panamericana de la Salud, 2015).

2.1.2 ¿Cómo se produce la ansiedad?

No se sabe a ciencia cierta dónde ni cómo se produce la ansiedad. Sin embargo, se han dado explicaciones, de carácter psicológico, y se han elaborado modelos biológicos para permitir identificar la ansiedad, cuantificarla y entender las bases biológicas de este trastorno (Pérez de la Mora, 2003).

El modelo denominado sistema de inhibición conductual propuesto por Gray & McNaughton, explica muchas de las características de la ansiedad, evaluando las condiciones del medio ambiente en búsqueda de estímulos potencialmente dañinos. La sensibilidad del sistema para responder a las situaciones potencialmente dañinas está determinada por la genética de la persona, y por las numerosas influencias ambientales que han actuado sobre él desde su nacimiento. De esta manera, no todos los individuos son igualmente sensibles a aquellos estímulos que provocan ansiedad. De acuerdo con este modelo, el trastorno de ansiedad generalizada se presentaría con un sistema de inhibición

conductual muy sensible, que toma eventos fisiológicos cotidianos como eventos potencialmente dañinos (Gray & McNaughton, 2000).

En términos biológicos cuando se percibe o se anticipa algún tipo de peligro, el cerebro envía un mensaje al sistema nervioso autónomo. Este sistema es una división del sistema nervioso periférico; el cual recibe información de los órganos sensoriales y controla los movimientos de los músculos esqueléticos. El sistema nervioso autónomo también afecta a los órganos y glándulas para regular el funcionamiento corporal. Debido a que este sistema trabaja sin conciencia, es descrito como un sistema involuntario o automático que se gobierna a sí mismo.

El sistema nervioso autónomo está conformado por dos sistemas anatómicamente separados: el sistema nervioso simpático y el sistema nervioso parasimpático. Los órganos del cuerpo son internados por ambas subdivisiones y cada una tiene un efecto diferente. Por ejemplo, el sistema nervioso simpático acelera el ritmo cardíaco; mientras que el sistema nervioso parasimpático lo reduce. Ambos sistemas funcionan de manera conjunta para mantener a la persona en un estado interno estable (Sotelo Narváez, 2016) (Carlson, 2010). Los cambios sufridos por el sistema nervioso autónomo se pueden cuantificar en manifestaciones mediadas como la frecuencia cardíaca, la presión arterial, la frecuencia respiratoria o la conductancia eléctrica de la piel, que refleja su grado de humedad (sudoración).

A. Sistema nervioso simpático

Este sistema nervioso activa al cuerpo en caso de algún tipo de peligro o estrés (Carlson, 2010). Excita al organismo para la acción defensiva, movilizándolo su energía en situaciones de estrés. Incrementa las activaciones simpáticas cuando se experimenta un estímulo o algo importante está por pasar. Se encarga también de acelerar el ritmo cardíaco, hacer más lento el proceso de digestión, aumenta la concentración sanguínea de glucosa, dilata las arterias y refresca mediante la transpiración, alertando y preparando al organismo para la acción (Empatica, s.f.) (Myers, 2005). Este sistema es el responsable de la respuesta lucha y/o huida; ya que se encarga de liberar la energía preparando el cuerpo para la acción. Cuando el sistema nervioso simpático activa al cuerpo ante alguna respuesta de lucha-huida, todos los órganos que lo conforman tienen a producir una respuesta. Es por esto que,

en la mayoría de los episodios de ataques de pánico, la persona sufre de manera súbita un intenso miedo junto a diversos síntomas (ver Figura 2.1) (Psicosalud,2015).

B. Sistema nervioso parasimpático

Este sistema nervioso produce los efectos contrarios del sistema nervioso parasimpático cuando el estrés disminuye. Conserva la energía mientras disminuye la velocidad de los procesos que han sido acelerados por la activación del sistema nervioso simpático, como disminuir el ritmo cardíaco, reducir la concentración sanguínea de glucosa, entre otros procesos. Es decir, apoya actividades que intervienen en el aumento del abastecimiento corporal de energía almacenada (Empatica, s.f.) (Myers, 2005). Este sistema es el responsable de la restauración y relajación; ya que devuelve al cuerpo a un estado de normalidad previo a la activación (ver Figura 2.1) (Psicosalud, 2015).

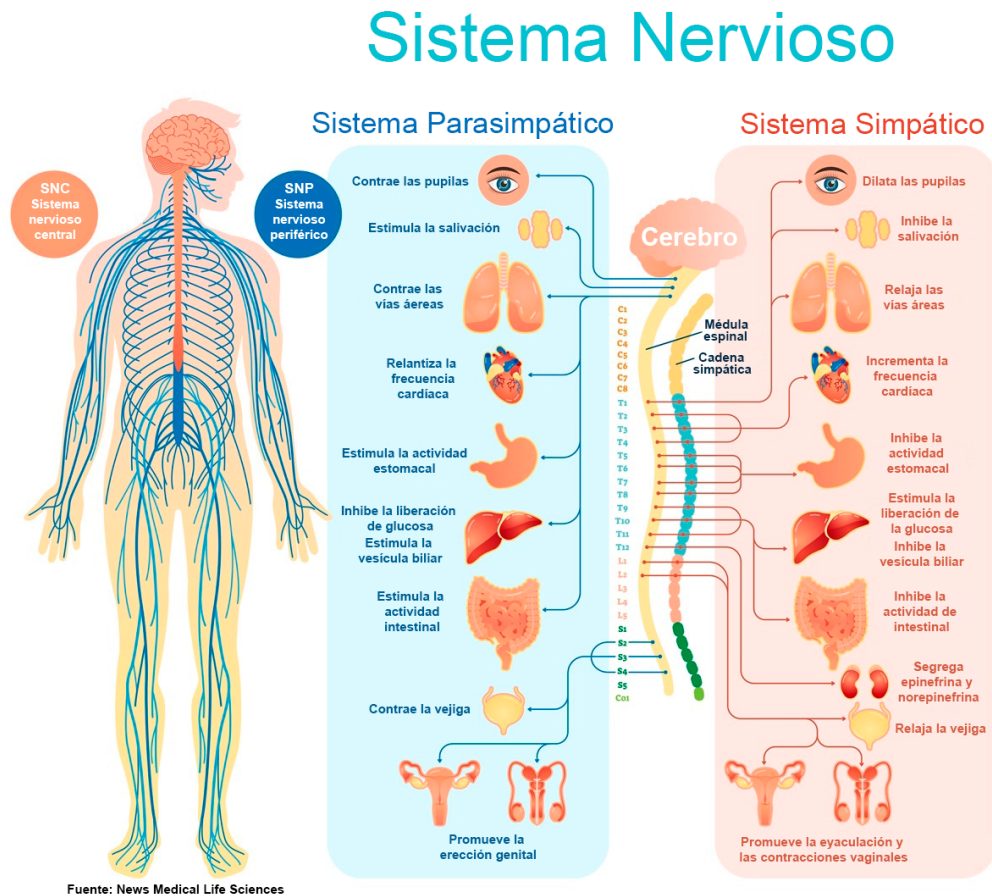


Figura 2.1. Sistema Nervioso: Sistema Nervioso Parasimpático y Sistema nervioso Simpático.

Cuando el sistema nervioso autónomo recibe la señal de un posible peligro, activa al sistema nervioso simpático y éste libera dos productos químicos a través de las glándulas adrenales que están situadas en los riñones. Estos dos productos químicos llamados adrenalina y noradrenalina, viajan libremente por el torrente sanguíneo y son usadas como mensajeros por el sistema nervioso simpático para continuar activando otros sistemas del organismo (Psicosalud, 2015). Una vez que el sistema nervioso simpático se active, continuará incrementando de nivel los síntomas. Sin embargo, este nivel de los síntomas no seguirá creciendo hasta niveles dañinos; porque la actividad del sistema nervioso simpático siempre cesa y lo hace por dos razones distintas:

- la adrenalina y noradrenalina son destruidas por otros productos químicos en el cuerpo. Su vida media en la sangre dura unos minutos hasta que son recaptados;
- cuando se activa el sistema nervioso parasimpático, éste restaura la sensación de relajación.

Cuando el sistema nervioso simpático se activa en exceso, los síntomas no seguirán incrementando; por el contrario, llegará un momento en el que el cuerpo tendrá un límite con la respuesta huida-lucha y activará al sistema nervioso parasimpático para restaurar la sensación de relajación (Psicosalud, 2015).

2.1.3 Manifestaciones sintomatológicas de la ansiedad

Las respuestas fisiológicas relacionadas a la ansiedad son diversas y pueden clasificarse en diferentes grupos (Baezas, Balaguer & Belchi 2008):

Físicos: Se presenta como taquicardia, palpitaciones, opresión en el pecho, falta de aire, temblores, sudoración, molestias digestivas, náuseas, vómito, nudo en el estómago, sensación de obstrucción en la garganta, agarrotamiento, tensión y rigidez muscular, cansancio, hormigueo, sensación de mareo e inestabilidad. También pueden aparecer alteraciones del sueño, alimentación y respuesta sexual (Baezas, Balaguer & Belchi 2008).

Psicológicos: Inquietud, agobio, sensación de amenaza o peligro, ganas de huir o atacar, inseguridad, sensación de vacío, sensación de extrañeza o

despersonalización, temor a perder el control, celos, sospechas, incertidumbre, dificultad para tomar decisiones. Incluso temor a la muerte, a la locura o al suicidio (Baezas, Balaguer & Belchi 2008).

De conducta: Estado de alerta e hipervigilancia, inhibición, bloqueos, torpeza o dificultad para estar quieto y en reposo. Estos síntomas van acompañados de cambios en la expresión corporal y el lenguaje corporal: posturas cerradas, rigidez, movimientos torpes de manos y brazos, cambios en la voz, expresión facial de asombro, duda o crispación, etc (Baezas, Balaguer & Belchi 2008).

Intelectuales o cognitivos: Dificultades de atención, concentración y memoria, aumento de los despistes y descuidos, preocupación excesiva, expectativas negativas, pensamientos distorsionados e importunos, incremento de las dudas y la sensación de confusión, tendencia a recordar sobre todo cosas desagradables, sobrevaloración de pequeños detalles desfavorables, abuso de la prevención y de la sospecha, interpretaciones inadecuadas, susceptibilidad, etc (Baezas, Balaguer & Belchi 2008).

Sociales: Irritabilidad, hostilidad, ensimismamiento, dificultades para iniciar o seguir una conversación, verborrea, bloquearse o quedarse en blanco a la hora de preguntar o responder, dificultades para expresar las propias opiniones o hacer valer los propios derechos, temor excesivo a posibles conflictos, etc (Baezas, Balaguer & Belchi 2008).

2.2 Marcos regulatorios

Los marcos regulatorios del trastorno de ansiedad generalizada son el Manual de diagnóstico y estadístico de trastornos mentales, quinta edición (DSM-V), y la Clasificación estadística internacional de enfermedades y problemas relacionados con la salud (CIE-10). Estos marcos regulatorios son descritos brevemente a continuación.

2.2.1 Manual de diagnóstico y estadístico de trastornos mentales (DSM-V)

El manual de diagnóstico y estadístico de trastornos mentales, quinta edición, es la actualización del manual de diagnóstico y estadísticos de trastornos mentales. Esta quinta edición es la herramienta taxonómica y de diagnóstico publicada por la Asociación Americana de Psiquiatría que sirve como la principal autoridad para los diagnósticos psiquiátricos (American Psychiatric Association, 2013).

2.2.2 Clasificación estadística internacional de enfermedades y problemas relacionados con la salud (CIE-10)

La clasificación estadística internacional de enfermedades y problemas relacionados con la salud es publicada por la Organización Mundial de la Salud como un sistema de clasificación de los trastornos mentales con ejes diagnósticos variables. Se ocupa para fines estadísticos relacionados con la morbilidad y mortalidad. (Organización Panamericana de la Salud, 2015).

2.3 Instrumentos psicológicos de valoración para determinar el trastorno de ansiedad generalizada

En el área de psicología existen diversos instrumentos psicológicos de valoración que permiten detectar el trastorno de ansiedad generalizada. Se recuerda que el trastorno de ansiedad generalizada es la alteración por excelencia de una preocupación patológica; ya que este trastorno se caracteriza por la preocupación ansiosa que la persona refiere como incontrolable y muy frecuente (American Psychiatric Association, 2013).

Los instrumentos de valoración seleccionados para la investigación fueron elegidos para evaluar el grado de preocupación excesiva y la presencia sintomatológica que caracterizan al trastorno de ansiedad generalizada. A continuación, se listan los instrumentos de valoración para determinar el trastorno de ansiedad generalizada utilizados en esta investigación:

A. Escala de Detección del Trastorno de Ansiedad Generalizada

Su nombre en inglés es Screening Scale for DSM-IV y fue elaborado por Carroll y Davidson (Carroll & Davidson, 2000). Evalúa la presencia sintomatológica del trastorno de ansiedad generalizada de acuerdo con los criterios del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-IV) (Padros-Blazquez et al, 2018). Los síntomas deben haberse experimentado la mayoría de los días en los últimos 6 meses, además de valorar la interferencia producida por los mismos (Bados, 2017).

B. Inventario de Preocupación del Estado de Pensilvania

Su nombre en inglés es Penn State Worry Questionnaire y fue elaborado por Meyer, Miller, Metzger y Borkovec (Meyer et al, 1990). Evalúa la tendencia a preocuparse en general, aunque no evalúa las áreas de preocupación. Está redactado de tal forma que sus ítems indican preocupación y falta de problemas de preocupación (Bados, 2017).

2.4 Dispositivos vestibles

Los dispositivos vestibles son aquellos que se llevan puestos y están siempre encendidos. Pueden actuar como una extensión del cuerpo o de la mente del usuario; porque se pueden colocar sobre, debajo o dentro de la ropa. Los dispositivos vestibles que tienen biosensores ópticos y electrónicos integrados proporcionan datos electrofisiológicos o estados bioquímicos en tiempo real. Los datos físicos y bioquímicos pueden ser enviados vía *wireless* a otros dispositivos vestibles para lograr un sistema terapéutico de circuito cerrado logrando un monitoreo en tiempo real de cualquier discapacidad o desorden metabólico (Yetisen et al, 2018).

Los dispositivos vestibles requieren de software y hardware para obtener datos del usuario en tiempo real. Los datos obtenidos de los sensores son procesados, analizados, mostrados, o interpretados para generar información procesable. La ventaja principal de los dispositivos vestibles es el de proporcionar retroalimentación (*feedback*) inmediata de cada paciente. Gracias a los continuos avances tecnológicos en dispositivos vestibles, el potencial de estos dispositivos se ha expandido a la vigilancia médica, al cuidado médico no invasivo, y a

monitoreos móviles de salud y bienestar. Los dispositivos vestibles son regulados de la misma manera que cualquier otro dispositivo médico por la Administración de Alimentos y Medicamentos (FDA).

2.4.1 Pulseras de actividad: Fitbit

La pulsera de actividad Fitbit es un dispositivo para monitorear y dar seguimiento al estado físico como la distancia que se recorrió caminando o corriendo, las calorías consumidas, la calidad de sueño y la frecuencia cardíaca. Este tipo de dispositivos, monitores de actividad del consumidor, son usados actualmente en investigaciones biomédicas para estudiar los efectos terapéuticos de un monitoreo, terapia de ejercicio e intervenciones conductuales. Este tipo de dispositivos tienen como objetivo el asociar la actividad física y las funciones metabólicas, la salud cognitiva y neurológica utilizando el seguimiento de actividades del consumidor.

2.5 Herramientas de software utilizadas en la investigación

A continuación, se describen las herramientas de software utilizadas en esta investigación. Para el caso de IBM SPSS Statistics, el programa fue utilizado para la validación de los instrumentos psicológicos; mientras que el software Anaconda y Jupyter fueron utilizados para la obtención de datos, el análisis exploratorio de los datos, el procesamiento de los datos, la obtención de atributos relevantes y del modelo de comportamiento, así como la evaluación de dicho modelo de comportamiento.

2.5.1 IBM SPSS Statistics

IBM SPSS Statistics es un software estadístico de pago que es utilizado por muchas organizaciones, empresas e investigadores; sobre todo aquellos que son procedentes de las ciencias socioeconómicas o sociales, con el objetivo de resolver una gran variedad de problemas de investigación.

2.5.2 Anaconda

Anaconda es una suite que facilita la tarea de instalar el ambiente e incluye el programa Jupyter. Anaconda es de distribución libre y abierta de los lenguajes Python y R, utilizada tanto en ciencia de datos como en aprendizaje automático.

2.5.3 Jupyter

Jupyter fue un producto desarrollado como parte del proyecto IPython. Este proyecto fue usado para proporcionar un acceso online interactivo a Python. El nombre de Jupyter deriva de la combinación de Julia, Python y R. En sí, Jupyter es una herramienta que permite a los científicos de datos realizar procesos completos de análisis, casi de la misma manera en que otros científicos utilizan una libreta o cuaderno para registrar sus procesos, resultados y conclusiones.

2.6 Librerías utilizadas en la investigación

A continuación, se describen las librerías más importantes utilizadas en la herramienta Jupyter a lo largo de esta investigación.

2.6.1 Pandas

Pandas es la biblioteca de análisis de datos de Python. También es un módulo abierto que permite el trabajar con la importación y la exportación de los formatos de ficheros de datos, así como con estructuras de datos de forma eficaz y fácil de usar.

2.6.2 Scikit-learn

Esta librería es una parte central de Python y es utilizada para aplicar a modelos de minería de datos; tales como los modelos de clasificación y de regresión que incluyen algoritmos como árboles de decisión, bosque aleatorio, vecinos más próximos, regresión múltiple, máquinas de soporte vectorial, entre otros.

2.7 Aprendizaje automático o machine learning

El aprendizaje automático o *machine learning* es una subdisciplina de la Inteligencia Artificial (IA) que se define como la capacidad que tienen las máquinas para aprender y aportar soluciones a problemas concretos; además de generar conocimiento a partir de información que se le proporciona a través del Big Data (Elsevier, 2018). Además, es la búsqueda de algoritmos que razonan desde instancias provistas para generar hipótesis generales, específicamente para supervisar las técnicas de clasificación que son utilizadas para datos entrenados (Pintelas et al, 2018).

2.7.1 Algoritmos de aprendizaje automático o machine learning

Los algoritmos de aprendizaje automático pueden ser clasificados a través de distintas dimensiones, de las fuentes donde se obtienen experiencias de aprendizaje y de la representación de conocimiento utilizada. Estos algoritmos se pueden dividir según la cantidad y el tipo de supervisión requerida durante la capacitación: aprendizaje supervisado (gran cantidad de datos de entrenamiento) y no supervisado (ausencia total de datos de entrenamiento) (Vega Calcines, 2014).

A continuación, se describe brevemente el algoritmo de aprendizaje automático utilizado en esta investigación.

2.7.1.1 Árboles de decisión

Los algoritmos de árboles de decisión permiten predecir los valores de las respuestas a través de reglas de decisión derivadas de las características. Son árboles recursivos que crean modelos basados en la estructura de árboles que utilizan un conjunto de datos de entrenamiento e intentan separar instancias pertenecientes a categorías separadas (Pintelas, et al, 2018).

Capítulo 3

Validación de instrumentos psicológicos para la detección del trastorno de ansiedad generalizada

En este capítulo se presenta la validación de instrumentos psicológicos para la detección del trastorno de ansiedad generalizada.

3.1 Validación de instrumentos para la detección del trastorno de ansiedad generalizada

La Fase 1 perteneciente a la metodología de solución presentada en la Figura 1.2, muestra el proceso de validación de los instrumentos psicológicos utilizados en esta investigación para detectar la ansiedad generalizada a una población de adultos jóvenes perteneciente a Cuernavaca, Morelos, México. La Figura 3.1 muestra detalladamente las tres etapas que integran esta validación.

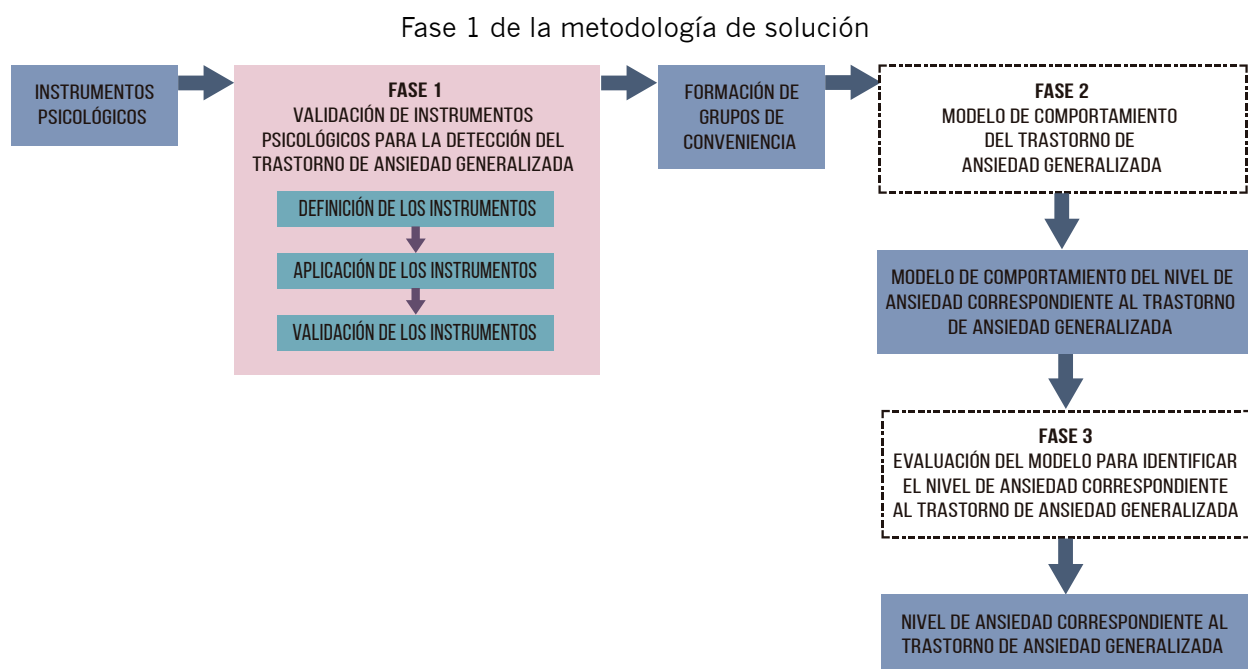


Figura 3.1. Validación de instrumentos psicológicos para detectar el trastorno de ansiedad generalizada.

A continuación, se describen brevemente las etapas que conforman la Fase 1 de la metodología de solución y que se explica a lo largo de este capítulo:

- **Etapa 1. Definición de los instrumentos.** En esta etapa se define el perfil sociodemográfico que debe tener el participante de la investigación. También se definen los instrumentos psicológicos que se utilizan para detectar el trastorno de ansiedad generalizada en la población muestra.

- **Etapa 2. Aplicación de los instrumentos.** En esta etapa se aplican los instrumentos psicológicos definidos en la Etapa 1 y se analizan los resultados obtenidos con técnicas estadísticas.
- **Etapa 3. Validación de los instrumentos.** En esta etapa los instrumentos psicológicos deben ser validados de forma general como una escala de valoración y de forma interna, evaluando cada uno de los ítems que lo componen con técnicas estadísticas.

3.2 Definición de los instrumentos

Esta etapa perteneciente a la Fase 1 de la metodología de solución, y mostrada en la Figura 3.1, contiene la definición de las características sociodemográficas del perfil del grupo de estudio, así como la definición de los instrumentos psicológicos de los que se hace uso en esta investigación para la detección del trastorno de ansiedad generalizada.

3.2.1 Definición del perfil del grupo de estudio

El perfil del grupo de estudio son características específicas que deben tener las personas que participan en el estudio. En esta investigación se definieron los criterios de inclusión y de exclusión para delimitar las variables sociodemográficas que aportarían información a la investigación.

Los participantes seleccionados fueron reclutados del Centro Nacional de Investigación y Desarrollo Tecnológico, ubicado en la ciudad de Cuernavaca, Morelos, México. El perfil del grupo de estudio comprende los siguientes criterios de inclusión: i) edad entre 23 años y 30 años; ii) sin distinción de género; y iii) estudiantes del programa de maestría de Ciencias de la Computación ofrecido por el Centro Nacional de Investigación y Desarrollo Tecnológico.

Los criterios de exclusión fueron los siguientes: i) mujeres embarazadas; ii) participantes mayores a 30 años o menores a 23 años; y iii) participantes en tratamiento por trastornos de ansiedad con medicamento recetado.

Los criterios de inclusión y exclusión fueron definidos de acuerdo a los propósitos de la investigación. Se descartó seleccionar participantes menores de 18 años (niños y adolescentes); porque al ser menores de edad, se requiere hacer intervenciones con permiso de sus padres, y, en caso de realizar pruebas de estudio, los padres o tutores deben estar presentes. A su vez, se descartó seleccionar a adultos mayores por el nivel de complejidad al detectar los trastornos de ansiedad a causa de las comorbilidades; es decir, con frecuencia se pueden asociar síntomas físicos con características propias del envejecimiento, o con efectos secundarios por la toma de medicamentos utilizados y requeridos para esta etapa de la vida (Ramos & Stanley 2018). Por otro lado, se debe recordar que los factores sociales pueden tener un papel predisponente en el desarrollo de los trastornos de ansiedad (Cárdenas, Feria & Palacios, 2010). Por lo tanto, al eliminar a participantes menores de edad y adultos mayores, se decidió trabajar con adultos jóvenes. No obstante, para realizar las pruebas de estudio se necesitaba tener cierto grado de control; es decir, se buscaban participantes que estuvieran rodeados por casi los mismos factores biológicos, psicológicos y sociales. De esta manera, se eligieron participantes que estuvieran estudiando un posgrado; por el nivel de estrés por el que pudieran verse afectados; siendo un posgrado más demandante que una carrera universitaria. En este sentido, se decidió excluir a mujeres embarazadas y participantes en tratamiento por trastornos de ansiedad con medicamentos recetado; ya que este tipo de participantes formarían parte de otras investigaciones al estar rodeados por diferentes factores biológicos, psicológicos, y sociales que pudieran entorpecer la investigación.

3.2.2 Instrumentos psicológicos seleccionados para la investigación

Para la definición de los instrumentos psicológicos se analizaron las siguientes escalas de valoración del trastorno de ansiedad generalizada para seleccionar aquellos que se adecuaban a los propósitos de la investigación:

1. entrevista estructurada para los trastornos de ansiedad según el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V);
2. entrevista diagnóstica internacional compuesta;
3. cuestionario de preocupación y ansiedad;
4. cuestionario de preocupación del estado de Pensilvania;

5. cuestionario del trastorno de ansiedad generalizada-IV;
6. escala del trastorno de ansiedad generalizada.

En esta investigación se utilizaron los siguientes cuestionarios para detectar el trastorno de ansiedad generalizada: el cuestionario de preocupación del estado de Pensilvania (ver Anexo I), y el cuestionario de detección del trastorno de ansiedad generalizada (ver Anexo II). El cuestionario de preocupación del estado de Pensilvania se utilizó para evaluar el rasgo de preocupación, principal elemento para la detección de ansiedad generalizada; mientras que el cuestionario de detección del trastorno de ansiedad generalizada se utilizó para la correlación de respuestas de déficit cognitivo conductual con el monitoreo de variables a través de sensores vestibles.

Se debe tener en cuenta que son las escalas de valoración definidas las que determinan los rangos de ansiedad para las pruebas de la investigación.

Los instrumentos seleccionados para esta investigación consisten en lo siguiente:

A. Cuestionario de preocupación del estado de Pensilvania

El cuestionario de preocupación del estado de Pensilvania elaborado por Meyer Miller, Metzger y Borkovec (Meyer et al, 1990), es uno de los instrumentos más utilizados para evaluar la preocupación como fenómeno excesivo, generalizado e incontrolable. Además, captura el aspecto fundamental de una preocupación patológica, principal característica del trastorno de ansiedad generalizada. Este instrumento fue diseñado específicamente para evaluar la preocupación como proceso cognitivo; pero no evalúa los síntomas somáticos (González et al, 2018). No obstante, es de gran relevancia para la detección, estudio y tratamiento del trastorno de ansiedad generalizada.

Este instrumento de valoración consta de 16 ítems con una escala tipo Likert de 5 puntos, que oscila entre 1 y 5. Las opciones de respuesta van de 1= Nada, 2= Algo, 3= Regular, 4= Bastante y 5= Mucho (Meyer et al, 1990).

La suma de los ítems que componen este instrumento de valoración va de los 16 a los 80 puntos. Se considera a una persona sin rasgos ansiosos

cuando las puntuaciones van por debajo de los 40. Una puntuación mayor a los 40 se considera una persona con rasgos ansiosos. No obstante, en una puntuación mayor a los 60, se considera a la persona con rasgos de ansiedad generalizada (Meyer et al, 1990). De esta manera, las puntuaciones pueden ir de la siguiente forma (ver Tabla 3.1):

Tabla 3.1. Rangos de ansiedad generalizada según el cuestionario del estado de Pensilvania.

Rangos de ansiedad generalizada según el cuestionario del estado de Pensilvania	
Bajo	Puntuación entre 16 y 39
Medio	Puntuación entre 40 y 59
Alto	Puntuación entre 60 y 80

En esta investigación la versión del cuestionario de preocupación del estado Pensilvania que se utilizó es el traducido a su versión en español por Nuevo, Montorio y Ruiz (Nuevo Benítez, Montorio Cerrato, & Ruiz Díaz, 2002) (ver Anexo I).

B. Cuestionario de detección del trastorno de ansiedad generalizada

El cuestionario de detección del trastorno de ansiedad generalizada de Carroll y Davidson (Carroll & Davidson, 2000), es un instrumento simple y rápido que fue creado para identificar personas con trastorno de ansiedad generalizada de acuerdo con los criterios del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-IV). Esta escala de valoración evalúa principalmente la presencia sintomatológica del trastorno de ansiedad generalizada. Los síntomas deben haberse experimentado la mayoría de los días en los últimos 6 meses; además de valorar la interferencia producida por los mismos.

Este instrumento de valoración consta de 12 ítems dicotómicos (Sí/No). El cuestionario se divide en 4 criterios: el criterio A se refiere a la expectativa aprensiva (ítems 1 y 2), el criterio B a la falta de control en la preocupación (ítems 3 y 4), el criterio C a los síntomas fisiológicos (ítems del 5 al 10); y finalmente, el criterio D se refiere al aspecto temporal y su afectación a la vida cotidiana (ítems 11 y 12) (Carroll & Davidson, 2000).

La suma de los ítems que componen este instrumento de valoración va de los 0 a los 12 puntos. Se considera una persona con síntomas no sugestivos al trastorno de ansiedad generalizada cuando la puntuación va de los 0 a los 5 puntos. Cuando la puntuación va de los 6 a los 12, la persona es considerada con síntomas sugestivos al trastorno de ansiedad generalizada (Carroll & Davidson, 2000). De esta manera, las puntuaciones pueden ir de la siguiente forma (ver Tabla 3.2):

Tabla 3.2. Rangos de ansiedad generalizada según la escala de detección del trastorno de ansiedad generalizada.

Rangos de ansiedad generalizada según la escala de detección del trastorno de ansiedad generalizada	
Bajo	Puntuación entre 0 y 5
Alto	Puntuación entre 6 y 12

En esta investigación la versión del cuestionario de detección del trastorno de ansiedad generalizada que se utilizó es el traducido a su versión en español por Bobes, García-Calvo, Prieto, García-García y Rico-Villademoros (Bobes, García-Calvo & Prieto, 2006) (ver Anexo II).

3.3 Aplicación de los instrumentos

El cuestionario de preocupación del estado de Pensilvania y la escala de detección del trastorno de ansiedad generalizada se aplicaron por primera vez a una población de 31 personas. Esta aplicación tuvo como objetivo el validar el instrumento de forma externa e interna (evaluación de las propiedades métricas); para que los resultados obtenidos se generalizaran al resto de la población cuando el cuestionario se aplicara nuevamente a una muestra totalmente diferente.

Se habla de una validación de forma externa cuando el instrumento psicológico se valida como un instrumento de valoración. Cuando se habla de una validación interna es cuando se valida cada uno de los ítems que lo componen. La validación de los instrumentos psicológicos requiere de la comprobación de su utilidad antes de su aplicación. Además de ser importante; porque asegura la

confiabilidad de la investigación, así como de los resultados obtenidos. Esta validación toma mayor relevancia si la investigación está relacionada a las ciencias de la salud.

En el caso de esta investigación, los instrumentos psicológicos definidos se validaron exclusivamente para la población de alumnos pertenecientes al posgrado de Ciencias de la Computación del Centro Nacional de Investigación y Desarrollo Tecnológico, ubicado en el estado de Morelos.

3.3.1 Equilibrio de datos con técnicas estadísticas

Los datos obtenidos de esta primera aplicación se revisaron para evaluar tanto el rango y la varianza de cada ítem, como también la fiabilidad de la puntuación de toda la escala; y a su vez, las correlaciones existentes entre ítems. Durante este paso, también se desarrolló un análisis para obtener la distribución de las respuestas por ítem con estadística descriptiva como medias y desviación estándar. En la Figura 3.2 se puede observar los datos obtenidos de la estadística descriptiva perteneciente al cuestionario de preocupación del estado de Pensilvania que consta de 16 ítems; mientras que la Figura 3.3 muestra los datos obtenidos de la estadística descriptiva perteneciente a la escala de detección del trastorno de ansiedad generalizada que consta de 12 ítems.

La Figura 3.2 muestra los resultados obtenidos de la estadística básica perteneciente al cuestionario de preocupación del estado de Pensilvania. Cuestionario que está compuesto por 16 ítems en escala tipo Likert de 5 puntos, que oscila entre 1 y 5. Las opciones de respuesta van de 1= Nada, 2= Algo, 3= Regular, 4= Bastante y 5= Mucho (Meyer et al, 1990). Las columnas 1-16 pertenecen a los 16 ítems del cuestionario.

Datos estadísticos pertenecientes al cuestionario del estado de Pensilvania

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
conteo	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000
media	2.935484	2.612903	2.612903	2.193548	2.870968	2.870968	2.032258	2.387097	2.129032	1.903226	1.806452	1.548387	2.096774	2.064516	1.806452	3.096774
σ (s)	1.314927	1.054433	1.115868	1.108132	1.586231	1.203936	1.048296	1.229564	1.231312	1.106190	1.137815	0.809885	1.220788	1.123550	1.013882	1.075634
min	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
25%	2.000000	2.000000	2.000000	1.500000	1.500000	2.000000	1.000000	2.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	2.500000
50%	3.000000	3.000000	2.000000	2.000000	2.000000	3.000000	2.000000	2.000000	2.000000	2.000000	1.000000	1.000000	2.000000	2.000000	2.000000	3.000000
75%	4.000000	3.000000	3.500000	2.500000	4.500000	4.000000	2.500000	3.000000	3.000000	2.000000	2.000000	2.000000	2.500000	2.000000	2.000000	4.000000
max	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	4.000000	5.000000	5.000000	5.000000	5.000000

Figura 3.2. Datos estadísticos obtenidos del cuestionario de preocupación del estado de Pensilvania pertenecientes a los ítems 1-16. Los datos estadísticos obtenidos son: el número de encuestados (conteo = 31 personas), la media obtenida de cada ítem, la desviación estándar obtenida de cada ítem, el valor mínimo obtenido de cada ítem; los valores percentiles correspondientes al 25%, 50 y 75% de cada ítem; y el valor máximo obtenido de cada ítem.

La Figura 3.3, por otro lado, muestra los resultados obtenidos de la estadística básica perteneciente a la escala de valoración del trastorno de ansiedad generalizada. Escala que está compuesta por 12 ítems dicotómicos (sí/no). Sin embargo, para esta escala de valoración se discretizaron los valores obtenidos de los 12 ítems, dándole un valor de 0 a las respuestas correspondientes a sí, y un valor de 1 a las respuestas correspondientes a no. Las columnas 1-12 pertenecen a los 12 ítems del cuestionario (Carroll & Davidson, 2000).

Datos estadísticos pertenecientes a la escala de detección del trastorno de ansiedad generalizada

	1	2	3	4	5	6	7	8	9	10	11	12
conteo	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000
media	0.741935	0.709677	0.806452	0.870968	0.806452	0.322581	0.451613	0.677419	0.806452	0.645161	0.677419	0.838710
σ (s)	0.444803	0.461414	0.401610	0.340777	0.401610	0.475191	0.505879	0.475191	0.401610	0.486373	0.475191	0.373878
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.500000	0.000000	1.000000	1.000000	1.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	1.000000
50%	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000
75%	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Figura 3.3. Datos estadísticos obtenidos de la escala de detección del trastorno de ansiedad generalizada pertenecientes a los ítems 1-12. Los datos obtenidos son: el número de encuestados (conteo = 31 personas), la media obtenida de cada ítem, la desviación estándar obtenida de cada ítem; los valores percentiles correspondientes al 25%, 50 y 75% de cada ítem; y el valor máximo obtenido de cada ítem.

3.4 Validación de los instrumentos

Una vez se obtuvieron los resultados del análisis estadístico (Figura 3.2 y 3.3) con el programa de Jupyter Notebook (comando *describe*), se procedió a validar las propiedades métricas de los instrumentos psicológicos para la valoración del trastorno de ansiedad generalizada. Este paso tuvo como objetivo validar el cuestionario en una población de un lugar específico, que, para el caso de esta investigación, fueron validados para la ciudad de Cuernavaca, Morelos, México.

3.4.1 Análisis factorial

Para determinar la estructura interna (validación de cada uno de los ítems) de los instrumentos psicológicos del trastorno de ansiedad generalizada utilizadas en esta investigación, se consideró el uso de técnicas estadísticas avanzadas como el análisis factorial. El análisis factorial se incorpora a la metodología cuantitativa relacionando variables latentes con variables observadas (Salmerón Enciso, 2016). Las variables latentes son aquellas que no pueden medirse de una manera directa; es decir, se estiman a través de variables observadas; tales como el nivel de ansiedad, el nivel socioeconómico o el grado de satisfacción. En cambio, las variables observadas son las que pueden dar respuesta a un reactivo de un examen, el número de aciertos en un examen o el número de reactivos de una encuesta.

Los resultados obtenidos en el análisis factorial corresponden al cuestionario de preocupación del estado de Pensilvania y a la escala de detección del trastorno de ansiedad generalizada, en donde se utilizó el índice Kaiser-Meyer-Olkin, la prueba de esfericidad de Barlett y el método de rotación Oblimin por la falta de normalización en los datos. Este análisis factorial se realizó con el programa de IBM SPSSStatistics con el objetivo de determinar un número reducido de factores que puedan representar a las variables originales.

La ventaja de utilizar este programa de estadística perteneciente a IBM, es que permite importar los resultados de las encuestas en formato CSV y hacer todo el procedimiento de manera amigable para el usuario, eligiendo los índices como el de Kaiser-Meyer-Olkin, la prueba de esfericidad de Barlett y el coeficiente alfa de Cronbach de una lista predeterminada desde el mismo software, eliminando la

necesidad de introducir fórmulas o generar código en programas como Jupyter Notebook o R. Los resultados obtenidos en este programa de IBM se muestran como figuras y tablas.

A. Análisis factorial: cuestionario de preocupación del estado de Pensilvania

El análisis factorial realizado para el cuestionario de preocupación del estado de Pensilvania obtuvo un coeficiente Kaiser-Meyer-Olkin de 0.875 y un coeficiente de la prueba de esfericidad de Barlett de 0.000 (ver Figura 3.4).

Medida de adecuación muestra de Kaiser-Meyer-Olkin		.875
Prueba de esfericidad De Barlett	Chi-cuadrada	618.470
	df.	105
	Sig.	.000

Figura 3.4. Datos obtenidos de la prueba de Kaiser-Meyer-Olkin y Barlett perteneciente al cuestionario de preocupación del estado de Pensilvania. Se obtiene 0.875 para la medida de adecuación muestra de Kaiser-Meyer-Olkin; mientras que en la prueba de esfericidad de Barlett se obtiene 618.470 de Chi-cuadrada, 105 grados de libertad y un p-valor de 0.000.

El estadístico Kaiser-Meyer-Olkin contrasta si las correlaciones parciales entre las variables son pequeñas, variando entre 0 y 1. Los valores menores que 0.5 obtenidos en el estadístico Kaiser-Meyer-Olkin indican que el análisis factorial no debe utilizarse con los datos muestrales que se están analizando. Por otro lado, la prueba de esfericidad de Barlett contrasta si la matriz de correlaciones es una matriz de identidad, indicando si el modelo factorial es inadecuado.

Al haber obtenido un coeficiente Kaiser-Meyer-Olkin de 0.875 (ver Figura 3.4), se demuestra que, para el cuestionario de preocupación del estado de Pensilvania, es factible realizar un análisis factorial. Parte de este análisis factorial, es la obtención de la covarianza y la correlación de la matriz anti-imagen. En este aspecto, la correlación de la matriz anti-imagen toma una gran importancia; ya que al observar los resultados obtenidos y destacados con un superíndice a , se pueden eliminar aquellos ítems que pertenecen a los instrumentos psicológicos con valores que están por abajo de 0.5. De

esta manera, si se eliminan los ítems con valores por debajo de 0.5, se logra obtener una mejora en el coeficiente Kaiser-Meyer-Olkin. El objetivo es tener valores que estén cercanos a 1, con un coeficiente Kaiser-Meyer-Olkin por arriba del 0.5. Al realizar el análisis factorial para el cuestionario de preocupación del estado de Pensilvania, se eliminó el ítem 1 por obtener un valor menor a 0.5. De esta forma, se logró obtener un coeficiente Kaiser-Meyer-Olkin de 0.875.

La Figura 3.5 muestra los 15 ítems restantes de este instrumento psicológico, donde los valores obtenidos están destacados con un superíndice a, cuyos valores son mayores a 0.5 y cercanos a 1. También muestra la matriz de covarianzas anti-imagen y la matriz de correlaciones anti-imagen. La matriz de covarianzas anti-imagen contiene los negativos de las covarianzas parciales; mientras que la matriz de correlaciones anti-imagen contiene los coeficientes de correlación parcial cambiados de signo. En la diagonal de la matriz de correlaciones anti-imagen se encuentran las medidas de adecuación para cada variable. Si el modelo factorial elegido es el adecuado para explicar los datos, los elementos de la diagonal de la matriz (identificados con un superíndice a) deben tener un valor cercano a 1; mientras que los demás elementos deben ser pequeños.

Matriz anti-imagen perteneciente al cuestionario de preocupación del estado de Pensilvania

	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	
Covarianza de matriz anti-imagen	V2	.236	-.024	-.048	.007	-.079	-.041	-.055	-.001	-.025	-.016	-.003	.029	.026	.026	-.030
	V3	-.024	.429	.002	-.126	-.009	-.018	-.044	-.018	-.044	-.001	-.003	.026	.029	-.015	-.003
	V4	-.048	.002	.120	-.039	-.049	-.054	.007	-.026	-.010	-.025	-.021	.017	.004	.007	.051
	V5	.007	-.126	-.039	.365	-.082	.001	-.004	.010	.077	-.018	.021	-.036	-.015	.031	-.035
	V6	-.079	-.009	-.049	-.082	.384	.047	.002	.017	-.016	.035	.021	-.007	-.013	-.039	-.022
	V7	-.041	-.018	-.054	.001	.047	.093	.029	.004	.008	.023	.029	-.033	-.025	-.015	-.010
	V8	-.055	-.044	.007	-.004	.002	.029	.077	-.032	.008	.015	.001	-.032	-.022	.003	.033
	V9	-.001	-.018	-.026	.010	.017	.004	-.032	.074	-.011	-.016	.010	-.003	.005	.006	-.043
	V10	-.025	-.044	-.010	.077	-.016	.008	.008	-.011	.096	.002	.011	-.012	-.018	-.008	-.009
	V11	-.016	-.001	-.025	-.018	.035	.023	.015	-.016	.002	.053	.009	-.010	-.016	-.023	.007
	V12	-.003	-.003	-.021	.021	.021	.029	.001	.010	.011	.009	.048	-.018	-.022	-.007	-.023
	V13	.029	.026	.017	-.036	-.007	-.033	-.032	-.003	-.012	-.010	-.018	.046	.016	-.009	-.004
	V14	.026	.029	.004	-.015	-.013	-.025	-.022	.005	-.018	-.016	-.022	.016	.027	.000	-.002
	V15	.026	-.015	.007	.031	-.039	-.015	.003	.006	-.008	-.023	-.007	-.009	.000	.043	-.011
	V16	-.030	-.003	.051	-.035	-.022	-.010	.033	-.043	-.009	.007	-.023	-.004	-.002	-.011	.080
	Correlación de matriz anti-imagen	V2	.719 ^a	-.074	-.287	.024	-.263	-.274	-.409	-.011	-.165	-.141	-.024	.277	.328	.255
V3		-.074	.882 ^a	.010	-.317	-.023	-.091	-.242	-.101	-.219	-.004	-.020	.188	.266	-.112	-.015
V4		-.287	.010	.838 ^a	-.184	-.227	-.512	.073	-.281	-.090	-.319	-.270	.233	.079	.093	.515
V5		.024	-.317	-.184	.853 ^a	-.219	.005	-.023	.063	.409	-.128	.160	-.275	-.152	.246	-.203
V6		-.263	-.023	-.227	-.219	.910 ^a	.249	.014	.102	-.085	.246	.158	-.053	-.132	-.302	-.127
V7		-.274	-.091	-.512	.005	.249	.841 ^a	.338	.042	.080	.329	.437	-.507	-.490	-.234	-.120
V8		-.409	-.242	.073	-.023	.014	.338	.852 ^a	-.424	.090	.228	.020	-.545	-.491	.052	.418
V9		-.011	-.101	-.281	.063	.102	.042	-.424	.917 ^a	-.129	-.261	.170	-.055	.105	.098	-.560
V10		-.165	-.219	-.090	.409	-.085	.080	.090	-.129	.940 ^a	.028	.169	-.176	-.357	-.125	-.097
V11		-.141	-.004	-.319	-.128	.246	.329	.228	-.261	.028	.902 ^a	.173	-.205	-.412	-.474	.105
V12		-.024	-.020	-.270	.160	.158	.437	.020	.170	.169	.173	.873 ^a	-.383	-.599	-.154	-.365
V13		.277	.188	.233	-.275	-.053	-.507	-.545	-.055	-.176	-.205	-.383	.865 ^a	.453	-.207	-.058
V14		.328	.266	.079	-.152	-.132	-.490	-.491	.105	-.357	-.412	-.599	.453	.839 ^a	-.006	-.042
V15		.255	-.112	.093	.246	-.302	-.234	.052	.098	-.125	-.474	-.154	-.207	-.006	.924 ^a	-.187
V16		-.220	-.015	.515	-.203	-.127	-.120	.418	-.560	-.097	.105	-.365	-.058	-.042	-.187	.865 ^a

Figura 3.5. Datos obtenidos de la matriz anti-imagen perteneciente al cuestionario de preocupación del estado de Pensilvania. Los valores de las columnas (v2-v16) representan los 15 ítems seleccionados para el cuestionario de preocupación del estado de Pensilvania (ítems 2 al16).

Por otra parte, se definieron 3 factores para dividir los 15 ítems seleccionados del cuestionario de preocupación del estado de Pensilvania (ver Figura 3.6) para obtener una varianza acumulativa del 86.283% (ver Figura 3.7).

La Figura 3.6 (a) sirve para determinar el número óptimo de factores. Estos factores se definieron gracias a la gráfica de sedimentación, los cuales están resaltados por líneas punteadas en la Figura 3.6. La Figura 3.6 (b); por otro lado, muestra la matriz de patrones, en donde los componentes son divididos en los factores previamente definidos. Los ítems que pertenecerán a cada factor son elegidos por sus valores cercanos a 1. De esta forma, en el factor 1 se encuentran los ítems 7, 8, 9, 10, 11, 12, 13, 14, 15 y 16. En el factor 2 se encuentran los ítems 2 y 4; mientras que en el factor 3 se encuentran los ítems 3 y 5.

Definición de factores y matriz de patrones pertenecientes al cuestionario de preocupación del estado de Pensilvania

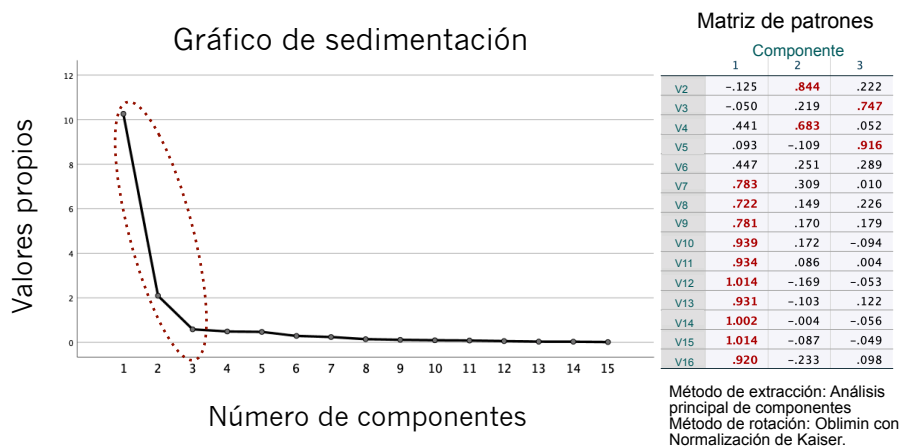


Figura (a) Gráfico de sedimentación

Figura (b) Matriz de patrones

La Figura 3.6(a) muestra el número de componentes con sus valores propios; mientras que la Figura 3.6(b) muestra los ítems divididos en los 3 factores (los ítems para cada factor son elegidos por su valor cercano a 1). En el factor 1 se encuentran los ítems 7, 8, 9, 10, 11, 12, 13, 14, 15 y 16. En el factor 2 se encuentran los ítems 2 y 4; mientras que en el factor 3, se encuentran los ítems 3 y 5.

Figura 3.6. Datos obtenidos en cuanto a la selección de 3 factores pertenecientes al cuestionario de preocupación del estado de Pensilvania.

La Figura 3.7 muestra los autovalores asociados a cada factor, y su respectivo porcentaje de varianza explicada. Los autovalores mostrados en la Figura 3.7 están ordenados de mayor a menor, siendo el primer autovalor el mayor de los posibles, el segundo autovalor es el segundo mayor, y así sucesivamente. En la sección de autovalores propios iniciales se encuentran los ítems 2-16 del cuestionario de preocupación del estado de Pensilvania con su respectivo porcentaje de varianza y varianza acumulada. Por otro lado, la sección suma de extracción de cargas cuadradas muestra los tres factores anteriormente definidos con su respectivo porcentaje de varianza y varianza acumulada. En esta sección, si un autovalor se aproxima a cero, significa que el factor correspondiente a este autovalor es incapaz de explicar una cantidad relevante de la varianza total.

Porcentajes de varianza perteneciente al cuestionario de preocupación del estado de Pensilvania

Componente	Autovalores propios iniciales			Suma de extracción de cargas cuadradas			Sumas de rotación de cargas cuadradas Total
	Total	% de varianza	% acumulada	Total	% de varianza	Cumulativa %	
1	10.261	68.409	68.409	10.261	68.409	68.409	9.840
2	2.095	13.964	82.373	2.095	13.964	82.373	3.230
3	.586	3.910	86.283	.586	3.910	86.283	4.834
4	.488	3.256	89.539				
5	.470	3.135	92.674				
6	.289	1.925	94.598				
7	.240	1.600	96.199				
8	.143	.954	97.153				
9	.114	.762	97.915				
10	.094	.630	98.545				
11	.083	.553	99.098				
12	.058	.388	99.486				
13	.033	.220	99.706				
14	.030	.197	99.903				
15	.015	.097	100.000				

Figura 3.7. Datos obtenidos de los porcentajes de varianza perteneciente a los 15 ítems seleccionados del cuestionario de preocupación del estado de Pensilvania.

B. Análisis factorial: escala de detección del trastorno de ansiedad generalizada

Para el caso de la escala de detección del trastorno de ansiedad generalizada se obtuvo un coeficiente Kaiser-Meyer-Olkin de 0.744 y un coeficiente de la prueba de esfericidad de Barlett de 0.000 (ver Figura 3.8).

Prueba de Kaiser-Meyer-Olkin y esfericidad de Barlett

Medida de adecuación muestra de Kaiser-Meyer-Olkin		.744
Prueba de esfericidad De Barlett	Chi-cuadrada	63.641
	df.	15
	Sig.	.000

Figura 3.8. Datos obtenidos de la prueba de Kaiser-Meyer-Olkin y Barlett perteneciente a la escala de detección del trastorno de ansiedad generalizada. Se obtiene 0.744 para la medida de adecuación muestra de Kaiser-Meyer-Olkin; mientras que en la prueba de esfericidad de Barlett se obtiene 63.641 de Chi-cuadrada, 15 grados de libertad y un p-valor de 0.000.

El estadístico Kaiser-Meyer-Olkin contrasta si las correlaciones parciales entre las variables son pequeñas variando entre 0 y 1. Los valores menores que 0.5 obtenidos en el estadístico Kaiser-Meyer-Olkin indican que no debe utilizarse el análisis factorial con los datos muestrales que se están analizando. Por otro lado, la prueba de esfericidad de Barlett contrasta si la

matriz de correlaciones es una matriz de identidad, indicando si el modelo factorial es inadecuado.

Al haber obtenido un coeficiente Kaiser-Meyer-Olkin de 0.744 (ver Figura 3.8), se demuestra que, para la escala de detección del trastorno de ansiedad generalizada, es factible realizar un análisis factorial. Parte de este análisis factorial, es la obtención de la covarianza y la correlación de la matriz anti-imagen. En este aspecto, la correlación de la matriz anti-imagen toma una gran importancia; ya que al observar los resultados obtenidos y destacados con un superíndice α , se pueden eliminar aquellos ítems que pertenecen a los instrumentos psicológicos con valores que están por abajo de 0.5. De esta manera, si se eliminan los ítems con valores por debajo de 0.5, se logra obtener una mejora en el coeficiente Kaiser-Meyer-Olkin. El objetivo es tener valores que estén cercanos a 1, con un coeficiente Kaiser-Meyer-Olkin por arriba del 0.5.

Al realizar el análisis factorial para la escala de detección del trastorno de ansiedad generalizada, se eliminaron los ítems 6, 7, 8, 9, 11 y 12 por obtener un valor menor a 0.5. De esta forma, se logró obtener un coeficiente Kaiser-Meyer-Olkin de 0.744.

La Figura 3.9 muestra los 6 ítems restantes de este instrumento psicológico, donde los valores obtenidos están destacados con un superíndice α , cuyos valores son mayores a 0.5 y cercanos a 1. También muestra la matriz de covarianzas anti-imagen y la matriz de correlaciones anti-imagen. La matriz de covarianzas anti-imagen contiene los negativos de las covarianzas parciales; mientras que la matriz de correlaciones anti-imagen contiene los coeficientes de correlación parcial cambiados de signo. En la diagonal de la matriz de correlaciones anti-imagen se encuentran las medidas de adecuación para cada variable. Si el modelo factorial elegido es el adecuado para explicar los datos, los elementos de la diagonal de la matriz (identificados con un superíndice α) deben tener un valor cercano a 1; mientras que los demás elementos deben ser pequeños.

Matriz anti-imagen perteneciente a la escala de detección del trastorno de ansiedad generalizada

		r1	r2	r3	r4	r5	r10
Covarianza de matriz anti-imagen	r1	.404	-.187	.070	-.058	-.183	-.129
	r2	-.187	.592	-.031	-.107	.034	-.138
	r3	.070	-.031	.531	-.181	-.216	.121
	r4	-.058	-.107	-.181	.582	-.078	.059
	r5	-.183	.034	-.216	-.078	.383	-.079
	r10	-.129	-.138	.121	.059	-.079	.688
Correlación de matriz anti-imagen	r1	.734 ^a	-.382	.150	-.119	-.466	-.246
	r2	-.382	.794 ^a	-.055	-.181	.071	-.216
	r3	.150	-.055	.655 ^a	-.326	-.479	.201
	r4	-.119	-.181	-.326	.832 ^a	-.166	.093
	r5	-.466	.071	-.479	-.166	.724 ^a	-.155
	r10	-.246	-.216	.201	.093	-.155	.746 ^a

Figura 3.9. Datos obtenidos de la matriz anti-imagen perteneciente a la escala de detección del trastorno de ansiedad generalizada. Los valores de las columnas (r1, r2, r3, r4, r5 y r10) representan los 6 ítems seleccionados para la escala de detección del trastorno de ansiedad generalizada (ítems 1, 2, 3, 4, 5 y 10).

Por otra parte, se definieron 2 factores para dividir los 6 ítems seleccionados de la escala de detección del trastorno de ansiedad generalizada (ver Figura 3.11) para obtener una varianza acumulativa del 72.264%.

La Figura 3.10 (a) sirve para determinar el número óptimo de factores. Estos factores se definieron gracias a la gráfica de sedimentación; los cuales están resaltados por líneas punteadas en la Figura 3.10. La Figura 3.10 (b), por otro lado, muestra la matriz de patrones, en donde los componentes son divididos en los factores previamente definidos. Los ítems que pertenecerán a cada factor son elegidos por sus valores cercanos a 1. De esta forma, en el factor 1 se encuentran los ítems 3, 4 y 5; mientras que en el factor 2 se encuentran los ítems 1, 2 y 10.

Definición de factores y matriz de patrones pertenecientes a la escala de detección del trastorno de ansiedad generalizada

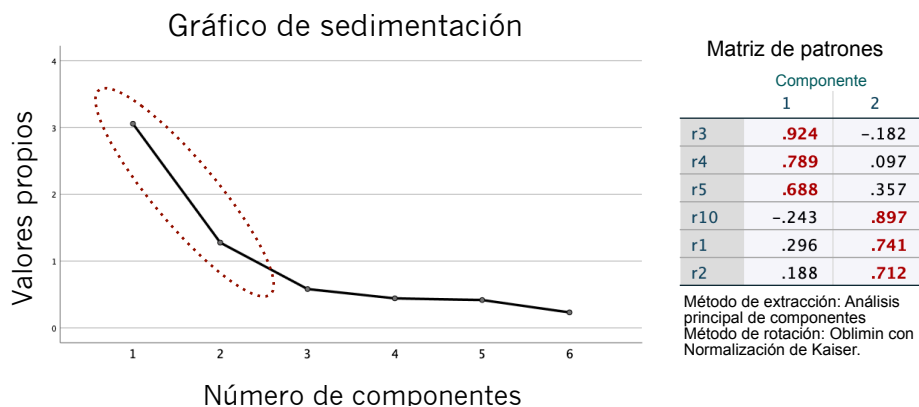


Figura (a) Gráfico de sedimentación

Figura (b) Matriz de patrones

La Figura 3.10(a) muestra el número de componentes con sus valores propios; mientras que la Figura 3.10(b) muestra los ítems divididos en los 2 factores (los ítems para cada factor son elegidos por su valor cercano a 1). En el factor 1 se encuentran los ítems 3, 4, y 5. En el factor 2 se encuentran los ítems 1, 2 y 10.

Figura 3.10. Datos obtenidos en cuanto a la selección de 2 factores pertenecientes a la escala de detección del trastorno de ansiedad generalizada.

La Figura 3.11 muestra los autovalores asociados a cada factor, y su respectivo porcentaje de varianza explicada. Los autovalores mostrados en la Figura 3.11 están ordenados de mayor a menor, siendo el primer autovalor el mayor de los posibles, el segundo autovalor es el segundo mayor, y así sucesivamente. En la sección de autovalores propios iniciales se encuentran los ítems 1, 2, 3, 4, 5 y 10 de la escala de detección del trastorno de ansiedad generalizada con su respectivo porcentaje de varianza y varianza acumulada. Por otro lado, la sección suma de extracción de cargas cuadradas muestra los dos factores anteriormente definidos con su respectivo porcentaje de varianza y varianza acumulada. En esta sección, si un autovalor se aproxima a cero, significa que el factor correspondiente a este autovalor es incapaz de explicar una cantidad relevante de la varianza total.

Porcentajes de varianza perteneciente a la escala de detección del trastorno de ansiedad generalizada

Componente	Autovalores propios iniciales			Suma de extracción de cargas cuadradas			Sumas de rotación de cargas cuadradas
	Total	% de varianza	% acumulada	Total	% de varianza	% acumulada	Total
1	3.055	50.915	50.915	3.055	50.915	50.915	2.467
2	1.276	21.264	72.179	1.276	21.264	72.179	2.375
3	.581	9.678	81.857				
4	.441	7.355	89.212				
5	.416	6.935	96.147				
6	.231	3.853	100.000				

Figura 3.11. Datos obtenidos de los porcentajes de varianza perteneciente a los 6 ítems seleccionados de la escala de detección del trastorno de ansiedad generalizada.

3.4.2 Análisis de fiabilidad

Una vez se realizó el análisis factorial, se procedió a realizar un análisis de fiabilidad. Este análisis estuvo enfocado a evaluar la fiabilidad de ambos instrumentos psicológicos utilizando el coeficiente alfa de Cronbach; el cual es una medida de la consistencia interna de los puntajes de los ítems; es decir, es la medida en la que los resultados de cada ítem están correlacionados entre sí. Principalmente, es una función de las correlaciones entre los elementos y el número total de elementos. Al final, la encuesta será considerada como válida cuando los puntajes de los instrumentos psicológicos de valoración de ansiedad sean confiables.

El coeficiente alfa de Cronbach toma valores entre 0 y 1. Cuanto más se aproxime a su valor máximo 1, mayor será la fiabilidad de la escala. Sin embargo, valores superiores a 0.7 o 0.8 son suficientes para garantizar la fiabilidad de la escala.

Los resultados que se obtuvieron en el análisis de fiabilidad corresponden al cuestionario de preocupación del estado de Pensilvania y a la escala de detección del trastorno de ansiedad generalizada. Este análisis se realizó con el programa de IBM SPSSStatistics para obtener el coeficiente alfa de Cronbach de ambas escalas de valoración de ansiedad. El uso de este programa perteneciente a IBM ofrece elegir la obtención del coeficiente alfa de Cronbach de una lista

predeterminada, eliminando la necesidad de introducir fórmulas o generar código en programas como Jupyter Notebook o R.

A. Análisis de fiabilidad: cuestionario de preocupación del estado de Pensilvania

Para el caso del cuestionario de preocupación del estado de Pensilvania se obtuvo un coeficiente alfa de Cronbach de 0.963 para los 15 ítems resultantes (ítems 2-16) del análisis factorial (ver Figura 3.12).

Coeficiente alfa de Cronbach perteneciente al cuestionario de preocupación del estado de Pensilvania

Alfa de Cronbach	No. de ítems
.963	15

Figura 3.12. Datos obtenidos del coeficiente alfa de Cronbach perteneciente al cuestionario de preocupación del estado de Pensilvania.

B. Análisis de fiabilidad: escala de detección del trastorno de ansiedad generalizada

Para el caso de la escala de detección del trastorno de ansiedad generalizada se obtuvo un coeficiente alfa de Cronbach de 0.791 para los 6 ítems dicotómicos resultantes (ítems 1, 2, 3, 4, 5 y 10) del análisis factorial (ver Figura 3.13).

Coeficiente alfa de Cronbach perteneciente a la escala de detección del trastorno de ansiedad generalizada

Alfa de Cronbach	No. de ítems
.791	6

Figura 3.13. Datos obtenidos del coeficiente alfa de Cronbach perteneciente a la escala de detección del trastorno de ansiedad generalizada.

En ambos casos, el coeficiente alfa de Cronbach que se obtuvo demuestra que ambas escalas de valoración tienen una buena fiabilidad y se aceptan para

realizar la aplicación final a una población muestra para formar los grupos de conveniencia.

3.4.3 Resultados

Una vez terminados el análisis factorial y el análisis de fiabilidad, se procedió a generar ambos instrumentos solamente con los ítems que ambos análisis presentaron como relevantes para la obtención de un alfa de Cronbach mayor a 0.5. En el caso del cuestionario de preocupación del estado de Pensilvania se eliminó el ítem 1, quedando los ítems del 2 al 16 (ver Anexo III). Por otro lado, en el caso de la escala de detección del trastorno de ansiedad generalizada se eliminaron los ítems 6, 7, 8, 9, 11 y 12, quedando solamente los ítems 1, 2, 3, 4, 5 y 10 (ver Anexo IV).

Capítulo 4

Modelo de comportamiento del
trastorno de ansiedad generalizada

En este capítulo se presenta la Fase 2 de la metodología de solución perteneciente a la investigación. La Fase 2 presenta el proceso llevado a cabo para la obtención del modelo de comportamiento del trastorno de ansiedad generalizada. La Figura 4.1 presenta de manera ordenada y detallada cada etapa perteneciente a la Fase 2 de la metodología presentada en la Figura 1.2. Se recuerda que parte del objetivo general correspondiente a esta investigación es obtener un modelo de comportamiento del trastorno de ansiedad generalizada para adultos jóvenes.

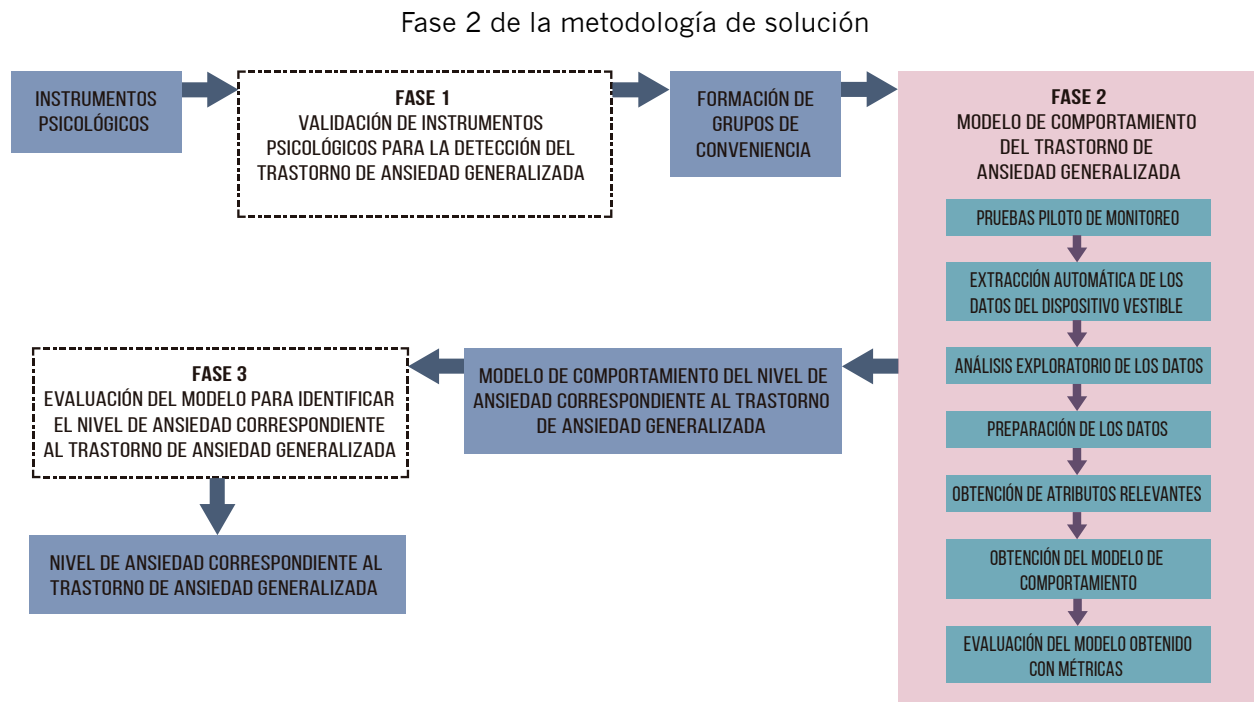


Figura 4.1. Determinación del modelo de comportamiento del nivel de ansiedad.

A continuación, se describen brevemente las principales etapas que conforman la Fase 2 correspondiente a la metodología de solución:

Para iniciar la Fase 2, los instrumentos psicológicos del trastorno de ansiedad generalizada se aplicaron por primera vez sólo para su validación como escalas de valoración. Una vez validadas, se aplicaron sólo con los ítems obtenidos por el análisis factorial (15 ítems para el cuestionario de preocupación del estado de Pensilvania y 6 ítems para la escala de detección del trastorno de ansiedad generalizada) a una muestra diferente a la validación de cuestionarios. Así, con los resultados obtenidos de la segunda aplicación de los dos instrumentos

psicológicos se formaron los grupos de conveniencia con respecto a los niveles de ansiedad correspondientes al trastorno de ansiedad generalizada definidos para esta investigación como: nivel bajo, nivel medio y nivel alto.

- **Formación de grupos de conveniencia.** Para iniciar la *Fase 2: determinación del modelo de comportamiento de ansiedad generalizada* se requiere formar los grupos de conveniencia para identificar a los participantes que formarán parte de las pruebas piloto de monitoreo.
- **Etapa 1. Pruebas piloto de monitoreo.** En esta etapa se inician las pruebas de monitoreo para los participantes elegidos de los diferentes grupos de conveniencia definidos anteriormente: nivel bajo, medio y alto.
- **Etapa 2. Extracción automática de los datos del dispositivo vestible.** En esta etapa se extraen los datos de la pulsera de actividad Fitbit para su posterior análisis con estadísticas descriptivas.
- **Etapa 3. Análisis exploratorio de los datos.** En esta etapa se realiza un análisis exploratorio de los datos para comprender los datos descargados del dispositivo vestible de manera gráfica o analítica con ayuda de métodos de visualización, estadísticas descriptivas y medidas de dispersión.
- **Etapa 4. Preparación de los datos.** En esta etapa los datos descargados de la pulsera son tratados para la selección y obtención de atributos relevantes. En esta fase se obtiene el modelo de comportamiento del trastorno de ansiedad generalizada. Esta etapa es de suma importancia; ya que los datos se deben limpiar y preparar para obtener el modelo general del comportamiento de ansiedad generalizada.
- **Etapa 5. Obtención de atributos relevantes.** Los datos anteriormente obtenidos de la etapa de preparación de los datos, son utilizados para generar el modelo de comportamiento de ansiedad generalizada con técnicas de aprendizaje automático: algoritmo árbol de decisión.
- **Etapa 6. Evaluación del modelo obtenido con métricas.** En esta etapa se evalúa el rendimiento del modelo obtenido por el algoritmo árbol de clasificación a

través de las métricas de evaluación pertenecientes al aprendizaje automático: exactitud, sensibilidad, precisión, valor F-1 y matriz de confusión.

4.1 Formación de grupos de conveniencia

De acuerdo con la Dra. en psicología Alicia Muniz (Muniz, 2018) la evaluación psicológica implica no sólo el valorar, validar, medir o comparar; también implica el conocer las características del evaluado. En este aspecto, para realizar una evaluación de aptitudes y características; tanto cognitivas como emocionales, es necesario partir de un paradigma que considere a la evaluación como un punto de partida y nunca como un destino final. Considerando este aspecto, la formación de grupos de conveniencia tuvo como objetivo identificar a posibles participantes que presentaran respuestas de déficit cognitivo conductual asociadas al trastorno de ansiedad generalizada.

4.1.1 Definición del perfil del grupo de estudio

El perfil del grupo de estudio consistió en cumplir con los criterios de inclusión y exclusión; además de ser alumnos pertenecientes al Centro Nacional de Investigación y Desarrollo Tecnológico, ubicado en la ciudad de Cuernavaca, Morelos, México. Los criterios de inclusión fueron los siguientes: i) edad entre 23 años y 30 años; ii) sin distinción de género; y iii) estudiantes del programa de maestría de Ciencias de la Computación ofrecido por el Centro Nacional de Investigación y Desarrollo Tecnológico. Los criterios de exclusión fueron los siguientes: i) mujeres embarazadas; ii) participantes mayores a 30 años o menores a 23 años; y iii) participantes en tratamiento por trastornos de ansiedad con medicamento recetado.

En este sentido, 9 participantes fueron excluidos: 1 participante femenino por presentar un estado de gestación, 2 participantes por ser menores de 23 años, 5 participantes por ser mayores a 30 años, y 1 participante por estar en tratamiento con medicamento recetado por trastornos de ansiedad. En total, sólo 25 participantes fueron tomados en cuenta para la formación de grupos de conveniencia.

4.1.2 Grupos de conveniencia

La formación de los grupos de conveniencia consistió en aplicar dos instrumentos psicológicos para la valoración de ansiedad: el cuestionario de preocupación del estado de Pensilvania para evaluar el rasgo de preocupación, y la escala de detección del trastorno de ansiedad generalizada para evaluar la presencia sintomatológica del trastorno de ansiedad generalizada. De estos dos instrumentos psicológicos aplicados, sólo se tomaron en cuenta los ítems seleccionados, y resultantes de su validación a través de los análisis factorial y de fiabilidad: 15 ítems para el caso del cuestionario de preocupación del estado de Pensilvania (ítems 2 al 16), y 6 ítems para el caso de la escala de detección del trastorno de ansiedad generalizada (ítems 1, 2, 3, 4, 5 y 10) (ver Anexo III y IV).

Un total de 25 participantes, 26 hombres y 8 mujeres contestaron ambos instrumentos psicológicos; además de proporcionar en una encuesta sus datos sociodemográficos (edad, sexo y grado de estudios). Estos participantes se clasificaron en tres niveles de ansiedad definidos por el cuestionario de preocupación del estado de Pensilvania: nivel bajo, medio y alto.

Los resultados para el cuestionario de preocupación del estado de Pensilvania fueron los siguientes: 17 participantes obtuvieron un nivel bajo de ansiedad, 6 participantes obtuvieron un nivel medio de ansiedad y sólo 2 obtuvieron un nivel alto de ansiedad. Del nivel bajo, 13 fueron hombres y 3 mujeres. Del nivel medio, 4 fueron hombres y 3 mujeres; mientras que del nivel alto 1 fue hombre y 1 mujer.

Los resultados para la escala de detección del trastorno de ansiedad generalizada fueron los siguientes: 18 participantes no obtuvieron presencia sintomatológica del trastorno de ansiedad generalizada; mientras que 6 participantes sí obtuvieron presencia sintomatológica. De los 18 participantes, 13 fueron hombres y 5 mujeres; mientras que de los 6 participantes, 4 fueron hombres y 2 mujeres.

De esta forma, de los 25 participantes que fueron clasificados de acuerdo a los resultados de los instrumentos de valoración, se seleccionaron 5 participantes con una edad promedio de 24.5 años; de los cuales 4 fueron hombres y 1 mujer.

Los grupos de conveniencia, así como los participantes seleccionados fueron los siguientes:

- 3 participantes del sexo masculino con una edad promedio de 23 años fueron seleccionados para el nivel de ansiedad bajo reportado por el cuestionario de preocupación del estado de Pensilvania;
- 1 participante del sexo femenino de 26 años fue seleccionado para el nivel de ansiedad medio reportado por el cuestionario de preocupación del estado de Pensilvania;
- 1 participante del sexo masculino de 25 años fue seleccionado para el nivel de ansiedad alto reportado por el cuestionario de preocupación del estado de Pensilvania.

Una vez los posibles participantes fueron identificados por el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada definidos para esta investigación (nivel bajo, nivel medio y nivel alto), se procedió a iniciar con las pruebas piloto de monitoreo.

4.2 Modelo de comportamiento

La obtención del modelo de comportamiento se logra gracias a diferentes etapas, las cuales son: i) pruebas piloto de monitoreo; ii) extracción automática de los datos del dispositivo vestible; iii) análisis exploratorio de los datos; iv) preparación de los datos; y v) obtención de atributos relevantes.

4.2.1 Pruebas piloto de monitoreo

Las pruebas piloto de monitoreo empiezan cuando se le proporciona a la persona seleccionada un consentimiento informado para que tenga noción de las etapas que conforman la prueba piloto de la que será partícipe. El participante debe firmar este consentimiento para proceder a realizar la prueba piloto de monitoreo.

Una vez que el consentimiento esté firmado, se procede a realizar las pruebas de monitoreo con el dispositivo vestible (para el caso de esta investigación la pulsera

de actividad Fitbit Inspire HR), descargar los datos, y analizarlos con el objetivo de obtener el modelo de comportamiento del trastorno de ansiedad generalizada.

Para esta investigación, la duración de las pruebas piloto se fijó por un periodo de 15 días con la finalidad de poder detectar si existe alguna alteración en los ciclos de sueño, así como permitir un mejor análisis a los registros de la frecuencia cardíaca.

4.2.1.1 Consentimiento informado

El consentimiento informado es un documento para actos médicos e investigaciones en salud que declara la voluntad expresada por el paciente, en donde se redacta información clara, precisa y adecuada respecto al procedimiento de las pruebas de monitoreo, así como de los beneficios esperados y los posibles riesgos, molestias y efectos adversos. Para el caso de esta investigación, el consentimiento informado (ver Anexo V) se redactó, y entregó a los participantes seleccionados después de haber definido y aplicado los instrumentos psicológicos para la valoración del trastorno de ansiedad generalizada (cuestionario de preocupación del estado de Pensilvania y escala de detección del trastorno de ansiedad generalizada).

El consentimiento informado perteneciente a esta investigación contiene lo siguiente:

A. Introducción

Este apartado contiene una breve introducción del tema a investigar, así como el nombre y el objetivo de la investigación. En este caso, al ser la ansiedad generalizada el tema principal de esta investigación, se introduce al participante sobre este trastorno de ansiedad, el cómo afecta la vida de las personas, y datos importantes de la investigación como el nombre y el objetivo.

B. Descripción del estudio

En este apartado se da una breve descripción general de las pruebas de monitoreo a realizar. En el caso de esta investigación se menciona un breve resumen en términos generales de la investigación.

C. Descripción de las pruebas de estudio

Este apartado contiene las diferentes etapas de las pruebas de monitoreo, así como una pequeña descripción para informar al participante de cuáles serán las actividades que le competen a su persona.

Para esta investigación se describen las etapas que conforman las pruebas de monitoreo, así como su descripción, y una pequeña leyenda explicando que se prestará equipo para las pruebas de monitoreo.

D. Riesgos y beneficios de la prueba de estudio

En este apartado se le informa al participante qué riesgos asociados hay respecto a las pruebas de monitoreo, y qué beneficios puede obtener al ser partícipe de la investigación.

E. Derechos

Este apartado le comunica al participante los derechos que posee al ser partícipe de las pruebas de monitoreo. Además de informarle que su participación es totalmente voluntaria.

Para las pruebas de monitoreo correspondientes a la investigación, se le hace una aclaración al participante que es posible retirarse de las pruebas de monitoreo en cualquier momento, o que se puede abstener de participar en cualquier fase.

F. Confidencialidad

En este apartado se le informa al participante que todos los datos obtenidos en las pruebas son totalmente confidenciales. También, se añade un pequeño texto con las personas con las que se puede comunicar para cualquier duda, queja, aclaración, o sugerencia.

Para esta investigación se hace la aclaración que toda la información recolectada será únicamente con fines propios a la investigación, y que se mantendrá plena confidencialidad de los datos personales. También, se añaden números telefónicos, y los e-mails de los investigadores que forman parte de las pruebas de monitoreo.

G. Firma

Este apartado contiene el nombre del participante, la fecha en la que se le da el consentimiento, su firma, y el documento de identificación oficial que avala su participación. Así como el nombre del responsable de la evaluación, su firma y la fecha en la que se da el consentimiento a firmar.

Para la investigación es muy importante obtener el consentimiento del participante para iniciar las pruebas de monitoreo. Sin embargo, también es importante explicarle en qué consiste la investigación de la que será partícipe, así como los riesgos y beneficios que pueden llegar a tener las pruebas de monitoreo.

4.2.2 Extracción automática de los datos del dispositivo vestible

Los datos extraídos de la pulsera de actividad (Fitbit Inspire HR) otorgada a los participantes que forman parte de las pruebas de monitoreo, son almacenados en un archivo CSV para su posterior procesamiento, con el objetivo de obtener el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada de cada participante. Así como obtener el modelo de comportamiento del trastorno de ansiedad generalizada.

4.2.2.1 Definición de variables

Para iniciar la extracción de los datos se necesitan definir las variables que se van a monitorear. Para el caso de esta investigación, las variables se definieron de acuerdo con los síntomas reportados por el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V), y la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE-10). De los síntomas reportados por ambos documentos, se eligieron los más representativos del trastorno de ansiedad generalizada y que pueden ser monitoreados a través de un dispositivo vestible.

Los síntomas que fueron elegidos como variables (reportadas por el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V), y la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la

Salud (CIE-10), y definidas por su grado de importancia al detectar los trastornos de ansiedad), se mencionan a continuación:

A. Frecuencia Cardíaca

La frecuencia cardíaca es uno de los síntomas primordiales cuando se trata de detectar la ansiedad (Pintelas et al, 2018); ya que la aparición súbita de ésta se puede manifestar a través de palpitaciones y un incremento a la frecuencia cardíaca (American Psychiatric Association, 2013).

B. Calidad del sueño

Uno de los síntomas que más afecta a las personas que sufren del trastorno de ansiedad generalizada es la dificultad para quedarse dormido o para continuar durmiendo, así como un sueño inquieto e insatisfactorio (American Psychiatric Association, 2013).

En esta investigación las variables correspondientes a la calidad del sueño son 10, las cuales son: duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos dormidos en etapa *sueño profundo* del ciclo del sueño, minutos dormidos en etapa *sueño ligero* del ciclo del sueño, minutos dormidos en etapa *sueño REM* del ciclo del sueño, los minutos despiertos durante el ciclo del sueño, el número de veces que hubo una interrupción en el ciclo del sueño y el tiempo transcurrido en cama (ver Glosario de variables).

C. Actividad física

En esta investigación se toma en cuenta la actividad física para analizar si hay alguna relación entre la actividad realizada y el trastorno de ansiedad generalizada. Artículos diversos en la literatura se han encargado de estudiar a profundidad la relación existente entre la actividad física y los trastornos de ansiedad (Guerra, Gutiérrez & Zavala, 2017). Dentro de esta variable global se encuentran diferentes variables específicas, tales como: el total de las calorías quemadas, las calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), el número de pasos dados, los pisos escalados, la distancia total recorrida, la elevación de los pasos, los minutos en actividad *sedentaria*, los minutos en actividad *ligeramente activa*, los minutos en actividad *moderadamente activa*, y los minutos en actividad *muy*

activa. En total son 10 las variables pertenecientes a la actividad física (ver Glosario de variables).

D. Sociodemográficas

Las variables sociodemográficas tomadas en cuenta del perfil de los participantes son 3, las cuales son: sexo, grado (maestría), y edad.

A partir de estas variables correlacionadas se pretende determinar el modelo de comportamiento de ansiedad general correspondiente al trastorno de ansiedad generalizada.

4.2.2.2 Definición del dispositivo vestible

Una vez definidas las variables a monitorear, se procede a buscar en el mercado el tipo de dispositivos que son capaces de medirlas. Para la comodidad de los participantes, se busca un dispositivo que no sea invasivo para el desarrollo de sus actividades diarias. En el caso de esta investigación, se decidió utilizar un dispositivo comercial y accesible económicamente de tipo pulsera de la marca Fitbit, modelo Inspire HR 2019. La marca Fitbit fue elegida al ser reportada en la literatura como un dispositivo vestible capaz de monitorear estrés, actividad física, calidad de sueño y trastornos de ansiedad (Peake, Kerr & Sullivan, 2018), (Tiwari, Cassani, & Narayanan, 2019). El modelo fue elegido al presentar mejoras en la tecnología y en sus algoritmos para monitorear las etapas de sueño (Fitbit, 2019). Con esta pulsera se monitorearon las variables pertenecientes a la frecuencia cardíaca, calidad del sueño y actividad física. Al ser un dispositivo de tipo pulsera a prueba de agua, resulta ser conveniente para evitar estresar a los participantes, y afectar de alguna manera sus actividades diarias.

4.2.2.3 Descarga de los datos de la pulsera Fitbit

La descarga de los datos se logra a través de la API proporcionada por el sitio dev.Fitbit.com (ver el sitio web <https://dev.fitbit.com/login>), así como del registro de una aplicación para obtener los datos almacenados en cualquier pulsera Fitbit. Una vez registrada la aplicación, los datos pueden ser descargados a través del cliente Python-Fitbit; el cual logra descargarlos de una manera simple

gracias al acceso de la API Fitbit utilizando Python (ver el sitio web <https://github.com/orcasgit/python-fitbit>).

Los datos descargados son almacenados en diferentes archivos CSV (frecuencia cardíaca, resumen total de actividad y etapas del ciclo del sueño). Estos archivos contienen toda la información recabada por el dispositivo Fitbit Inspire HR; entre los cuales destacan las variables pertenecientes a la actividad realizada, el registro de las horas dormidas, así como las diferentes etapas de sueño y su duración. Estos archivos también incluyen los registros de la frecuencia cardíaca.

4.2.3 Análisis exploratorio de los datos

El análisis exploratorio de los datos conocido por su nombre en inglés como *Exploratory Data Analysis*, tiene su primera etapa antes de la fase preparación de los datos; mientras que la segunda etapa del análisis es llevada a cabo una vez la fase de preparación de los datos ha terminado satisfactoriamente.

El análisis exploratorio de los datos es el primer paso para la construcción de cualquier modelo. Este análisis ayuda a entender los datos que se están descargando del dispositivo vestible de manera gráfica o analítica con ayuda de métodos de visualización, estadísticas descriptivas y medidas de dispersión.

En este análisis exploratorio de los datos es común que haya pérdida de datos. No obstante, la importancia de realizar este análisis antes de la preparación de los datos es para comprender a profundidad el comportamiento de los datos y los datos descargados del dispositivo vestible.

El análisis exploratorio de los datos se divide en la etapa listada a continuación:

A. Análisis univariante

El análisis univariante explora las variables una por una. El método a utilizar dependerá del tipo de variable (categórica o continua). En este tipo de análisis se implementan estadísticas descriptivas como la media, mediana, moda, los mínimos y máximos; así como las medidas de dispersión como el rango, rango intercuartílico, cuartiles, varianza, curtosis y asimetría; y los métodos de visualización como los histogramas y diagramas de cajas y

bigotes. Con este análisis se busca ver la cantidad de datos perdidos, de datos atípicos y la normalidad de los datos, así como el comportamiento de los datos (media, mediana, moda, entre otros). Además de comprender los datos que están descargando, se busca detectar si los datos requerirán un tratamiento especial (en el caso de datos atípicos o datos no normalizados).

En el análisis exploratorio de los datos perteneciente a esta investigación, se importan los datos por separado de los 5 participantes seleccionados para las pruebas de monitoreo en formato CSV, previamente descargados gracias a la API Python-Fitbit y al programa Jupyter Notebook para su visualización y análisis. Cada archivo perteneciente a cada participante es tratado aparte. El objetivo de realizar este análisis es para comprender más a profundidad cómo se están descargando los datos; además de ofrecer una visión general de su comportamiento.

La obtención de estadísticas descriptivas ayuda a comprender cómo se están comportando los datos, y si llegasen a existir datos atípicos, los diagramas de cajas y bigotes se encargarán de su detección temprana; aunque el tratamiento de datos atípicos se realiza en el procesamiento de los datos.

4.2.3.1 Análisis univariante

El análisis univariante perteneciente a esta investigación se realizó en diferentes etapas para cada uno de los 5 participantes previamente seleccionados. La primera etapa consistió en analizar los datos correspondientes a la frecuencia cardíaca, la segunda etapa consistió en analizar los datos correspondientes a las variables pertenecientes a la actividad física, y la tercera etapa consistió en analizar los datos correspondientes a las variables pertenecientes a la calidad del sueño. Las etapas fueron elegidas por el grado de importancia y complejidad, siendo los datos de la frecuencia cardíaca los que requerían de más tiempo para analizar, así como los datos más relevantes para la investigación.

A. Frecuencia cardíaca

El primer paso de este análisis univariante fue obtener estadísticas descriptivas de la frecuencia cardíaca de cada participante. En este paso se hace uso de las funciones automáticas del programa Jupyter Notebook.

De acuerdo con la Asociación Americana del Corazón (American Heart Association por su nombre en inglés) los parámetros de una frecuencia cardíaca normal se encuentran entre los 60 y los 100 latidos por minuto con una ligera diferencia entre hombres y mujeres. Se considera que una frecuencia cardíaca por arriba de los 100 latidos por minutos puede atribuirse a una arritmia, taquicardias inusuales o crisis de ansiedad (Association, 2015). Un corazón sano no late con la regularidad de un reloj. El corazón se acelera y disminuye para adaptarse a la necesidad cambiante de oxígeno dependiendo de las actividades de una persona a lo largo del día. En este sentido, el ritmo cardíaco o la frecuencia cardíaca *normal* varía de persona a persona (Harvard, 2018). La diferencia antes mencionada de la frecuencia cardíaca entre hombres y mujeres se debe al tamaño del corazón, el cual suele ser más pequeño en las mujeres que en los hombres. El corazón femenino bombea menos sangre con cada latido; pero necesita latir a un ritmo más rápido; ya que poseen una ritmicidad intrínseca diferente (European Heart Journal, 1998). Esto provocará que la frecuencia cardíaca, así como los datos atípicos, sean totalmente diferentes para cada participante, recalcando la ligera diferencia de la frecuencia cardíaca entre hombres y mujeres.

Parte de las estadísticas descriptivas obtenidas en el análisis perteneciente a los archivos de la frecuencia cardíaca son la media, la desviación estándar, los valores máximos y mínimos; la mediana y los valores percentiles correspondientes al 25%, 50% y 75%. Estas estadísticas descriptivas fueron obtenidas por cada uno de los 5 participantes con el comando *describe*, perteneciente a Jupyter Notebook.

El siguiente paso en el análisis univariante fue la obtención de datos atípicos para cada uno de los 5 participantes a través de un diagrama de cajas y bigotes para su futuro tratamiento. Una de las ventajas de usar diagramas de cajas y bigotes como método de visualización para datos atípicos, es la manera conveniente y rápida de mostrar, y examinar visualmente el conjunto de datos pertenecientes a la frecuencia cardíaca.

Como ejemplo, parte de los resultados obtenidos se observan en la Figura 4.2. Estos datos corresponden a los registros de la frecuencia cardíaca de un participante de los 5 monitoreados.

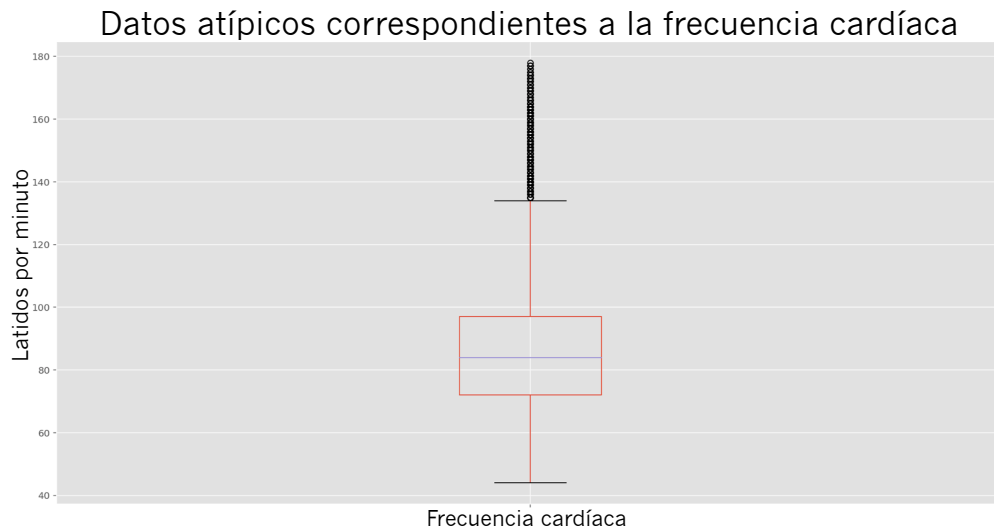


Figura 4.2. Datos atípicos correspondientes a la frecuencia cardíaca de un participante con nivel de ansiedad bajo correspondiente al trastorno de ansiedad generalizada.

En el diagrama de cajas y bigotes que se muestra en la Figura 4.2, se observa que los datos atípicos empiezan a ser notorios aproximadamente de los 135 a 180 latidos por minuto; mientras que la media de la frecuencia cardíaca ronda aproximadamente los 85 latidos por minuto. Gracias a este tipo de diagramas, es posible detectar tempranamente los datos atípicos.

B. Actividad física

En el caso de la etapa dedicada a las 10 variables correspondientes a la actividad (ver Glosario de variables); además de obtener estadísticas descriptivas, se obtienen gráficos que ayudan a comprender mejor los datos descargados.

En este sentido es importante entender cómo funcionan los diferentes niveles de actividad: sedentario, ligeramente activo, moderadamente activo y muy activo; ya que forman parte de las variables monitoreadas en esta investigación; además de ser variables que se descargan de la aplicación Fitbit.

La Figura 4.3 muestra los niveles de actividad por día de la semana perteneciente a un participante de los 5 monitoreados. Al igual que el análisis univariante perteneciente a los datos de la frecuencia cardíaca, el análisis univariante perteneciente a la actividad física fue realizado para cada uno de los 5 participantes monitoreados. En este análisis el objetivo es el ayudar a entender el nivel de actividad del participante; es decir, analizar y comprender las siguientes variables monitoreadas: actividad sedentaria, ligeramente activa, moderadamente activa y muy activa.

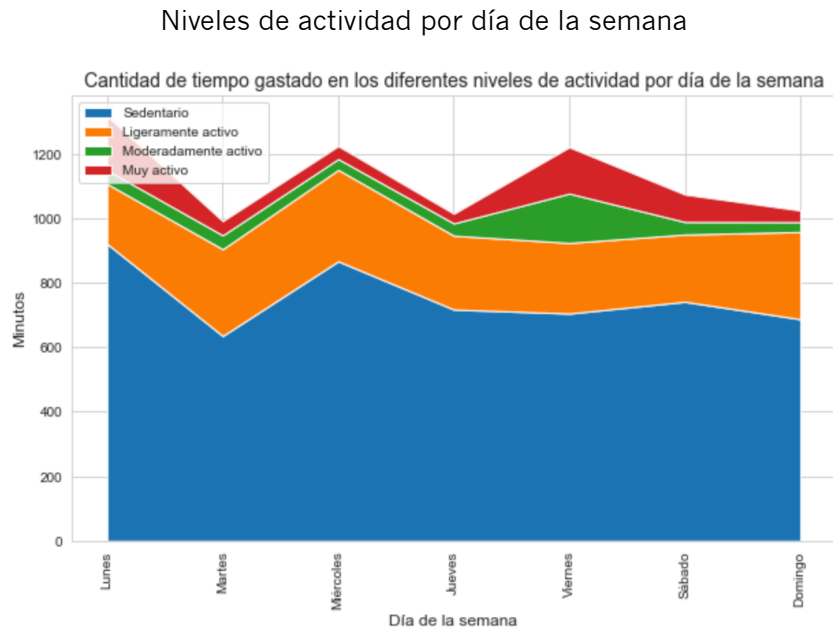


Figura 4.3. Niveles de actividad por día de la semana de un participante con nivel de ansiedad bajo correspondiente al trastorno de ansiedad generalizada.

Como se observa en la Figura 4.3, las variables pertenecientes a la actividad sedentaria, ligeramente activa, moderadamente activa y muy activa están dadas en minutos, al igual que cuando son descargadas desde el servidor de Fitbit. Gracias a este tipo de gráfica es posible observar el nivel de actividad que pertenece a cada participante, y con esto, verificar si alguna de estas variables están relacionadas con el nivel de ansiedad reportado para el participante. Si alguna de estas variables estuviera relacionada con el nivel de ansiedad, entonces se podría llegar a la conclusión de que la actividad física es un atributo relevante para el trastorno de ansiedad generalizada.

C. Calidad del sueño

En el caso de la etapa dedicada a las 10 variables correspondientes a la calidad del sueño; además de obtener estadísticas descriptivas; también se obtiene el número de horas en cama, así como la eficiencia del ciclo del sueño y el porcentaje de las etapas de sueño.

La eficiencia del sueño es un porcentaje que se obtiene de los datos descargados de Fitbit, así como de las demás variables obtenidas correspondientes al ciclo del sueño. Visualizar el porcentaje de las etapas de sueño y la eficiencia de los ciclos del sueño es importante; porque una de las características principales del trastorno de ansiedad generalizada es tener trastornos de sueño. Este trastorno se caracteriza por tener problemas de sueño; ya sea dificultad para dormir o para continuar durmiendo; así como un sueño inquieto e insatisfactorio. Al realizar el análisis univariante es posible entender cómo funciona el ciclo del sueño respecto a los datos descargados del dispositivo wearable. Al igual que los análisis correspondientes a la frecuencia cardíaca y a la actividad física, el perteneciente al ciclo del sueño fue realizado para cada uno de los 5 participantes monitoreados.

Tal como se muestra en la Figura 4.4, el participante tiene registro de 12 días dormidos (*count*=12), con una media de 6 horas, siendo el valor mínimo de 4 horas, dando un total de 74 horas dormidas aproximadamente en el periodo de monitoreo.

Se puede observar que 3 días durmió cerca de 4 horas, 2 días cerca de 5 horas, 3 días cerca de 6 horas, 2 días cerca de 7 horas y 1 día cerca de 9 horas. Sin embargo, hay un día que no se tiene registro de las horas dormidas y existen pérdidas de datos de los 3 días restantes del monitoreo.

Obtención de medidas centrales

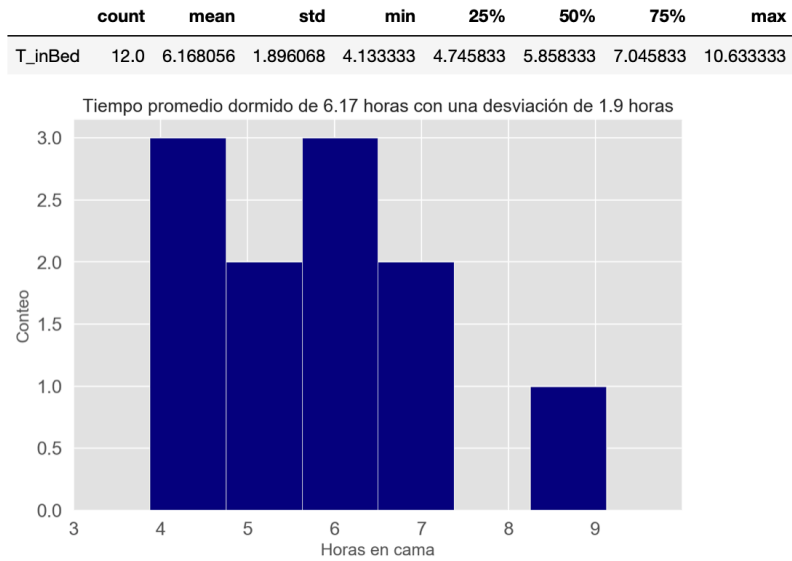


Figura 4.4. Obtención de medidas centrales (conteo = número de días monitoreados (12), media de horas dormidas, desviación estándar, valores máximos y mínimos, percentiles 25%, 50% y 75%) correspondientes a la etapa de calidad del sueño. La Figura 4.5 muestra 3 días donde el participante durmió 4 horas, 2 días donde el participante durmió 5 horas, 3 días donde el participante durmió 6 horas, 2 días donde el participante durmió 7 horas y 1 día donde el participante durmió 9 horas.

Gracias a la Figura 4.4 es posible saber fácilmente cuántas horas por día está durmiendo el participante monitoreado, detectando así, si existe alguna alteración en el ciclo del sueño perteneciente al participante.

Otra de las gráficas obtenidas muestra la eficiencia del ciclo del sueño de los 12 días registrados para el participante (ver Figura 4.5). La Figura 4.5 muestra la eficiencia en escala 0-100, siendo 0 una eficiencia nula, y 100 el valor máximo de un ciclo del sueño satisfactorio. En este caso, los valores registrados en eficiencia van del 86% al 96%.

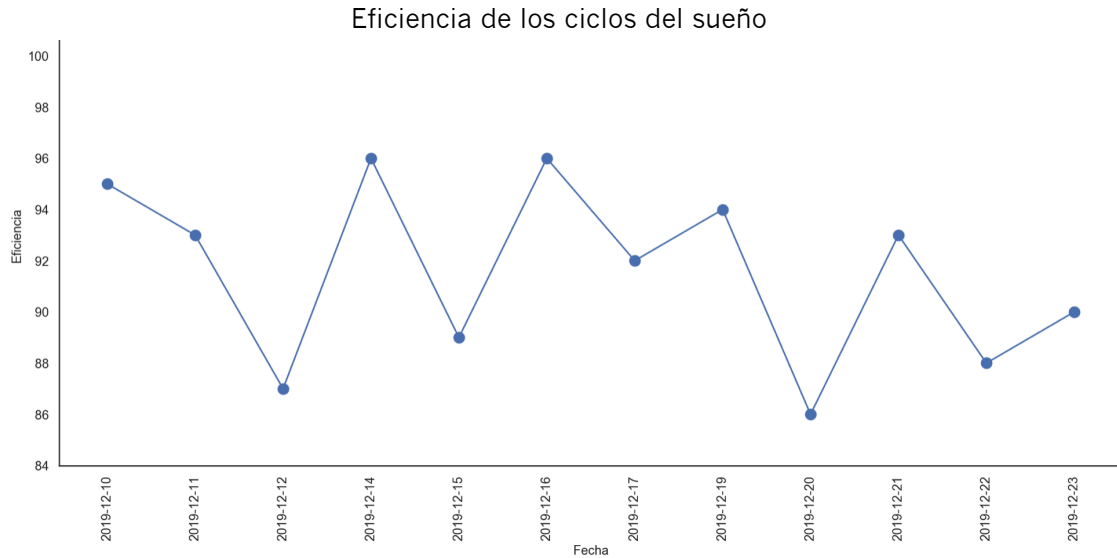


Figura 4.5. Eficiencia de los ciclos del sueño de un participante con nivel de ansiedad bajo correspondiente al trastorno de ansiedad generalizada.

La eficiencia del sueño, según explica Fitbit en su aplicación móvil, tiene relación con la calidad y la duración del ciclo del sueño. Esta eficiencia está clasificada de la siguiente manera:

- excelente > 90;
- bueno 80-90;
- favorable 60-79;
- pésimo < 60.

Con la gráfica perteneciente a la Figura 4.5 es posible saber fácilmente la eficiencia del ciclo del sueño correspondiente a cada participante por día monitoreado, detectando de manera sencilla si existe alguna alteración en el ciclo del sueño que pueda dar algún indicio relacionado a los trastornos del sueño. Por otro lado, no sólo las horas dormidas y la eficiencia del sueño son relevantes al analizar las variables pertenecientes al ciclo del sueño. Analizar los porcentajes de las diferentes fases del ciclo del sueño es importante; ya que algunas de estas fases ayudan a obtener el coeficiente de la eficiencia del sueño obtenido por cada día monitoreado.

Todas las personas tienen relojes biológicos que sincronizan sus ciclos de sueño. Estos relojes biológicos controlan en su mayoría los ritmos

circadianos que se repiten a intervalos de aproximadamente 24 horas. Estos ritmos circadianos, a su vez, son controlados por la luz. La incapacidad de los ojos de detectar la luz entorpecen a los ritmos circadianos, lo que deriva en problemas relacionados a los sueños crónicos (NINDS, 2018). Esto quiere decir que los patrones de sueño se regulan mediante los procesos que ocurren juntos: la necesidad de dormir y el reloj circadiano.

Cuando una persona duerme, el cerebro pasa por ciclos con cinco fases distintivas: fase 1, 2, 3, 4 y el sueño REM (sueño de movimientos oculares rápidos). Algunas de estas fases son necesarias para ayudar a que las personas se sientan descansadas y enérgicas al despertar; mientras que otras son necesarias para el proceso de almacenamiento de recuerdos y aprendizaje. El cerebro, al dormir, pasa de la fase 1 al sueño REM, y luego el ciclo vuelve a comenzar con la fase 1 (NINDS, 2018).

Las fases son brevemente descritas a continuación:

- Sueño no REM (75% del sueño): A medida que la persona comienza a quedarse dormida, entra en el sueño no REM, que consiste en las fases 1 a 4.
 - ✦ Fase 1
 - * etapa entre despierto y quedándose dormido.
 - ✦ Fase 2
 - * etapa con un sueño más profundo;
 - * la temperatura corporal desciende;
 - * se desconecta del entorno.
 - ✦ Fase 3 y 4
 - * etapa con un sueño más profundo y restaurador;
 - * el ritmo respiratorio desciende;
 - * los músculos se relajan;
 - * la energía se restablece.

- Sueño REM (25% del sueño): El sueño REM da inicio aproximadamente a los 70 a 90 minutos de que la persona se queda dormida, y después a intervalos sucesivos de aproximadamente 90 a 110 minutos. El sueño REM se hace más largo conforme transcurre la noche. Algunas de sus características son:
 - ✦ la temperatura corporal no está tan regulada;
 - ✦ el cuerpo se relaja y queda inmóvil;
 - ✦ el cerebro está activo y sueña.

De acuerdo con la información explicada por el Instituto Nacional de Enfermedades Neurológicas y Accidentes Cerebrovasculares (NINDS por sus siglas en inglés), el sueño REM debe ser el 25% del ciclo del sueño, y observando la Figura 4.6, el porcentaje obtenido de este sueño REM por un participante de los 5 monitoreados, es del 20.2%; mientras que el resto de las etapas deberían sumar un 75%, cuando en total se obtiene un 79.8% de sueño no REM (NINDS, 2018).

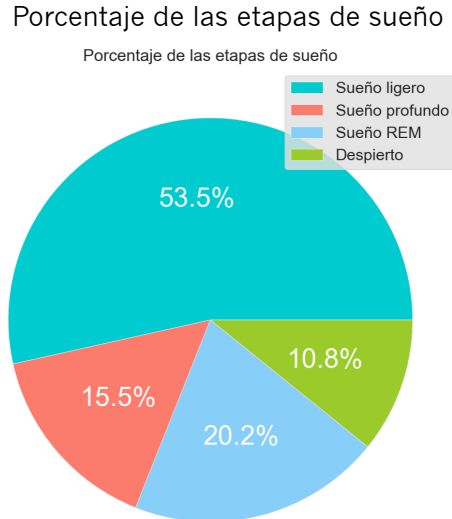


Figura 4.6. Porcentaje de las diferentes etapas de los ciclos del sueño.

Aumentar o disminuir este porcentaje perteneciente a las etapas de sueño REM y de sueño no REM está fuera de control de cualquier persona. Sin embargo, es posible que aplicando ciertos hábitos de salud se puedan

obtener mejorías: no consumir cafeína ni alcohol, o evitar hacer ejercicio cerca de la hora de dormir.

Por otra parte, el obtener este tipo de gráficas, como la de la Figura 4.6, es importante para detectar posibles anomalías en el ciclo del sueño perteneciente a cada participante monitoreado.

Como ya se ha mencionado a lo largo del desarrollo del análisis univariante perteneciente a las variables de la frecuencia cardíaca, a la actividad física y al ciclo del sueño, el objetivo es entender un poco más a profundidad los datos que son descargados de la pulsera Fitbit; además de observar el comportamiento de los datos, de detectar tempranamente si hay existencia de datos atípicos en los registros de la frecuencia cardíaca; y si existen anomalías en las horas dormidas como en la eficiencia del ciclo del sueño.

4.2.4 Preparación de los datos

El formato de los datos suele ser muy variable; ya que es probable que éstos tengan alguna inconsistencia; sobre todo cuando se tienen que descargar de dispositivos vestibles a través de alguna API. En este aspecto, es posible que los datos descargados sean inconsistentes, erróneos, o puede que haya una pérdida de datos o éstos estén incompletos, dificultando así la tarea de la preparación de los datos para la obtención de atributos relevantes, así como del modelo de comportamiento.

La fase de preparación de los datos es un proceso de varias etapas que comprenden diversos pasos individuales. Estos pasos son los siguientes:

A. Extracción de características y portabilidad

La extracción de características y portabilidad se vuelve una etapa importante; porque le da un formato homogéneo al conjunto de datos descargados. En ocasiones los datos sin procesar no están en un formato adecuado para su procesamiento. En algunos otros casos, los datos obtenidos provienen de diferentes fuentes, y deben integrarse en una única base de datos para su futuro procesamiento.

B. Limpieza de los datos

La limpieza de datos consiste en identificar los datos perdidos o incompletos, reemplazándolos con valores razonables, obtener estadísticas descriptivas como la media, mediana y moda, la desviación estándar, los valores mínimos y máximos, los cuartiles y la varianza (segunda parte del análisis exploratorio de los datos). También, se identifican, y se eliminan los valores atípicos denominados en inglés como *outliers*, así como la remoción de valores duplicados encontrados, resolviendo así todas las inconsistencias encontradas en los datos descargados.

C. Reducción, selección y transformación de los datos

La reducción, selección, y transformación de los datos es la fase donde el tamaño de los datos es reducido mediante la selección de subconjuntos de datos, la selección de subconjuntos de características o la transformación de datos. Una de las mayores ventajas de esta fase es buscar que los algoritmos sean más eficientes a través de la reducción del tamaño de los datos. Además, la calidad del proceso de minería de datos mejorará si se eliminan características o registros irrelevantes.

4.2.4.1 Extracción de características y portabilidad

La extracción de características y portabilidad es la creación de un conjunto de características con las que se pueda trabajar fácilmente en la preparación de los datos. En la mayoría de los casos los datos estarán en un formato crudo y no estructurado; es decir, serán datos que no tengan formato alguno. En otros casos, los datos se obtendrán en diferentes archivos o mantendrán un formato mixto de atributos que tendrán que ser discretizados; es decir, los datos tendrán valores numéricos.

En el caso de esta investigación, los archivos CSV descargados (resumen total de actividad y calidad del sueño) del dispositivo vestible para cada uno de los 5 participantes monitoreados (ver Glosario de variables), fueron combinados en un sólo archivo, dando un total de 20 variables. Un ejemplo del archivo con las 20 variables se puede descargar de la siguiente liga: <https://drive.google.com/file/d/1d09u7yyz--BWM1zcUhDpcxHBSO5qLCZg/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: ID del participante, duración

del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo del sueño, tiempo transcurrido en cama, total de calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados; distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria, minutos en actividad ligeramente activa, minutos en actividad moderada y minutos en actividad muy activa.

Por otro lado, el archivo CSV perteneciente a los registros de la frecuencia cardíaca fue tratado aparte. Los datos pertenecientes a la frecuencia cardíaca son descargados de los servidores de Fitbit en un archivo aparte; porque se descargan los valores de la frecuencia cardíaca por cada dos o tres segundos. Esto hace que no se pueda descargar en el mismo archivo que los datos pertenecientes a la actividad física o a la calidad del sueño. Un ejemplo del archivo descargado de los servidores de Fitbit se puede descargar de la siguiente liga: <https://drive.google.com/file/d/1Aal2ZS3zsUrvfI9IXieJtKMrRebWA4-X/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: tiempo y frecuencia cardíaca.

La variable de frecuencia cardíaca se dividió en 2 variables: media y desviación estándar. Estas dos variables fueron obtenidas de los registros pertenecientes a la frecuencia cardíaca de cada participante. De esta forma, se agregaron 2 variables a las 20 pertenecientes a la actividad física y al ciclo del sueño, dando un total de 22 variables (ver Glosario de variables). Un ejemplo del archivo con las 22 variables se puede descargar de la siguiente liga: <https://drive.google.com/file/d/1VVWnBTTt1efKv0B1DtAN5z5nsI7lw4Kk/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: ID del participante, media de la frecuencia cardíaca por día, desviación estándar de la frecuencia cardíaca por día, duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo

del sueño, tiempo transcurrido en cama, total de calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados, distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria minutos en actividad ligeramente activa, minutos en actividad moderada y minutos en actividad muy activa.

En el archivo combinado de las variables pertenecientes a la actividad física, calidad del sueño y frecuencia cardíaca, se añadieron las variables sociodemográficas: sexo, edad y grado de estudios (maestría o doctorado) del participante, dando un total de 25 variables (ver Glosario de variables). Un ejemplo del archivo con las 25 variables se puede descargar de la siguiente liga: https://drive.google.com/file/d/1yDpskkmQt8DVdhPR_rsvDLs1HmBMIn02/view?usp=sharing. Las variables contenidas, por orden de izquierda a derecha son: ID del participante, sexo (discretizado: femenino = 0 y masculino = 1), grado (maestría = 0), edad (23-25 años = 0 y 26-30 años = 1), media de la frecuencia cardíaca por día, desviación estándar de la frecuencia cardíaca por día, duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo del sueño, tiempo transcurrido en cama, total de calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados, distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria, minutos en actividad ligeramente activa, minutos en actividad moderada y minutos en actividad muy activa.

4.2.4.2 Limpieza de los datos

El objetivo de esta etapa es limpiar los datos obtenidos del dispositivo vestible, así como el resolver cualquier inconsistencia encontrada en los datos. Estos datos pueden estar incompletos; ya sea por errores computacionales, errores humanos, o errores del dispositivo vestible al no estar validado clínicamente por la Administración de Alimentos y Medicamentos, mejor conocida en sus siglas en inglés como FDA.

A pesar de que algunos algoritmos pueden trabajar con datos incompletos, siempre es mejor hacer la limpieza de datos para obtener mejores resultados. Sin embargo, el tratamiento de los datos siempre va a depender del tipo de algoritmo elegido para la obtención de atributos relevantes y del modelo de comportamiento.

Los datos perdidos significan una pérdida de información; es decir, son valores que no fueron almacenados para un atributo. Estos datos ocultan la habilidad de explicar y entender a fondo el problema a resolver. En el caso de la pérdida de datos en dispositivos vestibles se puede deber a falta de pila, falta de conectividad, y errores debidos a la obtención de datos del servidor de la compañía del dispositivo vestible. La mayoría de los dispositivos comerciales tienden a tener un rango de error que favorece a la pérdida de datos; ya que sus lecturas no suelen ser igual de precisas que los dispositivos dedicados a la investigación clínica o a los dedicados a las ciencias de la salud.

La limpieza de los datos empieza identificando los valores perdidos o incompletos, valores que son etiquetados como NAN (*not a number*, de acuerdo con sus siglas en inglés). Una vez que estos valores son identificados, pueden ser ignorados o reemplazados por constantes, por la media del atributo, o por la media de la clase.

Para esta investigación, se realizó la limpieza de los datos en dos diferentes etapas. La primera etapa se realizó con el archivo correspondiente a los registros de la frecuencia cardíaca; mientras que la segunda etapa correspondió al archivo combinado correspondiente al resumen total de actividad y a la calidad del sueño (20 variables).

Los registros de la frecuencia cardíaca tomados en cuenta para las pruebas piloto de monitoreo fueron únicamente aquellos en donde el participante se mantuvo despierto; es decir, que los registros pertenecientes al ciclo del sueño fueron eliminados. Los registros de la frecuencia cardíaca reportados por Fitbit como pertenecientes a alguna actividad física, también fueron eliminados. En este aspecto, uno de los archivos descargados de la API perteneciente a Fitbit ofrece las actividades físicas realizadas por el participante (algunas de las actividades físicas que reconoce Fitbit son: caminar, correr, nadar, o montar en bici (Fitbit,

2019)), haciendo más fácil la eliminación de los registros de la frecuencia cardíaca asociados a una actividad física. Un ejemplo del archivo descargado con los registros pertenecientes a la actividad física se puede descargar de la siguiente liga: <https://drive.google.com/file/d/1BS7u1WgMxelZ7EOaS8MI0fbRtClfMtb7/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: fecha, día de la semana, tiempo, nombre de la actividad, calorías quemadas durante la actividad, duración de la actividad, media de la frecuencia cardíaca durante la actividad, minutos en la zona fuera de zona perteneciente a la frecuencia cardíaca, minutos en la zona quema de grasa perteneciente a la frecuencia cardíaca, minutos en la zona cardio perteneciente a la frecuencia cardíaca y minutos en la zona pico perteneciente a la frecuencia cardíaca.

Para la primera etapa correspondiente a los registros de la frecuencia cardíaca se localizaron los valores perdidos, que gracias al programa Jupyter Notebook fueron reemplazados, en caso de existir, por la etiqueta NAN (*not a number*, de acuerdo con sus siglas en inglés). La Figura 4.7(a) muestra un ejemplo de la tabla correspondiente a los registros de la frecuencia cardíaca importada al programa Jupyter Notebook. El siguiente paso consistió en reemplazar los datos perdidos por la media del día correspondiente (ver Figura 4.7(b)).

Valores pertenecientes a la frecuencia cardíaca

Tiempo	hr	fecha	año	mes	día
0	12:00:03	66	20191219	2019	12 19
1	12:00:13	71	20191219	2019	12 19
2	12:00:18	NaN	20191219	2019	12 19
3	12:00:23	NaN	20191219	2019	12 19
4	12:00:28	69	20191219	2019	12 19

Tiempo	hr	fecha	año	mes	día
0	12:00:03	66	20191219	2019	12 19
1	12:00:13	71	20191219	2019	12 19
2	12:00:18	73.5	20191219	2019	12 19
3	12:00:23	73.5	20191219	2019	12 19
4	12:00:28	69	20191219	2019	12 19

Figura 4.7(a)

Figura 4.7(b)

La Figura 4.7(a) muestra valores pertenecientes a la frecuencia cardíaca como valores perdidos etiquetados como NAN; mientras que la Figura 4.7(b) muestra los valores perdidos reemplazados por la media de la frecuencia cardíaca: 73.5

Figura 4.7. Tabla correspondiente a los registros de la frecuencia cardíaca (hr).

La Figura 4.7(a) es un ejemplo de cómo se obtienen los datos de la frecuencia cardíaca al ser importados en el programa de Jupyter Notebook. Como se observa en la fila 2 y 4, hay valores perdidos etiquetados como NAN. La Figura

4.7(b) es un ejemplo de cómo los valores NAN son reemplazados con la media de la frecuencia cardíaca, que en este caso es de 73.5. La Figura 4.8 muestra el comando para reemplazar estos valores.

Comando correspondiente al reemplazo de los valores etiquetados como NAN

```
def summary_stats(stats_df, hr):
    stats_df = pd.DataFrame({
        'Mean' : str(round(np.mean(stats_df[hr]), 2))
    }, index=[0])
    return stats_df
```

```
summary_stats(df, 'hr')
```

	Mean
0	73.5

```
df = df.copy()
df['hr'].fillna(df['hr'].mean(), inplace = True)
```

Figura 4.8. Comando correspondiente al reemplazo de los valores pertenecientes a la frecuencia cardíaca etiquetados como NAN por la media de la frecuencia cardíaca (hr).

En la Figura 4.8 se puede apreciar el comando con el que los valores NAN de la frecuencia cardíaca fueron reemplazados por la media de la frecuencia cardíaca: 73.5, que pueden apreciarse en la Figura 4.7, en las filas 2 y 3, encerradas en un cuadro rojo. Después de reemplazar los valores perdidos se obtuvieron estadísticas descriptivas para ver el comportamiento de los datos después de tratar los datos perdidos (ver Figura 4.9 y 4.10).

La Figura 4.9 muestra los registros de la frecuencia cardíaca de un participante de los 5 monitoreados. La línea roja representa a la media: 85.65, la línea amarilla representa la mediana: 84, y la línea verde representa a la moda: 83. A simple vista, puede que los tres valores se mantengan en un rango cercano; sobre todo la media con un valor de 85.65, y la mediana con un valor de 84. Sin embargo, ambos valores no coinciden, demostrando que la muestra no es homogénea; ya sea por la presencia de sesgos en la distribución, o de valores extremos; los cuales se pueden observar en la gráfica a partir de los 140 latidos por minuto.

Estadística descriptiva perteneciente a los registros de la frecuencia cardíaca

Media: 85.65199898385845
Mediana: 84.0
Moda: 83

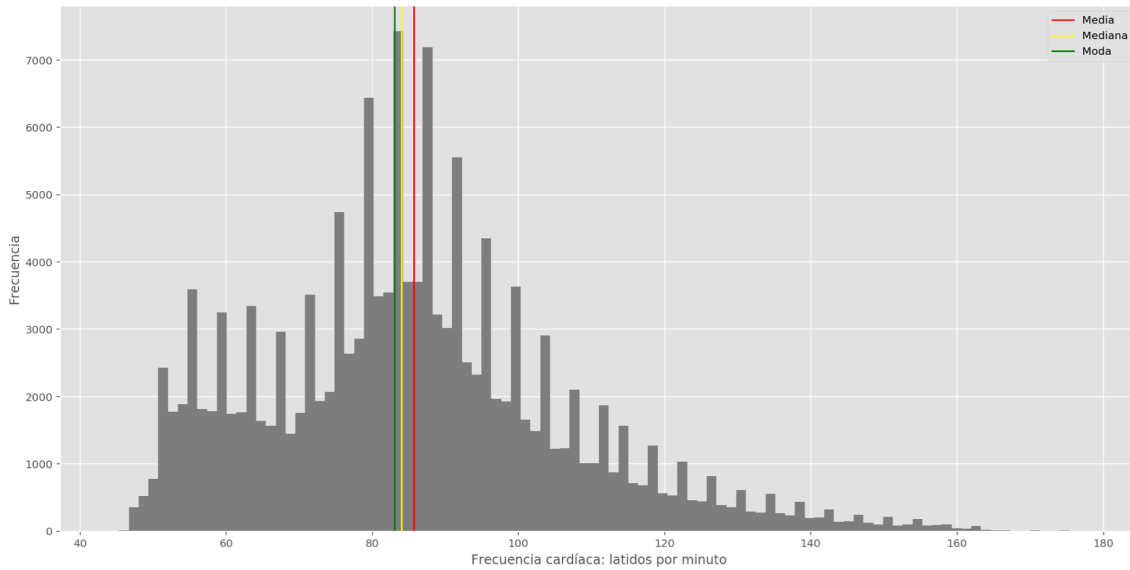


Figura 4.9. Estadística descriptiva perteneciente a los registros de la frecuencia cardíaca después de reemplazar datos perdidos.

En cambio, la Figura 4.10 muestra los mismos registros de la frecuencia cardíaca del mismo participante de los 5 monitoreados. No obstante, en esta Figura se observan los datos con su distribución actual y la distribución normal que deberían tener los datos. En este aspecto, observando la distribución de estos datos, se hace evidente que los datos pertenecientes a la frecuencia cardíaca deben ser normalizados antes de utilizar cualquier algoritmo de aprendizaje automático.

Cabe recalcar que la obtención de datos etiquetados como NAN, así como el reemplazo de los datos por la media y la obtención de estadísticas descriptivas, se realizó para cada uno de los 5 participantes.

Distribución de los datos correspondiente a los registros de la frecuencia cardíaca

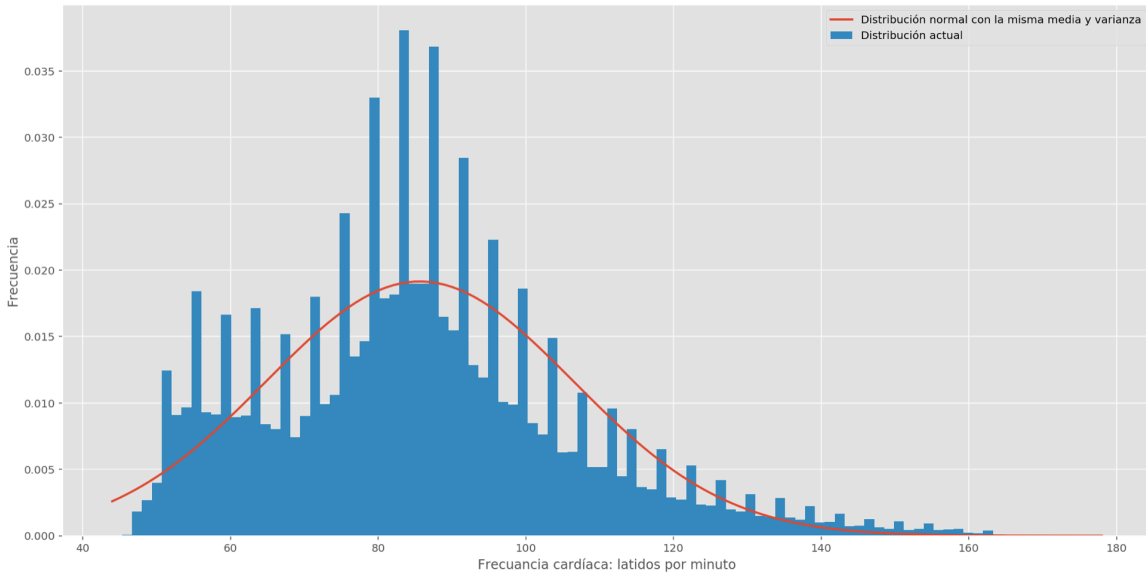


Figura 4.10. Distribución de los datos correspondiente a los registros de la frecuencia cardíaca después de reemplazar datos perdidos.

Posteriormente, se convirtió la variable *fecha* al formato *date time* 'año-mes-día' para agrupar los registros de la frecuencia cardíaca por día (ver Figura 4.11 (a)), y agrupar estos registros por la media del día (ver Figura 4.11 (b)). La agrupación de los registros de la frecuencia cardíaca por día y la obtención de la media por día, tiene el objetivo de obtener, también, la desviación estándar por día, y agregar estas dos columnas (media y desviación estándar de los registros pertenecientes a la frecuencia cardíaca) al archivo combinado correspondiente a las variables pertenecientes a la actividad física (10 variables), y a la calidad del sueño (10 variables) (ver Figura 4.11 (c)).

Transformación en el formato fecha

Tiempo	hr	fecha	año	mes	día
12:00:03	66	2019-12-19	2019	12	19
12:00:13	71	2019-12-19	2019	12	19
12:00:18	72	2019-12-19	2019	12	19
12:00:23	71	2019-12-19	2019	12	19
12:00:28	69	2019-12-19	2019	12	19

Figura 4.11(a)

fecha	media_hr	año	mes	día
2019-12-09	87.818593	2019.0	12.0	9.0
2019-12-10	80.135733	2019.0	12.0	10.0
2019-12-11	85.267189	2019.0	12.0	11.0
2019-12-12	82.573287	2019.0	12.0	12.0
2019-12-13	91.015732	2019.0	12.0	13.0

Figura 4.11(b)

fecha	std_hr
2019-12-09	15.364156
2019-12-10	17.025403
2019-12-11	17.939097
2019-12-12	14.560287
2019-12-13	18.371149

Figura 4.11(c)

La Figura 4.11(a) muestra la transformación en el formato de la fecha. La Figura 4.11(b) muestra la agrupación de la media por día perteneciente a la frecuencia cardíaca. La Figura 4.11(c) muestra la agrupación de la desviación estándar por día perteneciente a la frecuencia cardíaca.

Figura 4.11(a), (b) y (c) correspondientes a los registros de la frecuencia cardíaca. (hr significa frecuencia cardíaca; mientras que std_hr significa desviación estándar).

Por otro lado, la limpieza de los datos perteneciente al archivo combinado correspondiente a las variables pertenecientes a la actividad física (10 variables) y a la calidad del sueño (10 variables), se empezó con la discretización de los datos pertenecientes a las variables: sexo, grado (maestría) y edad.

En este aspecto, el sexo femenino se discretizó con el valor 0; mientras que el sexo masculino se discretizó con el valor 1. En cuanto al grado, maestría se discretizó con el valor 0. La Figura 4.12 muestra el antes y después del proceso de discretización. En este paso la variable edad se agrupó en rangos de 5 años; es decir, se agrupó de (20-25] años y (25-30] años. El primer rango (20-25] se discretizó con el valor 0; mientras que el segundo valor (25-30] se discretizó con el valor 1.

Proceso de discretización de los datos: antes y después

	ID	sexo	grado	edad
0	NH0K	Male	Master	23
1	NH0K	Male	Master	23
2	NH0K	Male	Master	23
3	NH0K	Male	Master	23
4	NH0K	Male	Master	23

Figura 4.12(a)

	ID	sexo	grado	edad
0	3	1	0	0
1	3	1	0	0
2	3	1	0	0
3	3	1	0	0
4	3	1	0	0

Figura 4.12(b)

La Figura 4.12(a) muestra el antes del proceso de discretización de los datos; mientras que la Figura 4.12(b) muestra el después del proceso de discretización de los datos.

Figuras 4.12(a) y (b) muestran el antes y después del proceso de discretización de los datos.

Posteriormente se procedió a detectar los datos perdidos que el programa Jupyter Notebook etiqueta automáticamente con la etiqueta NAN (*not a number*, de acuerdo con sus siglas en inglés). La Figura 4.13 muestra un ejemplo de datos perdidos que posteriormente fueron reemplazados por la media de cada variable.

Datos perdidos etiquetados automáticamente como NAN

Duration	Efficiency	M_Asleep	M_Awake	Sleep_Deep	Sleep_Light	Sleep_REM	Sleep_stage_Wake	Sleep_stage_Wake_Count	T_Bed	Calories_total	C
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1728
15600000.0	95.0	232.0	28.0	30.0	154.0	48.0	28.0	10.0	260.0	2472	
22080000.0	93.0	330.0	38.0	69.0	186.0	75.0	38.0	16.0	368.0	2481	
20940000.0	87.0	309.0	40.0	74.0	169.0	66.0	40.0	25.0	349.0	2204	
NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2843

Figura 4.13. Tabla correspondiente a datos perdidos etiquetados automáticamente como NAN por el programa Jupyter Notebook.

Los valores perdidos etiquetados como NAN en la Figura 4.13 se encuentran dentro de un cuadro rojo, mismos que pueden observarse por el reemplazo de la media de cada variable en la Figura 4.14. El comando utilizado para tratar los datos pertenecientes a las variables de actividad física (10 variables) y calidad de sueño (10 variables), fue el mismo que el mostrado en la Figura 4.8.

Es importante que los datos perdidos sean tratados antes de utilizar cualquier algoritmo de aprendizaje automático. También, discretizar los valores ayuda a implementar técnicas de aprendizaje automático; ya que varias de estas técnicas no son compatibles con variables categóricas de tipo nominal como las variables pertenecientes al perfil sociodemográfico de cada participante (edad, sexo y grado de estudios).

Una vez detectados estos datos perdidos, se procedió a reemplazarlos por la media de cada variable (ver Figura 4.14), terminando así la limpieza de los datos correspondientes a los registros de la frecuencia cardíaca, actividad física y calidad del sueño.

Reemplazo de datos perdidos por la media

Duration	Efficiency	M_Asleep	M_Awake	Sleep_Deep	Sleep_Light	Sleep_REM	Sleep_stage_Wake	Sleep_stage_Wake_Count	T_Bed	Calories_total	C
22292727.27273	91.72727	331.54545	40.00000	58.54545	199.18182	73.81818	40.00000	19.18182	371.54545	1728	
15600000.00000	95.00000	232.00000	28.00000	30.00000	154.00000	48.00000	28.00000	10.00000	260.00000	2472	
22080000.00000	93.00000	330.00000	38.00000	69.00000	186.00000	75.00000	38.00000	16.00000	368.00000	2481	
20940000.00000	87.00000	309.00000	40.00000	74.00000	169.00000	66.00000	40.00000	25.00000	349.00000	2204	
22292727.27273	91.72727	331.54545	40.00000	58.54545	199.18182	73.81818	40.00000	19.18182	371.54545	2843	

Figura 4.14. Tabla correspondiente al reemplazo de los datos perdidos por la media.

Un ejemplo del archivo resultante con la media y desviación estándar por día de la frecuencia cardíaca, y la limpieza de los datos correspondientes a las variables pertenecientes a la actividad física y calidad del sueño, se puede descargar de la

siguiente liga: <https://drive.google.com/file/d/1kd2551-aB3hotkmrZIYVDhiBfiiYEgs4/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: ID del participante, sexo (discretizado: femenino = 0 y masculino = 1), grado (maestría = 0), edad (23-25 años = 0 y 26-30 años = 1), media de la frecuencia cardíaca por día, desviación estándar de la frecuencia cardíaca por día, duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo del sueño, tiempo transcurrido en cama, total de calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados, distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria, minutos en actividad ligeramente activa, minutos en actividad moderada y minutos en actividad muy activa.

4.2.4.3 Tratamiento de datos atípicos

Los valores atípicos (en inglés llamados *outliers*) son valores que son numéricamente distantes del resto de los datos; y que pueden ser indicativos de datos que pertenecen a una población que es diferente del resto de las muestras establecidas. El método que se implementó en esta investigación para darle tratamiento a los datos atípicos fue la prueba de Tukey con ayuda del diagrama de cajas y bigotes. Este diagrama asocia cinco medidas descriptivas: la mediana, el primer cuartil (Q1), el tercer cuartil (Q3), el valor máximo y el valor mínimo; presentando al mismo tiempo información sobre la dispersión, la simetría de los datos de estudio y la tendencia central. Además de identificar con claridad las observaciones que se alejan de manera poco usual del resto de los datos, detectando así los datos atípicos, permitiendo comparar a la vez, diversos grupos de datos sin perder información.

El método para obtener el primer y tercer cuartil es el siguiente: primero se encuentra la media del conjunto de datos, después la mediana de los valores iguales o inferiores, así como la mediana de los valores superiores. Este será un valor de datos o será la mitad de dos valores, dependiendo si la cantidad de datos es impar o par. En el caso de un conjunto de datos impar, se debe incluir la

mediana en cada una de las dos mitades del conjunto de datos, para después encontrar el medio de cada mitad, dando como resultado el primer y tercer cuartil. En el caso de un conjunto de datos par, los datos se dividen en dos mitades, y se encuentra el medio de cada mitad. La obtención del rango intercuartil (IQR) está dado por: la diferencia del tercer cuartil (Q3) menos el primer cuartil (Q1), es decir $IQR = Q3 - Q1$.

Para la detección de los datos atípicos, la longitud máxima de cada uno de los bigotes debe ser 1.5 veces el rango intercuartil (IQR); es decir, $1.5(Q3 - Q1)$, tanto por debajo como por encima de los cuartiles. En este caso, las observaciones fuera de los bigotes son dibujadas separadamente y etiquetadas como valores atípicos.

En el caso de esta investigación, se obtienen los valores de los cuartiles, así como los valores del rango intercuartil (IQR) y los valores superior e inferior para los 5 participantes. La Figura 4.15 es un ejemplo de los valores obtenidos para un participante. En este caso el primer cuartil es una frecuencia cardíaca de 72 latidos por minuto, el segundo cuartil o la media, es una frecuencia cardíaca de 84 latidos por minuto, y el tercer cuartil es una frecuencia cardíaca de 97 latidos por minuto. El rango intercuartílico es de 25; mientras que el valor inferior es de 34.5 latidos por minuto, y el valor superior es de 134.5 latidos por minuto.

Valores correspondientes al primer, segundo, tercer cuartil, rango intercuartílico, y valores inferior y superior

```
# First Quartile
print('Primer cuartil', Q1)
print('Segundo cuartil', Q2)
print('Tercer cuartil', Q3)
print('IQR', iqr )
print('Inferior' , lower)
print('Superior' , upper)

Primer cuartil 72.0
Segundo cuartil 84.0
Tercer cuartil 97.0
IQR 25.0
Inferior 34.5
Superior 134.5
```

Figura 4.15. Obtención de los valores correspondientes al primer, segundo y tercer cuartil, así como del rango intercuartílico y los valores: inferior y superior.

Una vez obtenidos los valores anteriormente mencionados, se procede a filtrar los datos atípicos pertenecientes a la frecuencia cardíaca, utilizando una comparación entre el diagrama de cajas y bigotes obtenido inicialmente con el

diagrama resultante con los datos atípicos filtrados con la prueba de Tukey (ver Figura 4.16). Tal como lo muestra la Figura 4.16, los datos por arriba de los 134.5 latidos por minuto fueron filtrados de los registros pertenecientes a la frecuencia cardíaca. En este ejemplo específico, se logró tratar 3,789 datos considerados como atípicos de 145, 649. Del lado izquierdo de la Figura 4.16, se puede apreciar el diagrama correspondiente al conjunto de datos original, mostrando datos atípicos por arriba de los 134.5 latidos por minuto; mientras que del lado derecho, se puede apreciar el diagrama correspondiente al tratamiento de los datos atípicos con la prueba Tukey, reduciendo la cantidad de datos atípicos encontrados por arriba de los 134.5 latidos por minuto.

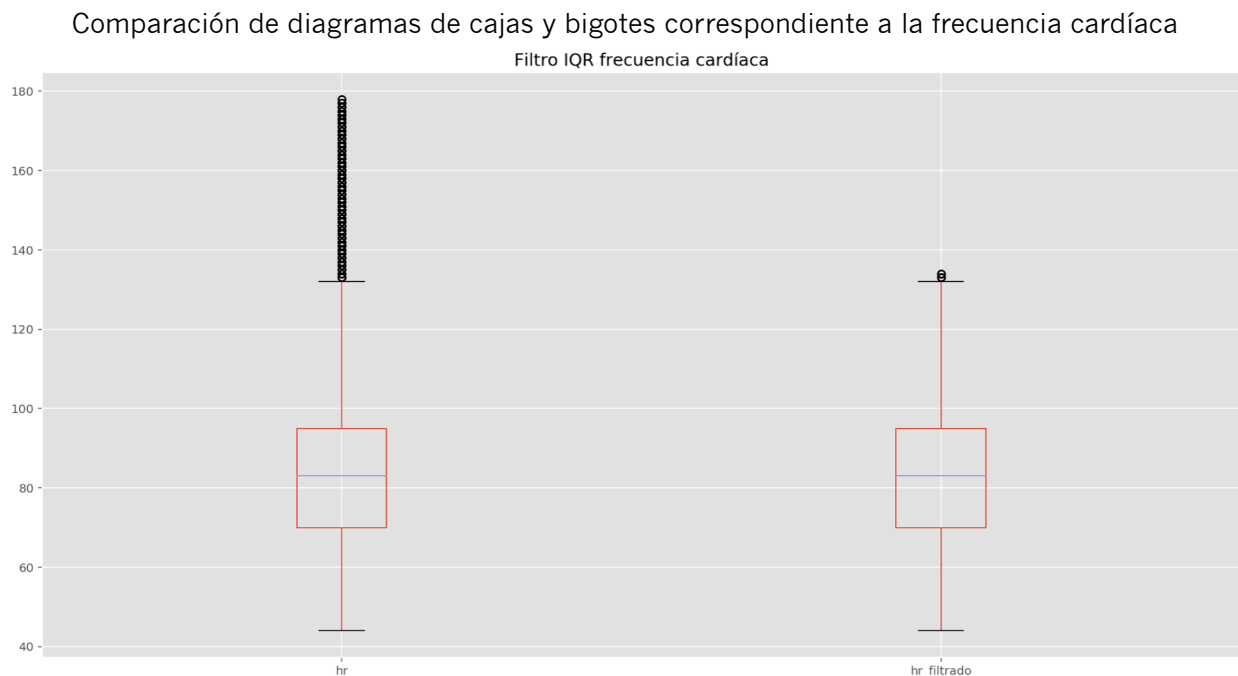


Figura 4.16. Comparación de diagramas de cajas y bigotes correspondiente a la frecuencia cardíaca. Del lado izquierdo se muestra el diagrama para el conjunto de datos original, mientras que del lado derecho se muestra el conjunto de datos filtrados utilizando la prueba Tukey.

4.2.4.4 Reducción, selección y transformación de los datos

La reducción, selección y transformación de los datos tiene como objetivo representar los datos de manera más compacta; ya que es más fácil aplicar algoritmos cuando el tamaño de los datos es menor. La reducción de los datos puede ser tanto en términos del número de filas (registros), como en términos del número de columnas (dimensionales).

En esta fase se llevó a cabo el tratamiento al desbalance de los datos; es decir, al tratamiento de la proporción de las diferentes clases, que, en el caso de esta investigación, se refiere a las clases pertenecientes a los distintos niveles de ansiedad correspondientes al trastorno de ansiedad generalizada: nivel bajo, nivel medio y nivel alto. Los clasificadores de aprendizaje automático suelen tener problemas con los conjuntos de datos desbalanceados; ya que son sensibles al conjunto de datos de una clase con un gran número de instancias y otra clase con muy pocas instancias. En algunas ocasiones, estos algoritmos favorecen a la clase mayoritaria, aquella clase que posee la mayor proporción de observaciones o valores; mientras que tienden a ignorar las clases pequeñas. Para mejorar esta proporción numérica de datos desbalanceados, se utilizó la aproximación de sobremuestreo sintético, mejor conocido como SMOTE, con el objetivo de añadir nuevas instancias a la clase minoritaria, detectada como el nivel medio de ansiedad.

Dentro de esta fase, también se llevó a cabo la normalización de los datos; porque en muchos escenarios, las diferentes características representan diferentes escalas de referencia, y, por lo tanto, pueden no ser comparables entre sí. La normalización de los datos perteneciente a las 25 variables monitoreadas en esa investigación se realizó con el comando *StandardScaler*, perteneciente a la librería *scikit-learn*. Este comando estandariza un atributo restando la media y luego escalado a la varianza de la unidad. La varianza de la unidad significa dividir todos los valores por la desviación estándar. Esta división da como resultado una distribución con una desviación estándar igual a 1. La varianza también es igual a 1; porque varianza es igual a la desviación estándar al cuadrado, y 1 al cuadrado es 1. *StandardScaler* hace que la media de la distribución sea 0. Por otro lado, cerca del 68% de los valores estarán entre -1 y 1.

La Figura 4.17 muestra los resultados de la normalización de los datos con el comando anteriormente mencionado de la librería *scikit-learn*. Los valores pertenecientes a las 25 variables monitoreadas tomaron un valor entre -1 y 1. Esta normalización de los datos es importante para implementar los algoritmos de aprendizaje automático, y así obtener los atributos relevantes para el modelo de comportamiento del trastorno de ansiedad generalizada.

Normalización de los datos

HR_Mean	HR_SD	Duration	Efficiency	M_Asleep	M_Awake	Sleep_Deep	Sleep_Light	Sleep_REM	Sleep_stage_Wake	Sleep_stage_Wake_Count	T_Bed
0.37370	-0.81885	-0.00000	0.00000	0.00000	0.00000	0.00000	-0.00000	-0.00000	0.00000	0.00000	0.00000
-0.97103	-0.44841	-1.10725	1.06509	-1.07635	-1.18735	-1.08915	-0.83116	-1.11106	-1.18735	-1.73601	-1.10725
-0.07287	-0.24466	-0.03519	0.41420	-0.01671	-0.19789	0.39889	-0.24249	0.05086	-0.19789	-0.60159	-0.03519
-0.54439	-0.99810	-0.22380	-1.53847	-0.24378	0.00000	0.58967	-0.55522	-0.33645	0.00000	1.10004	-0.22380
0.93330	-0.14832	-0.00000	0.00000	0.00000	0.00000	0.00000	-0.00000	-0.00000	0.00000	0.00000	0.00000
-0.44531	-0.33458	-0.66056	1.39054	-0.64384	-0.69262	0.70413	-1.10710	-0.76679	-0.69262	-0.22345	-0.66056
-1.23232	-1.05458	0.49091	-0.88758	0.52392	0.09895	-0.70760	1.11881	0.26603	0.09895	-0.22345	0.49091
1.82514	2.28677	-1.22637	1.39054	-1.26017	-0.69262	-0.70760	-1.53021	-0.63769	-0.69262	-1.35787	-1.22637
0.22367	-0.74336	2.64494	0.08876	2.53508	3.16626	2.80265	2.05701	2.11650	3.16626	2.61260	2.64494

Figura 4.17. Tabla correspondiente a la normalización de los datos.

4.2.5 Obtención de atributos relevantes

La obtención de atributos relevantes empieza importando los datos resultantes de la fase de procesamiento de los datos. Una vez que los datos estén limpios y normalizados, se procede a aplicar técnicas de aprendizaje automático para la obtención del modelo de comportamiento del nivel de ansiedad. No obstante, primero se deben identificar las variables predictoras (X) y la variable objetivo (y) para obtener aquellos atributos que serán relevantes para la obtención del modelo de comportamiento del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.

4.2.5.1 Identificación de variables predictoras (X) y objetivo (y)

La identificación de las variables predictoras (X) y objetivo (y) es un paso importante para la obtención de los atributos relevantes, y la obtención del modelo de comportamiento correspondiente al trastorno de ansiedad generalizada.

En el caso de esta investigación se identificó como variable objetivo (y) el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada (3 en total): nivel bajo, medio y alto; mientras que las variables predictoras (X) se identificaron como aquellas pertenecientes al perfil sociodemográfico del participante (3 en total): ID, sexo, grado y edad; las variables pertenecientes a la frecuencia cardíaca (2 en total): media y desviación estándar por día de la frecuencia cardíaca; las variables pertenecientes a la actividad física (10 variables): total de

calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados, distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria, minutos en actividad ligeramente activa, minutos en actividad moderada y minutos en actividad muy activa; así como las variables pertenecientes a las etapas del sueño de cada participante (10 variables): duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo del sueño, tiempo transcurrido en cama. En total, son 25 las variables predictoras (X) tomadas en cuenta para la selección de atributos relevantes.

Un ejemplo de las variables predictoras (X) pueden descargarse de la siguiente liga: <https://drive.google.com/file/d/1kd2551-aB3hotkmrZIYVDhiBfiiYEgs4/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: ID del participante, sexo (discretizado: femenino = 0 y masculino = 1), grado (maestría = 0), edad (23-25 años = 0 y 26-30 años = 1), media de la frecuencia cardíaca por día, desviación estándar de la frecuencia cardíaca por día, duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo del sueño, tiempo transcurrido en cama, total de calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados, distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria, minutos en actividad ligeramente activa, minutos en actividad moderada y minutos en actividad muy activa.

La variable objetivo (y) puede descargarse de la siguiente liga: <https://drive.google.com/file/d/1xEgytUb1uwWm0ZLu4fBHE06Q0IVQryPP/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: ID del participante y nivel de ansiedad discretizado: nivel bajo 1, nivel bajo 2 y nivel alto 3.

Cabe aclarar que las variables predictoras (X) pertenecen a la misma tabla que la variable objetivo (y). Los enlaces previamente mencionados muestran a las variables predictoras (X), y variable objetivo (y), separadas para ilustrar la diferencia entre ambos conceptos. También, se aclara que la variable objetivo (y) sumada a la tabla de las variables predictoras (X) sería la variable número 26. No obstante, se debe tener en cuenta que existe una diferencia del tipo de variable que es la variable objetivo (y) y las 25 variables predictoras (X).

Un ejemplo de la tabla combinada con la variable objetivo (y), y las 25 variables predictoras (X) puede descargarse de la siguiente liga: <https://drive.google.com/file/d/1cQfS4YDPBrq7uMfbOI9QNM1nau6ZjDt/view?usp=sharing>. Las variables contenidas, por orden de izquierda a derecha son: ID del participante, sexo (discretizado: femenino = 0 y masculino = 1), grado (maestría = 0), edad (23-25 años = 0 y 26-30 años = 1), media de la frecuencia cardíaca por día, desviación estándar de la frecuencia cardíaca por día, duración del ciclo del sueño, eficiencia del ciclo del sueño, minutos dormidos, minutos despierto, minutos en etapa de sueño profundo del ciclo del sueño, minutos en etapa de sueño ligero del ciclo del sueño, minutos en etapa de sueño REM del ciclo del sueño, minutos despierto durante el ciclo del sueño, número de veces que hubo una interrupción en el ciclo del sueño, tiempo transcurrido en cama, total de calorías quemadas, calorías quemadas cuando el cuerpo está en reposo (bmr: índice metabólico basal), número de pasos dados, pisos escalados, distancia total recorrida, elevación de los pasos, minutos en actividad sedentaria, minutos en actividad ligeramente activa, minutos en actividad moderada, minutos en actividad muy activa, y nivel de ansiedad discretizado: nivel bajo 1, nivel bajo 2 y nivel alto 3.

4.2.5.2 Obtención del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada

Una vez identificadas las variables predictoras (X) y objetivo (y), se procedió a obtener el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada. Para esto, se establecieron rangos previamente recomendados por profesionales en medicina, así como artículos encontrados en la literatura.

De acuerdo con el Centro Médico de la Universidad de Pittsburgh (UPMC, 2015), los parámetros de una frecuencia cardíaca normal se encuentran entre los 60 y

los 80 latidos por minuto. La Asociación Americana del corazón; por otra parte, considera que una frecuencia cardíaca por arriba de los 100 latidos por minutos puede atribuirse a una arritmia, taquicardias inusuales, o crisis de ansiedad (Association, 2015).

Por otro lado, según el Instituto Nacional de Enfermedades Neurológicas y Accidentes Cerebrovasculares (NINDS por sus siglas en inglés), la cantidad de sueño que se necesita depende de varios factores (incluidos la edad, los requerimientos individuales y si se ha estado durmiendo lo suficiente) (NINDS, 2018). En 2015, investigadores afiliados a la Fundación Nacional del Sueño (NFS por sus siglas en inglés) dieron a conocer la duración por edad: los adultos de 18 a 64 años de edad deben dormir entre las 7 y las 9 horas diarias (NIH, 2018). Por otro lado, Kryger, Roth y Dement dieron a conocer, a través de un estudio conducido por la Marina de Estados Unidos, que dormir menos de 6 horas diarias está asociado al trastorno de ansiedad generalizada (Kryger, Roth & Dement, 2017).

Estas variables predictoras (X), pertenecientes a la frecuencia cardíaca y a la duración del ciclo del sueño; y que son descargadas del dispositivo vestible Fitbit Inspire HR, son las más importantes para determinar si una persona está sufriendo del trastorno de ansiedad generalizada; de acuerdo a la Asociación Americana del Corazón, al Instituto Nacional de Enfermedades Neurológicas y Accidentes Cerebrovasculares, así como al Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V) (American Psychiatric Association, 2013).

Una vez se establecieron los rangos previamente aceptados por expertos en medicina y por artículos en la literatura, se procedió a desarrollar un algoritmo para obtener el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada de una manera automática. Este algoritmo está basado en los rangos establecidos para la frecuencia cardíaca por la Asociación Americana del Corazón (Association, 2015), y por el Centro Médico de la Universidad de Pittsburgh (UPMC, 2015). Por otro lado, para la duración del ciclo del sueño, este algoritmo se basa en los rangos establecidos por el Instituto Nacional de Enfermedades Neurológicas y Accidentes Cerebrovasculares (NIH, 2018), y por Kryger, Roth y Dement (Kryger, Roth & Dement, 2017).

Para el caso de la frecuencia cardíaca, los rangos se definieron de la siguiente manera: considerando que la frecuencia cardíaca normal se encuentra entre los 60 y los 80 latidos por minuto, se puede establecer este rango como un nivel bajo de ansiedad correspondiente al trastorno de ansiedad generalizada. Si la frecuencia cardíaca se encuentra por debajo de los 60 latidos por minuto es imposible que se asocie a un nivel de ansiedad correspondiente al trastorno de ansiedad generalizada. En cambio, si la frecuencia cardíaca se encuentra por encima de los 100 latidos por minuto, es probable, de acuerdo a la Asociación Americana del Corazón (Association, 2015) que se esté sufriendo de una crisis de ansiedad, de una arritmia o de taquicardias inusuales; por lo que se podría asociar a un nivel alto de ansiedad correspondiente al trastorno de ansiedad generalizada, dejando el rango de los 80 a 100 latidos por minuto como un nivel medio de ansiedad correspondiente al trastorno de ansiedad generalizada. Por lo que, las condiciones para establecer los niveles de ansiedad correspondientes al trastorno de ansiedad generalizada según a las reglas anteriormente definidas, son las siguientes:

- si frecuencia cardíaca < 60 no hay ningún nivel de ansiedad;
- si frecuencia cardíaca ≥ 60 pero < 80 corresponde al nivel bajo de ansiedad;
- si frecuencia cardíaca ≥ 80 pero < 100 corresponde al nivel medio de ansiedad;
- si frecuencia cardíaca ≥ 100 corresponde al nivel alto de ansiedad.

No obstante, con sólo estas reglas es imposible acotar clínicamente que alguien esté padeciendo de alguna crisis de ansiedad; por lo que es necesario agregar otras reglas para definir los niveles de ansiedad asociados al trastorno de ansiedad generalizada. De esta forma, tomando otra respuesta de déficit cognitivo conductual establecida como importante por el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V) (American Psychiatric Association, 2013), y monitoreada por el dispositivo vestible Fitbit Inspire HR, se agregan las horas dormidas a las reglas anteriormente descritas. En este aspecto, los rangos definidos para las horas dormidas son los siguientes: considerando que las horas de sueño que se deben dormir normalmente se encuentran entre las 7 y 9 horas diarias, se puede establecer este rango como un nivel bajo de ansiedad correspondiente al trastorno de ansiedad generalizada. Si las horas dormidas se encuentran por encima de las 9 horas diarias, es imposible que se asocie a un nivel de ansiedad correspondiente al trastorno de ansiedad

generalizada. En cambio, si las horas dormidas se encuentran por debajo de las 6 horas diarias, de acuerdo con Kryger, Roth y Dement (Kryger, Roth, & Dement, 2017), se puede establecer como un rango para el nivel de ansiedad alto correspondiente al trastorno de ansiedad generalizada. Por lo que, las condiciones para establecer los niveles de ansiedad correspondientes al trastorno de ansiedad generalizada según a las reglas anteriormente definidas, son las siguientes:

- si las horas dormidas ≥ 9 no hay ningún nivel de ansiedad;
- si las horas dormidas son ≥ 7 pero < 9 corresponde al nivel bajo de ansiedad;
- si las horas dormidas son ≥ 6 pero < 7 corresponde al nivel medio de ansiedad;
- si las horas dormidas son ≤ 6 corresponde al nivel alto de ansiedad.

La Figura 4.18 muestra parte del algoritmo con las reglas anteriormente descritas tomando en cuenta a la Asociación Americana del Corazón (Association, 2015), a Kryger, Roth y Dement (Kryger, Roth, & Dement, 2017), al Instituto Nacional de Enfermedades Neurológicas y Accidentes Cerebrovasculares (NIH, 2018), y al Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V) (American Psychiatric Association, 2013).

Algoritmo para obtener el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada

```
si frecuencia cardíaca < 60 y las horas dormidas son  $\geq 9$  no hay ningún nivel de ansiedad  
si frecuencia cardíaca  $\geq 60$  pero  $< 80$  y las horas dormidas son  $\geq 7$  pero  $< 9$  corresponde a nivel bajo  
si frecuencia cardíaca  $\geq 80$  pero  $< 100$  y las horas dormidas son  $\geq 6$  pero  $< 7$  corresponde a nivel medio  
si frecuencia cardíaca  $\geq 100$  y las horas dormidas son  $\leq 6$  el nivel de ansiedad es alto
```

Figura 4.18. Algoritmo para obtener el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.

4.2.5.3 Obtención de atributos relevantes

La selección de atributos, o la obtención de atributos relevantes, es el proceso en el que se seleccionan los atributos que tienen más importancia para predecir la variable objetivo (y), o para contribuir en una mejora en cuanto a la precisión del modelo. Tener atributos irrelevantes en los datos sólo provocará que la precisión del modelo disminuya; además, el modelo estaría siendo entrenado con estos atributos irrelevantes. El objetivo de esta selección es mejorar el rendimiento del modelo para obtener una mejor precisión, reducir el costo computacional, y el

tiempo de entrenamiento para obtener un modelo que sea simple de interpretar; es decir, reducir la complejidad del modelo.

Una vez identificadas las variables predictoras (X), la variable objetivo (y), y obtenido el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada, se procedió a obtener los atributos relevantes utilizando el algoritmo árbol de decisión para determinar la importancia de cada atributo, así como la obtención de las reglas pertenecientes al modelo de comportamiento.

Además, convenientemente, la implementación del algoritmo árbol de decisión con la librería de aprendizaje automático *scikit-learn* perteneciente a Python, ya recopila los valores de importancia de los atributos a través del comando *feature_importances*. Cabe mencionar, que la librería de aprendizaje automático *scikit-learn* en conjunto con el algoritmo árbol de decisión, implementa el comando *SelectFromModel* que selecciona los atributos basándose en el umbral especificado por el usuario después del ajuste del modelo; lo cual resulta conveniente si se quiere utilizar el árbol de decisión como un selector de atributos, y como un paso intermedio para conectar diferentes pasos de pre-procesamiento con un estimador para la evaluación del modelo.

La Figura 4.19 muestra parte del comando utilizado para el caso del algoritmo árbol de decisión; mientras que la Figura 4.20 muestra el resultado de los atributos relevantes seleccionados de mayor a menor por este algoritmo a partir de los 25 atributos para la obtención del modelo de comportamiento correspondiente al trastorno de ansiedad generalizada.

Parte del comando utilizado en la Figura 4.19 utiliza los parámetros para equilibrar a la clase minoritaria (*class_weight*), y para dividir los datos bajo el criterio Gini, así como una validación cruzada con 10 interacciones para balancear los datos pertenecientes a las 25 variables predictoras (X) y a la variable objetivo (y).

Comando utilizado para la obtención de los atributos relevantes

```
clf = tree.DecisionTreeClassifier(class_weight='balanced', criterion='gini', max_depth=7,
                                max_features=25, max_leaf_nodes=None, min_samples_leaf=1,
                                min_samples_split=2, min_weight_fraction_leaf=0.0,
                                presort=False, random_state=None, splitter='best')

cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
n_scores = cross_val_score(clf, X, y, scoring='accuracy', cv=cv, n_jobs=-1, error_score='raise')

clf = clf.fit(X,y)
clf_names = list(X.columns)
```

Figura 4.19. Comando utilizado para la obtención de los atributos relevantes. Se hace una validación cruzada; además de ajustar automáticamente el peso de cada clase con respecto a los datos pertenecientes a las 25 variables predictoras.

Después de implementar el comando perteneciente a la Figura 4.19 se obtuvieron los siguientes 5 atributos relevantes, ordenados de mayor a menor de acuerdo al peso de los atributos y mostrados en la Figura 4.20. Los atributos relevantes son: i) la media por día de la frecuencia cardíaca (HR_Mean); ii) la eficiencia del ciclo del sueño (Efficiency); iii) los minutos dormidos en etapa de *sueño REM* (Sleep_REM); iv) los minutos en actividad ligeramente activa (Light_min); y v) el número de pasos dados (Steps).

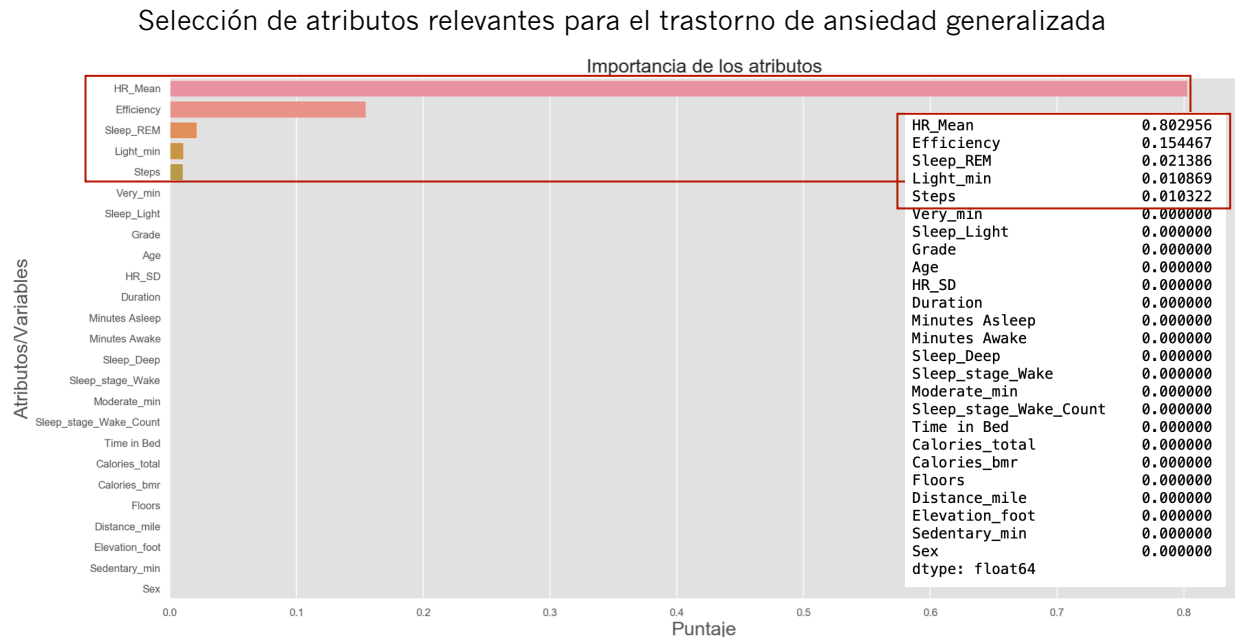


Figura 4.20. Selección de atributos relevantes para el trastorno de ansiedad generalizada. Los atributos están ordenados de mayor a menor. En total, la suma de los valores obtenidos para cada atributo es 1.

De los 25 atributos originales, sólo 5 fueron seleccionados, de los cuales, la mayor puntuación la obtuvo el atributo *media por día de la frecuencia cardíaca*

(*HR_Mean*) con un valor de 0.802956, seguido por la *eficiencia del ciclo del sueño* (*Efficiency*) con un valor de 0.154467, los *minutos dormidos en etapa de sueño REM* (*Sleep_REM*) con un valor de 0.021386, los *minutos en actividad ligeramente activa* (*Light_min*) con un valor de 0.10869, y el de menor puntuación fue el atributo *número de pasos dados* con un valor de 0.010322. Los 20 atributos restantes obtuvieron un valor de 0.0000000; por lo que no fueron tomados en cuenta.

La Figura 4.20 también muestra la lista con el valor obtenido por cada atributo. La suma de los valores obtenidos para cada atributo será 1, de acuerdo con los parámetros establecidos por el clasificador árbol de decisión.

4.2.6 Obtención del modelo de comportamiento

Para la obtención del modelo de comportamiento, se utilizaron diferentes algoritmos de clasificación pertenecientes al aprendizaje automático, con el fin de obtener un modelo con el mejor rendimiento. El proceso para hacer la comparación de desempeño entre los algoritmos de clasificación, se realizó con el programa Jupyter Notebook a los 5 atributos seleccionados como más relevantes. Los algoritmos clasificadores que se utilizaron fueron: *Decision-Tree* (DT), *Gradient-Boosting* (GB), *Adaboost* (AB), *Naive Bayes* (NB), *Support Vector Machine* (SVM), *K-Nearest Neighbor* (k-NN), *Linear-Regression* (LR) y *Gradient-Boosting Regressor* (GBR). Las pruebas de rendimiento se llevaron a cabo con una validación cruzada de 10 iteraciones evaluada 3 veces, obteniendo como resultado un 98.64% de rendimiento para el algoritmo árbol de decisión (*Decision-Tree*), seguido de un 98.54% de rendimiento para los algoritmos *Gradient-Boosting* (GB) y *Adaboost* (AB), el algoritmo *Gradient-Boosting Regressor* (GBR) obtuvo un 94.53% de rendimiento; mientras que el algoritmo *Naive Bayes* (NB) obtuvo un 83.78%. Los algoritmos *Linear-Regression* (LR), *K-Nearest Neighbor* (k-NN) y *Support Vector Machine* (SVM) obtuvieron un rendimiento menor al 75.5%.

La obtención del modelo se realizó al aplicar el algoritmo árbol de decisión a los 5 atributos seleccionados como relevantes, y pertenecientes a los datos recolectados (25 variables predictoras y 1 variable objetivo) por los 5 participantes monitoreados. El comando para la construcción del modelo se presenta en la Figura 4.20. Se recuerda que los árboles de decisión son

representaciones gráficas que dan solución a través de decisiones basadas en ciertas condiciones. Estos árboles tienen un primer nodo, que es la raíz (*root*). Los nodos no terminales del árbol; por otro lado, representan un atributo que tendrá tantas ramas como valores distintos pueda tomar y los nodos terminales representan a las clases. En cuanto a la clasificación de las instancias, se comienza por la raíz y se sigue el camino que indican los valores de los atributos hasta llegar a un nodo terminal, el cual indica la clase a la que pertenece. Al aplicar el algoritmo, se están analizando los datos recolectados en las pruebas de monitoreo para decidir la mejor forma de hacer las divisiones entre nodos. El algoritmo otorga un árbol de decisión óptimo que considere más acertado desde el punto de vista estadístico.

El programa Jupyter Notebook, obtuvo un árbol con el mejor desempeño perteneciente al modelo de comportamiento del trastorno de ansiedad generalizada, implementando el algoritmo árbol de decisión anteriormente mencionado. La librería Scikit-learn hace uso de una versión optimizada del algoritmo CART. Este algoritmo crea un árbol binario, donde cada nodo tiene dos salidas, encontrando la mejor característica numérica para dividir haciendo uso del criterio de impureza. En este aspecto, y para el caso de esta investigación, el criterio elegido fue *gini* para la división de nodos.

Los parámetros obtenidos del árbol de decisión, y mostrados en la Figura 4.21, son: *nombre del atributo*, *gini*, *samples*, *value* y *class*. El parámetro *gini* es la métrica que mide la probabilidad de escoger aleatoriamente un elemento (nivel de ansiedad) para ser correctamente clasificado. El parámetro *samples* son el número de muestras de los niveles de ansiedad que corresponden a esa decisión. El parámetro *value* describe la repartición de las muestras entre las 2 posibles clases correspondientes a los valores que puede tomar el nivel de ansiedad (bajo y medio): el valor de la izquierda pertenece a la primera clase (nivel bajo de ansiedad); mientras que el valor de la derecha pertenece a la segunda clase (nivel medio de ansiedad). El parámetro *class* es el valor que puede tomar el nivel de ansiedad y que es predicho por el árbol de decisión en el nodo raíz. Cabe aclarar que el algoritmo es el que elige qué atributos, y los mejores valores de los parámetros (*gini*, *samples*, *value* y *class*) para la toma de decisiones.

Modelo de comportamiento obtenido para el trastorno de ansiedad generalizada

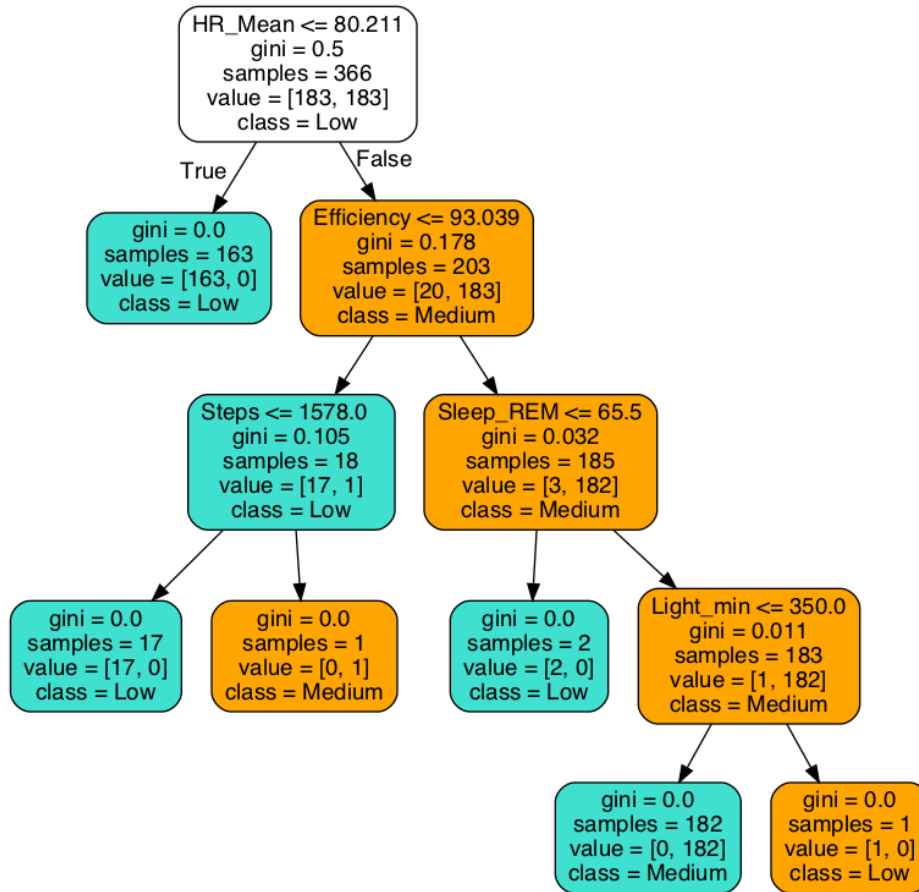


Figura 4.21. Modelo de comportamiento obtenido para el trastorno de ansiedad generalizada.

Matemáticamente hablando, el árbol de decisión basa su comportamiento y la división de los nodos en la fórmula correspondiente a la impureza gini, la función utilizada para evaluar la división de los datos. La impureza gini o el valor gini, es una métrica que mide la probabilidad de escoger un elemento al azar (para el caso de esta investigación un nivel de ansiedad correspondiente al trastorno de ansiedad generalizada) para ser incorrectamente clasificado; es decir, la probabilidad de escoger un elemento multiplicado por la probabilidad de ser mal clasificado. La impureza gini, o el valor gini, está dado por la siguiente fórmula:

$$Gini(D) = 1 - \sum_{i=1}^m P_i^2$$

Fórmula 1. Fórmula para la obtención del valor gini para la división de atributos en un árbol de decisión, donde P_i es la probabilidad de que un elemento particular pertenezca a una clase específica.

El valor de gini proporciona una idea de cómo se dividen los datos por la combinación de las clases en los grupos creados por la división; además de ser una variación del coeficiente gini; el cual es utilizado en economía para medir la dispersión de salud en una población, o la desigualdad de los ingresos dentro de un país. Normalmente, este coeficiente se utiliza para medir cualquier forma de distribución desigual. Una separación perfecta da como resultado una puntuación gini de 0; sin embargo, gini siempre dará como resultado valores entre 0 y 1, donde 0 denota que todos los elementos pertenecen a una clase en específico (o que la división es pura); mientras que 1 denota que los elementos están distribuidos aleatoriamente en varias clases. Si el valor de gini es 0.5, denota que los elementos están distribuidos equitativamente en algunas clases. Este valor considera una división binaria para cada atributo; es decir, para el caso de esta investigación, con tres posibles valores de ansiedad correspondientes al trastorno de ansiedad generalizada {bajo, medio, alto}; mientras que los posibles subconjuntos serían: {bajo, medio, alto}, {bajo, medio}, {bajo, alto}, {medio, alto}, {bajo}, {medio}, {alto} y {}. No obstante, al no presentarse algún episodio ansioso por parte de los participantes, catalogado como un nivel alto (valor discretizado de 3), la división binaria para cada atributo tendría los siguientes valores: {bajo, medio}; mientras que los posibles subconjuntos serían: {bajo, medio}, {bajo}, {medio} y {}.

Los pasos para calcular el valor gini para una división serían los siguientes:

1. calcular el valor gini para los subnodos, utilizando la Fórmula 1;
2. calcular el valor gini para la división utilizando la puntuación ponderada de gini para cada nodo de esa división.

Siguiendo los pasos anteriores y utilizando la Fórmula 2, la división de los subnodos del modelo obtenido y presentado en la 4.22 es la siguiente:

$$1 - p_b^2 - p_m^2$$

Fórmula 2. Fórmula para obtener la división de nodos/ramas perteneciente al modelo de comportamiento del trastorno de ansiedad generalizada. Donde P_b^2 es la probabilidad de obtener un nivel bajo de ansiedad correspondiente al trastorno de ansiedad generalizada; mientras P_m^2 es la probabilidad de obtener un nivel medio de ansiedad correspondiente al trastorno de ansiedad generalizada.

- Valor gini para el nodo raíz (media de la frecuencia cardíaca)

$$Gini_{fc} = 1 - \left(\frac{183}{366}\right)^2 - \left(\frac{183}{366}\right)^2 = 0.5$$

Fórmula 3. Fórmula para obtener la división del nodo raíz perteneciente al atributo frecuencia cardíaca.

Siendo 183 el número de muestras pertenecientes a la clase nivel bajo de ansiedad (valor discretizado de 1), 183 el número de muestras pertenecientes a la clase nivel medio de ansiedad (valor discretizado de 2), y 366 el número total de muestras correspondientes a las dos clases (nivel bajo y medio de ansiedad), se obtiene como resultado un valor gini de 0.5, permitiendo así una nueva división.

- Valor gini para el subnodo eficiencia del ciclo del sueño (*efficiency*)

$$Gini_e = 1 - \left(\frac{20}{203}\right)^2 - \left(\frac{183}{203}\right)^2 = 0.178$$

Fórmula 4. Fórmula para obtener la división del subnodo perteneciente al atributo eficiencia del ciclo del sueño.

Siendo 20 el número de muestras pertenecientes a la clase nivel bajo de ansiedad (valor discretizado de 1), 183 el número de muestras pertenecientes a la clase nivel medio de ansiedad (valor discretizado de 2), y 203 el número total de muestras correspondientes a las dos clases (nivel bajo y medio de ansiedad), se obtiene como resultado un valor gini de 0.178, permitiendo así una nueva división.

- Valor gini para el subnodo número de pasos dados (*steps*)

$$Gini_s = 1 - \left(\frac{17}{18}\right)^2 - \left(\frac{1}{18}\right)^2 = 0.105$$

Fórmula 5. Fórmula para obtener la división del subnodo perteneciente al atributo número de pasos dados.

Siendo 17 el número de muestras pertenecientes a la clase nivel bajo de ansiedad (valor discretizado de 1), 1 el número de muestras pertenecientes a la clase nivel medio de ansiedad (valor discretizado de 2), y 18 el número total de muestras correspondientes a las dos clases (nivel bajo y medio de ansiedad), se obtiene como resultado un valor gini de 0.105, permitiendo así una nueva división.

- Valor gini para el subnodo minutos dormidos en etapa de *sueño REM* (*sleep_REM*)

$$Gini_{REM} = 1 - \left(\frac{3}{185}\right)^2 - \left(\frac{182}{185}\right)^2 = 0.032$$

Fórmula 6. Fórmula para obtener la división del subnodo perteneciente al atributo minutos dormidos en etapa de sueño REM.

Siendo 3 el número de muestras pertenecientes a la clase nivel bajo de ansiedad (valor discretizado de 1), 182 el número de muestras pertenecientes a la clase nivel medio de ansiedad (valor discretizado de 2), y 185 el número total de muestras correspondientes a las dos clases (nivel bajo y medio de ansiedad), se obtiene como resultado un valor gini de 0.032, permitiendo así una nueva división.

- Valor gini para el subnodo minutos en actividad ligeramente activa (*light_min*)

$$Gini_l = 1 - \left(\frac{1}{183}\right)^2 - \left(\frac{182}{183}\right)^2 = 0.011$$

Fórmula 7. Fórmula para obtener la división del subnodo perteneciente al atributo minutos en actividad ligeramente activa.

Siendo 1 el número de muestras pertenecientes a la clase nivel bajo de ansiedad (valor discretizado de 1), 182 el número de muestras pertenecientes a la clase nivel medio de ansiedad (valor discretizado de 2), y 183 el número total de muestras correspondientes a las dos clases (nivel bajo y medio de ansiedad), se obtiene como resultado un valor gini de 0.011, permitiendo así una nueva división; aunque por el bajo valor obtenido, es probable que ya no haya una nueva división después, resultando así en una división perfecta con un valor gini de 0.

El modelo correspondiente al trastorno de ansiedad generalizada (ver Figura 4.21) tras aplicar el algoritmo árbol de decisión con los parámetros brevemente explicados en la Figura 4.19, obtuvo sólo las reglas para un nivel bajo y medio de ansiedad. La razón por la que no se obtuvieron las reglas para un nivel alto de ansiedad reside en que ningún participante sufrió de un episodio ansioso que pudiera servir al algoritmo para su entrenamiento. Cabe aclarar, que entre más datos se incorporen, los modelos predictivos representarán mejor los eventos raros, como en el caso de esta investigación, la incidencia de sufrir una crisis de ansiedad perteneciente al trastorno de ansiedad generalizada.

En este caso, la raíz del árbol, como se esperaba, es la media de la frecuencia cardíaca, seguida de la eficiencia del ciclo del sueño, el número de pasos dados, los minutos dormidos en etapa de *sueño REM*, y los minutos en actividad *ligeramente activa*. El modelo obtenido puede visualizarse como un árbol de decisión en la Figura 4.21, mostrando los 5 atributos seleccionados como relevantes por el algoritmo.

El árbol de decisión obtenido (ver Figura 4.21) muestra el nodo raíz que hace una primera división por frecuencia cardíaca (*HR_Mean*). El valor gini de 0.5 permite la división del nodo, el total de muestras (*Samples*) pertenecientes al nodo raíz es de 366; mientras que la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en partes iguales: 183 para la clase nivel bajo de ansiedad (*Low*), y 183 para la clase nivel medio de ansiedad (*Medium*). La suma de los valores *values* deben ser el total de muestras (*Samples*) pertenecientes al nodo, en este caso de 366. Las salidas del nodo raíz van a izquierda (*True*) con una frecuencia cardíaca menor o

igual a 80 latidos por minutos y a derecha (*False*) con una frecuencia mayor o igual a 80 latidos por minuto.

En el segundo nivel, las muestras (*Samples*) quedan repartidas en 163 para *True* y 203 para *False*. Sin embargo, la salida a *True* obtiene un coeficiente de gini de 0.0, lo que significa que el nodo ya no puede dividirse. Así, el nodo terminal representa a la clase nivel bajo de ansiedad (*Low*) por tener el mayor número de muestras (*Samples*). La clasificación (*Value*) de las muestras (*Samples*) pertenecientes a la salida *True* del nodo raíz, es de 163, número total de muestras que corresponden al nivel bajo de ansiedad (*Low*). Por otro lado, la salida a *False* representa un nodo no terminal del atributo eficiencia del sueño (*Efficiency*) con una salida a *True* si la eficiencia del sueño es menor o igual a 93.039 y una salida a *False* si la eficiencia del sueño es mayor a 93.039. El valor gini de 0.178 permite la división del nodo. El total de muestras (*Samples*) pertenecientes a este nodo es de 203; mientras que la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en 20 para la clase nivel bajo de ansiedad (*Low*), y 183 para la clase nivel medio de ansiedad (*Medium*).

En el tercer nivel las muestras (*Samples*) quedan repartidas en 18 para *True* y 185 para *False*. La salida a *True* representa al nodo no terminal número de pasos dados (*Steps*), con sus respectivas salidas a *True* y *False*, ambas con un valor gini de 0.0. La salida a *True* es el nodo terminal para la clase nivel bajo de ansiedad (*Low*) con un número de muestras (*Samples*) de 17; mientras la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en su totalidad para la clase nivel bajo de ansiedad (*Low*). La salida a *False* es el nodo terminal para la clase nivel medio de ansiedad (*Medium*) con un número de muestras (*Samples*) de 1; mientras la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en su totalidad para la clase nivel medio de ansiedad (*Medium*). Por otro lado, la salida a *False* del atributo eficiencia del sueño (*Efficiency*) representa al nodo no terminal el atributo minutos dormido en etapa de sueño REM (*Sleep_REM*), con una salida a *True* si los minutos dormidos en etapa de sueño REM es menor o igual a 65.5, y una salida a *False* si los minutos dormidos es mayor o igual a 65.5. El valor gini de 0.032 permite la división del nodo. El total de muestras (*Samples*) pertenecientes

a este nodo es de 185; mientras que la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en 3 para la clase nivel bajo de ansiedad (*Low*), y 182 para la clase nivel medio de ansiedad (*Medium*).

En el cuarto nivel las muestras (*Samples*) quedan repartidas en 2 para *True* y 183 para *False*. La salida a *True* representa el nodo terminal para la clase nivel bajo de ansiedad (*Low*) con un número de muestras (*Samples*) de 2; mientras la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en su totalidad para la clase nivel bajo de ansiedad (*Low*). La salida a *False* representa al nodo no terminal del atributo minutos en actividad ligeramente activa (*Light_min*). El valor gini de 0.011 permite la división del nodo. El total de muestras (*Samples*) pertenecientes a este nodo es de 183; mientras que la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en 1 para la clase nivel bajo de ansiedad (*Low*), y 182 para la clase nivel medio de ansiedad (*Medium*).

En el quinto nivel las muestras (*Samples*) quedan repartidas en 182 para *True* y 1 para *False*. La salida a *True* representa al nodo terminal para la clase nivel medio de ansiedad (*Medium*) con un número de muestras (*Samples*) de 182; mientras la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en su totalidad para la clase nivel medio de ansiedad (*Medium*). La salida a *False* es el nodo terminal para la clase nivel bajo de ansiedad (*Low*) con un número de muestras (*Samples*) de 1; mientras la clasificación (*Value*) de las muestras (*Samples*) en las clases pertenecientes a los niveles de ansiedad (*Low* y *Medium*), son repartidas en su totalidad para la clase nivel bajo de ansiedad (*Low*).

Las reglas proporcionadas por el modelo obtenido son explicadas a continuación: si el participante tiene una media de frecuencia cardíaca por arriba de los 80 latidos por minuto, entonces ya es considerado un nivel medio de ansiedad; si en cambio es menor, entonces no hay nivel de ansiedad. En cambio, si el participante tiene una eficiencia del ciclo del sueño menor del 93.039% es considerado como un nivel medio de ansiedad; si por el contrario es mayor al 93.039%, entonces se trata de un nivel bajo de ansiedad. En el caso del número

de pasos dados, si el participante tiene más de 1578 pasos al día, es considerado como un nivel medio de ansiedad; pero si tiene menos de 1578 pasos al día, entonces es considerado como un nivel bajo de ansiedad. Si el participante tiene menos de 65.5 minutos dormidos en etapa de *sueño REM* es considerado un nivel medio de ansiedad, si tiene más de 65.5 minutos dormidos en etapa de *sueño REM* tiene un nivel bajo de ansiedad. Por último, si el participante tiene menos de 350 minutos en actividad *ligeramente activa*, entonces es considerado como un nivel medio de ansiedad, si tiene más de 350 minutos en actividad *ligeramente activa*, ya es considerado como un nivel bajo de ansiedad.

El modelo de comportamiento obtenido del árbol de decisión también puede visualizarse como pseudocódigo (ver Figura 4.22). De esta manera, los cálculos para la construcción del árbol de decisión usando el valor gini son los siguientes:

Pseudocódigo para la construcción del árbol de decisión usando el valor gini

Índice Gini:

```
for cada rama en una división:
  Calcula el porcentaje que la rama representa #Utilizado para ponderar
  for cada clase en una rama:
    Calcula la probabilidad de esa clase en la rama dada
    Elevar al cuadrado la probabilidad de las clases
  Suma las probabilidades de las clases elevadas al cuadrado
  Resta la suma de 1 #Este es el índice Gini para cada rama
  Pondera cada rama en función de la probabilidad de referencia
  Suma el índice Gini ponderado para cada división
```

Figura 4.22. Pseudocódigo para la construcción del árbol de decisión usando el valor Gini.

De acuerdo con el pseudocódigo presentado anteriormente, cada rama tendrá una división en la que se calcula el porcentaje que representa. A su vez, para cada clase perteneciente a esa rama, se calcula la probabilidad de la clase y esta probabilidad se eleva al cuadrado, después se suman las probabilidades de cada clase elevadas al cuadrado y este resultado se le resta a 1. Se pondera cada rama en función de la probabilidad de referencia y se le suma el índice Gini obtenido para cada división.

4.2.6.1 Diagrama de flujo del modelo de comportamiento obtenido correspondiente al trastorno de ansiedad generalizada

El modelo de comportamiento obtenido al aplicar el algoritmo árbol de decisión, y mostrado en la Figura 4.21 puede visualizarse también como un diagrama de flujo en la Figura 4.23.

De acuerdo a las reglas obtenidas para el modelo de comportamiento presentado en la Figura 4.21, se necesita de una media de la frecuencia cardíaca arriba de los 80 latidos por minuto, tener una eficiencia del ciclo del sueño menor al 93.039%, tener más de 1578 pasos al día, haber dormido menos de 65.5 minutos en etapa de *sueño REM*, y tener menos de 350 minutos en actividad *ligeramente activa* para obtener un nivel medio de ansiedad; por el contrario, si la media de la frecuencia cardíaca está por abajo de los 82 latidos por minuto, se tiene una eficiencia del ciclo del sueño mayor al 93.039%, menos de 1578 pasos al día, más de 65.5 minutos dormidos en etapa de *sueño REM* y más de 350 minutos en actividad *ligeramente activa*, es un nivel bajo de ansiedad.

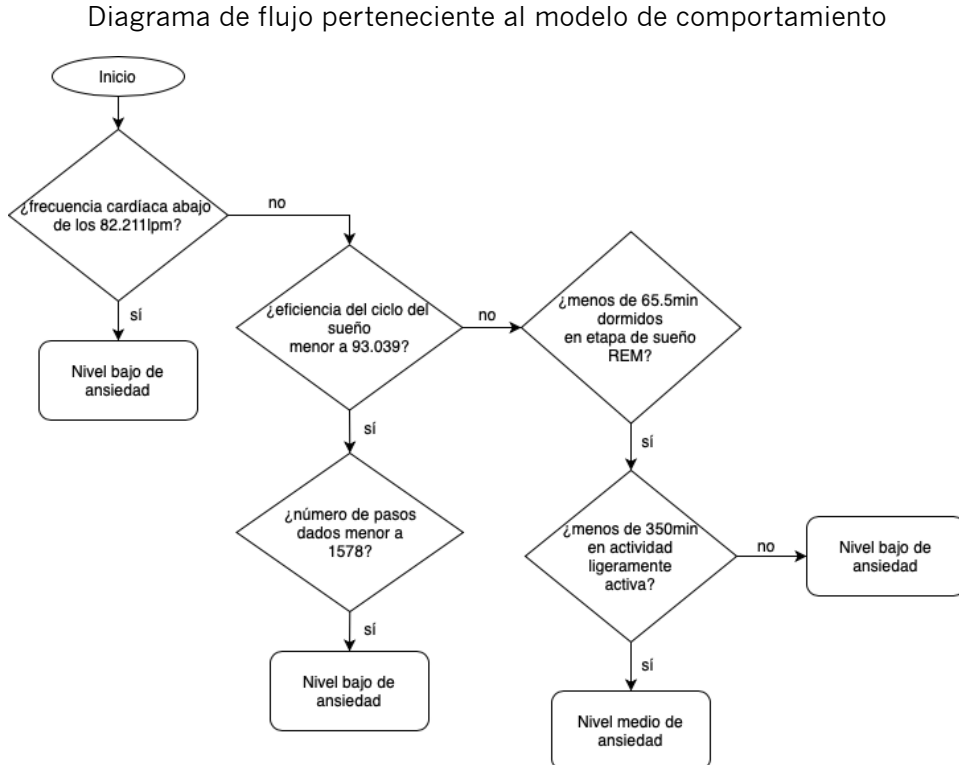


Figura 4.23. Diagrama de flujo perteneciente al modelo de comportamiento del trastorno de ansiedad generalizada.

4.3 Evaluación del modelo obtenido con métricas

El objetivo de la evaluación de un modelo de aprendizaje automático se basa en estimar la precisión de un modelo. Los métodos para evaluar el rendimiento de cualquier modelo requieren utilizar un conjunto de pruebas diferente al conjunto de datos que se utilizaron para construir el modelo. Esto se debe a que el modelo recordará todo el conjunto de entrenamiento; y por lo tanto, siempre predecirá la etiqueta correcta para cualquier punto del conjunto de entrenamiento.

Las métricas de evaluación son requeridas para evaluar el rendimiento del modelo. Estas métricas van desde la clasificación, hasta la regresión, la precisión-cobertura o el valor-F1. En el caso de esta investigación, una vez se obtuvo el modelo de comportamiento general correspondiente al trastorno de ansiedad generalizada con los atributos más relevantes, se procedió a evaluarlo con las métricas de evaluación precisión-cobertura, exactitud, matriz de confusión y el valor-F1.

La Figura 4.24 muestra los resultados del reporte obtenido con la librería *metrics* perteneciente a Python para el modelo correspondiente al trastorno de ansiedad generalizada (modelo con 5 atributos relevantes). Este reporte contiene los valores de exactitud, matriz de confusión, precisión, cobertura y valor-F1.

Los valores encerrados en el recuadro rojo representan las clases de la variable objetivo (y) pertenecientes al nivel bajo de ansiedad (discretizado con el valor 1) y nivel medio de ansiedad (discretizado con el valor 2). En ninguno de los 5 participantes monitoreados con la pulsera de actividad Fitbit se obtuvo un nivel de ansiedad alto (clase 3); por lo que la matriz de confusión sólo muestra los valores predichos de las clases 1 y 2, correspondientes a los niveles bajo y medio del trastorno de ansiedad generalizada respectivamente.

En total se realizaron 74 predicciones para evaluar el modelo obtenido (ver el recuadro verde perteneciente a la Figura 4.24), de las cuales 37 pertenecen a las predicciones obtenidas para la clase 1, correspondiente al nivel de ansiedad bajo y 37 para la clase 2, perteneciente al nivel de ansiedad medio (ver recuadro azul perteneciente a la Figura 4.24). El recuadro rojo (Figura 4.24) representa los dos

valores de ansiedad obtenidos del modelo de comportamiento: nivel bajo (discretizado con el valor 1) y nivel medio (discretizado con el valor 2).

Resultados obtenidos con las métricas de evaluación para el trastorno de ansiedad generalizada

```

Árbol de decisión
Exactitud: 1.0
Matriz de confusión
[[37  0]
 [ 0 37]]

```

	precision	recall	f1-score	support
1	1.00	1.00	1.00	37
2	1.00	1.00	1.00	37
accuracy			1.00	74
macro avg	1.00	1.00	1.00	74
weighted avg	1.00	1.00	1.00	74

Figura 4.24. Resultados obtenidos para el trastorno de ansiedad generalizada.

La Figura 4.24 muestra un valor de exactitud del 1.00 para todo el modelo. Por otro lado, en las dos clases (nivel bajo y nivel medio) la precisión también obtiene un valor de 1.0, al igual que la cobertura y el valor-F1. En la matriz de confusión se obtienen 37 verdaderos positivos, 0 falsos negativos, 0 falsos positivos y 37 verdaderos negativos. Los 37 verdaderos positivos son los valores que el modelo predijo correctamente la clase positiva (nivel de ansiedad bajo); mientras que los 37 verdaderos negativos son los valores que el modelo predijo correctamente la clase negativa (nivel de ansiedad medio).

Capítulo 5

Evaluación del modelo y resultados

En este capítulo se presenta la Fase 3 de la metodología de solución perteneciente a la investigación. La Fase 3 presenta el proceso llevado a cabo para la evaluación del modelo para identificar el nivel de ansiedad correspondiente al trastorno de ansiedad generalizada. La Figura 5.1 presenta de manera ordenada y detallada cada etapa perteneciente a la Fase 3 de la metodología presentada en la Figura 1.2.

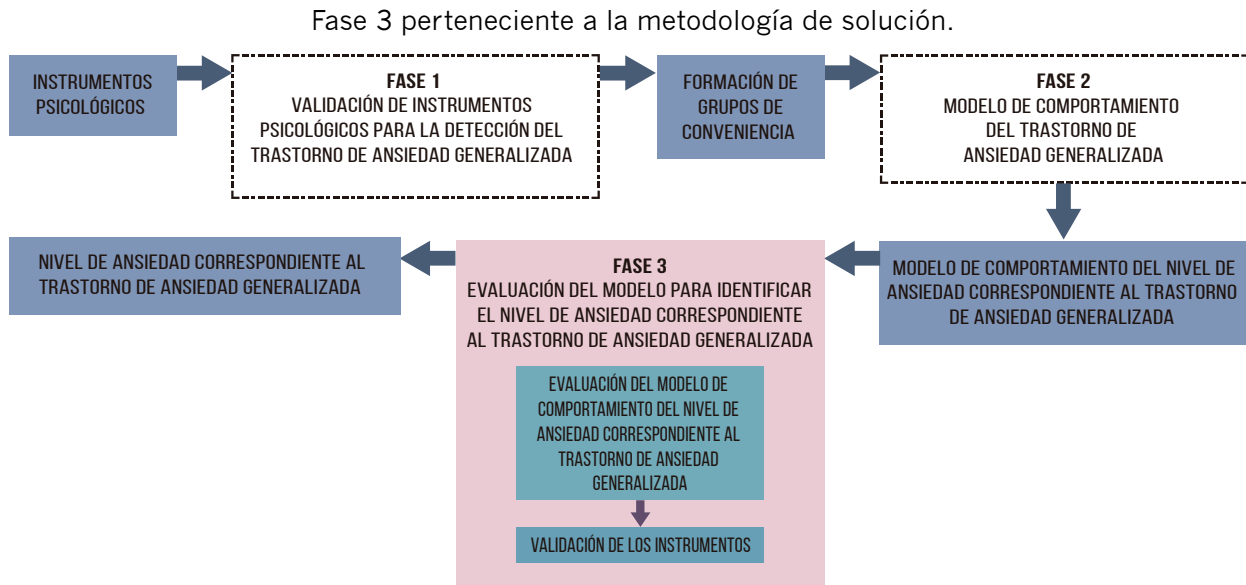


Figura 5.1. Evaluación del modelo de comportamiento del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada

A continuación, se describen brevemente principales las etapas que conforman la Fase 3 correspondiente a la metodología de solución:

- **Etapa 1. Evaluación del modelo de comportamiento del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.** En esta fase se evalúa el modelo de comportamiento obtenido en el capítulo 5, implementando las reglas proporcionadas por el algoritmo de clasificación árbol de decisión para la obtención del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.
- **Etapa 2. Comparación de resultados con escalas de valoración.** En esta fase los resultados de las pruebas de monitoreo son comparados con los resultados previamente obtenidos de la aplicación de los dos instrumentos psicológicos

definidos al inicio de la investigación (cuestionario de preocupación del estado de Pensilvania y la escala de detección del trastorno de ansiedad generalizada).

5.1 Descripción de las pruebas de monitoreo para la evaluación del modelo de comportamiento del trastorno de ansiedad generalizada obtenido

La evaluación del modelo de comportamiento obtenido, y perteneciente al trastorno de ansiedad generalizada requirió la implementación de pruebas de monitoreo a participantes totalmente ajenos a aquellos monitoreados para la obtención del modelo de comportamiento.

5.1.1 Perfil de los participantes

El perfil de los participantes consistió en cumplir con los criterios de inclusión y exclusión; además de ser alumnos pertenecientes al Centro Nacional de Investigación y Desarrollo Tecnológico, ubicado en la ciudad de Cuernavaca, Morelos, México. Los criterios de inclusión fueron los siguientes: i) edad entre 23 años y 30 años; ii) sin distinción de género; y iii) estudiantes del programa de posgrado (maestría y/o doctorado) de Ciencias de la Computación ofrecido por el Centro Nacional de Investigación y Desarrollo Tecnológico. Los criterios de exclusión fueron los siguientes: i) mujeres embarazadas; ii) participantes mayores a 30 años o menores a 23 años; y iii) participantes en tratamiento por trastornos de ansiedad con medicamento recetado.

5.1.2 Selección de los participantes

Para la selección de los participantes se aplicaron las mismas condiciones que las explicadas en el *capítulo 4 Modelo de comportamiento del trastorno de ansiedad generalizada, sección 4.1.2 Formación de grupos de conveniencia*. En este sentido, sólo para las pruebas de evaluación del modelo obtenido para el trastorno de ansiedad generalizada se seleccionaron 3 participantes.

De los 3 participantes seleccionados, 2 fueron hombres con una edad promedio de 23 años; mientras 1 fue mujer con 29 años. Cada participante obtuvo un nivel de ansiedad diferente reportado por el cuestionario de preocupación del estado

de Pensilvania. La participante obtuvo un nivel alto de ansiedad; mientras los otros dos participantes obtuvieron nivel medio y nivel bajo de ansiedad. Por otro lado, 2 de los participantes reportaron presencia sintomatológica reportada por la escala de valoración del trastorno de ansiedad generalizada; mientras que el participante con nivel bajo de ansiedad reportado por el cuestionario de preocupación del estado de Pensilvania no reportó presencia sintomatológica.

5.1.3 Pruebas de monitoreo

Una vez seleccionado a los participantes, las pruebas de monitoreo dieron comienzo al entregar un consentimiento informado para proporcionar los detalles de las pruebas de estudio, así como brindar el dispositivo (Fitbit Inspire HR) que se encargaría de monitorear sólo los 5 atributos seleccionados como relevantes: media de la frecuencia cardíaca, eficiencia del ciclo del sueño, el número de pasos dados, los minutos dormidos en etapa de *sueño REM* y los minutos en actividad *ligeramente activa*.

Las pruebas de monitoreo duraron 16 días por participante, obteniendo así datos pertenecientes a casi 48 días. Cada participante portó el dispositivo las 24 horas de los 16 días que duró su prueba de monitoreo.

Cuando las pruebas de monitoreo finalizaron, se descargaron los datos del dispositivo. Estos datos fueron tratados debidamente en la etapa de procesamiento de los datos para proceder a la evaluación del modelo con métricas de evaluación, y así obtener los valores de rendimiento del modelo obtenido. Después, se procedió a implementar las reglas generadas por el modelo de comportamiento general perteneciente al trastorno de ansiedad generalizada para su evaluación con pruebas de monitoreo.

5.2 Evaluación del modelo de comportamiento del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada

La evaluación del modelo de comportamiento tiene como objetivo implementar las reglas obtenidas por el árbol de decisión generado para el trastorno de ansiedad generalizada con los 5 atributos más relevantes de los 25 tomados en

cuenta para monitorear (ver Figura 4.21). Estos 5 atributos relevantes fueron seleccionados gracias al algoritmo de clasificación árbol de decisión (media de la frecuencia cardíaca, eficiencia del ciclo del sueño, el número de pasos dados, los minutos dormidos en etapa de *sueño REM* y los minutos en actividad *ligeramente activa*).

5.2.1 Implementación de reglas obtenidas del árbol de decisión correspondiente al trastorno de ansiedad generalizada

Las reglas del modelo de comportamiento del trastorno de ansiedad generalizada se obtuvieron gracias al algoritmo árbol de decisión (ver Figura 5.2). Estas reglas involucran los 5 atributos seleccionados como relevantes de acuerdo con el algoritmo árbol de decisión: media de la frecuencia cardíaca, eficiencia del ciclo del sueño, el número de pasos dados, los minutos dormidos en etapa de *sueño REM* y los minutos en actividad *ligeramente activa*.

Las reglas son las siguientes:

Reglas obtenidas del árbol de decisión correspondiente al trastorno de ansiedad generalizada

```
Low = [(df.HR_Mean <= 80.211) & (df.Efficiency >= 93.039) & (df.Steps <= 1578)
      & (df.Sleep_REM >= 65.5) & (df.Light_min >= 350)]
Medium = [(df.HR_Mean >= 80.211) & (df.Efficiency <= 93.039) & (df.Steps >= 1578)
          & (df.Sleep_REM <= 65.5) & (df.Light_min <= 350)]
```

Figura 5.2. Reglas obtenidas del modelo correspondiente al nivel de ansiedad del trastorno de ansiedad generalizada.

Estas reglas obtenidas engloban los dos atributos importantes (frecuencia cardíaca y calidad del sueño) para la detección del trastorno de ansiedad generalizada de acuerdo a los marcos regulatorios como el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V), la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE-10), así como a la Asociación Americana del Corazón y al Instituto Nacional de Enfermedades Neurológicas y Accidentes Cerebrovasculares.

Las reglas obtenidas por el algoritmo árbol de decisión respetan en cierta medida los rangos establecidos inicialmente en la investigación: frecuencia cardíaca ideal entre 60 y 80 latidos por minutos y frecuencia cardíaca atribuida como señal de ansiedad arriba de los 100 latidos por minuto. De esta manera, si el participante

tiene una media de frecuencia cardíaca por arriba de los 80 latidos por minuto, entonces ya es considerado un nivel medio de ansiedad; si en cambio es menor, entonces es un nivel bajo de ansiedad. En cambio, si el participante tiene una eficiencia del ciclo del sueño menor del 93.039% es considerado como un nivel medio de ansiedad; si por el contrario es mayor al 93.039%, entonces se trata de un nivel bajo de ansiedad. En el caso del número de pasos dados, si el participante tiene más de 1578 pasos al día, es considerado como un nivel medio de ansiedad; pero si tiene menos de 1578 pasos al día, entonces es considerado como un nivel bajo de ansiedad. Si el participante tiene menos de 65.5 minutos dormidos en etapa de *sueño REM* es considerado un nivel medio de ansiedad, si tiene más de 65.5 minutos dormidos en etapa de *sueño REM* tiene un nivel bajo de ansiedad. Por último, si el participante tiene menos de 350 minutos en actividad *ligeramente activa*, entonces es considerado como un nivel medio de ansiedad, si tiene más de 350 minutos en actividad *ligeramente activa*, ya es considerado como un nivel bajo de ansiedad.

En términos de pseudocódigo, las reglas anteriormente descritas se pueden observar en la Figura 5.3:

Pseudocódigo propuesto para la obtención del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada

```

If media de la frecuencia cardíaca <= 80.211 latidos por minuto
    nivel ansiedad = "Low"
Else if media de la frecuencia cardíaca >= 80.211 latidos por minuto
    nivel ansiedad = "Medium"
If eficiencia del ciclo del sueño >= 93.039%
    nivel ansiedad = "Low"
Else if eficiencia del ciclo del sueño <= 93.039%
    nivel ansiedad = "Medium"
If número de pasos dados <= 1578
    nivel ansiedad = "Low"
Else if número de pasos dados >= 1578
    nivel ansiedad = "Medium"
If minutos dormidos en etapa de sueño REM >= 65.5
    nivel ansiedad = "Low"
Else if minutos dormidos en etapa de sueño REM <= 65.5
    nivel ansiedad = "Medium"
If minutos en actividad ligeramente activa >= 350
    nivel ansiedad = "Low"
Else if minutos en actividad ligeramente activa <= 350
    nivel ansiedad = "Medium"
    End if
    End if
    End if
    End if
    End if

```

Figura 5.3. Pseudocódigo propuesto para la obtención del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada.

5.2.2 Obtención del nivel de ansiedad con las reglas obtenidas del árbol de decisión correspondiente al trastorno de ansiedad generalizada

Una vez implementadas las reglas obtenidas por el modelo de comportamiento correspondiente al trastorno de ansiedad generalizada se obtienen los niveles de ansiedad de los participantes. De esta manera, con las reglas implementadas, se obtienen 48 valores pertenecientes a los niveles de ansiedad por día,

correspondientes al trastorno de ansiedad generalizada. De estos 48 niveles, 16 pertenecen al participante con nivel bajo de ansiedad, 16 al participante con nivel medio de ansiedad y 16 al participante con nivel alto de ansiedad.

La Tabla 5.1 muestra los registros obtenidos con la implementación de las reglas proporcionadas por el algoritmo árbol de decisión que pertenecen a los participantes por cada nivel de ansiedad correspondiente al trastorno de ansiedad generalizada. En esta Tabla se muestran los 5 atributos seleccionados como los más relevantes: media de la frecuencia cardíaca, eficiencia del ciclo del sueño, el número de pasos dados, los minutos dormidos en etapa de *sueño REM* y los minutos en actividad *ligeramente activa* (ver Glosario de variables).

En la Tabla 5.1 también se puede observar que los niveles de ansiedad correspondientes al trastorno de ansiedad generalizada obtenidos tras la implementación de las reglas pertenecientes al modelo de comportamiento, resultan ser, en su gran mayoría, nivel bajo (Low). Esto puede deberse a que los 3 participantes, a pesar de llevar el dispositivo vestible puesto, no presentaron una cantidad considerable de episodios de ansiedad en sus respectivas pruebas de monitoreo según las reglas anteriormente definidas. Sin embargo, también se puede corroborar observando los valores de la variable eficiencia del ciclo del sueño donde el valor máximo es una eficiencia del 100%; mientras que el valor mínimo es una eficiencia del 84%; y en los valores de la variable de la media de la frecuencia cardíaca donde el valor máximo es de 97 latidos por minuto, mientras el valor mínimo es de 57 latidos por minuto.

Si se toman en cuenta los datos proporcionados por Fitbit, se considera una buena eficiencia si los valores se encuentran en un rango entre el 80% y 90%; por lo que, los registros pertenecientes a la eficiencia del ciclo del sueño se encuentran dentro de los rangos normales. Sólo observando la eficiencia del ciclo del sueño perteneciente a los 3 participantes se puede corroborar, que, en su periodo de monitoreo, en la gran mayoría de los días, estuvieron lo suficientemente relajados como para presentar un cuadro ansioso de nivel alto (valor discretizado de 3); no obstante, hubo 2 ocasiones que se detectó un nivel de ansiedad medio registrado por el mismo participante (ID 1000). Por otra parte, el valor mínimo detectado por los participantes en minutos dormidos en etapa de *sueño REM* es de 22 minutos, mientras el valor máximo es de 108

minutos. Siendo el participante con ID 1000 el que tiene la mayoría de los registros menores a los 82 minutos dormidos en etapa *REM*, caso contrario a los participantes con ID 1001 y 1010. Para el caso del número de pasos dados y los minutos en actividad *ligeramente activa*, el participante con ID 1000 es el que registra menos pasos y menos minutos en actividad *ligeramente activa*, a comparación de los participantes con ID 1001 y 1010, que en ambas variables registran más pasos y más minutos en actividad *ligeramente activa*.

Tabla 5.1. Resultados del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada obtenidos con las reglas proporcionadas por el algoritmo árbol de decisión.

ID	HR_Mean	Efficiency	Sleep_REM	Steps	Light_min	A_Level
1000	82.4371	95.0	82.0	7243.0	192.0	Low
1000	78.2271	88.7692	58.4615	6038.0	171.0	Low
1000	76.6032	89.0	54.0	5781.0	126.0	Low
1000	77.6627	87.0	22.0	5323.0	156.0	Low
1000	83.6441	92.0	41.0	5631.0	227.0	Medium
1000	77.5737	91.0	58.0	2376.0	104.0	Low
1000	78.3646	85.0	46.0	1848.0	106.0	Low
1000	77.4877	84.0	84.0	6139.0	159.0	Low
1000	76.9038	88.0	48.0	6827.0	189.0	Low
1000	76.1916	88.0	59.0	6799.0	168.0	Low
1000	74.3091	84.0	58.0	4599.0	156.0	Low
1000	75.0927	89.0	83.0	6028.0	151.0	Low
1000	72.5429	90.0	77.0	994.0	78.0	Low
1000	75.2267	88.7692	58.4615	5587.0	142.0	Low
1000	70.8928	92.0	48.0	4378.0	59.0	Low
1000	87.2338	88.7692	58.4615	10833.0	276.0	Medium
1001	85.0633	98.0	63.0	3279.0	165.0	Low
1001	83.4023	98.0	44.0	7484.0	239.0	Low
1001	82.5042	96.0	105.0	2936.0	151.0	Low

ID	HR_Mean	Efficiency	Sleep_REM	Steps	Light_min	A_Level
1001	83.1084	96.0	60.0	6613.0	235.0	Low
1001	81.5380	94.0	87.0	10151.0	261.0	Low
1001	83.8478	95.0	102.0	11513.0	289.0	Low
1001	82.9375	96.125	69.5	7412.4667	217.4667	Low
1001	80.3491	96.125	69.5	14802.0	356.0	Low
1001	87.3688	96.125	69.5	13829.0	262.0	Low
1001	84.0770	96.125	69.5	3356.0	153.0	Low
1001	83.5677	96.125	69.5	2231.0	150.0	Low
1001	81.3167	96.125	69.5	1872.0	98.0	Low
1001	79.2009	96.125	69.5	6785.0	264.0	Low
1001	79.0889	96.125	69.5	10632.0	224.0	Low
1001	84.3798	96.125	69.5	7412.4667	217.4667	Low
1001	82.9914	96.125	69.5	7412.4667	217.4667	Low
1001	77.0869	96.0	55.0	13296.0	252.0	Low
1001	71.9345	99.0	64.0	4582.0	129.0	Low
1010	69.4444	98.0	77.0	1126.0	56.0	Low
1010	84.5622	98.0	72.0	29016.0	304.0	Low
1010	72.0295	98.1538	76.69230	9746.0	102.0	Low
1010	72.8667	99.0	73.0	17505.0	141.0	Low
1010	75.8680	98.0	94.0	18413.0	143.0	Low
1010	67.1340	98.0	87.0	10358.4	170.4	Low
1010	86.8020	98.1538	76.6923	22108.0	334.0	Low
1010	57.9580	99.0	58.0	8908.0	199.0	Low
1010	75.1616	98.1538	76.6923	8981.0	181.0	Low
1010	68.2039	99.0	91.0	4376.0	152.0	Low
1010	64.3356	100.0	108.0	2419.0	73.0	Low
1010	63.1834	97.0	67.0	1559.0	53.0	Low

ID	HR_Mean	Efficiency	Sleep_REM	Steps	Light_min	A_Level
1010	67.2749	97.0	69.0	5951.0	171.0	Low
1010	71.6123	98.0	82.0	7390.0	266.0	Low

5.3 Comparación de resultados con instrumento psicológico

Una vez obtenidos los resultados diarios de las pruebas de evaluación del modelo de comportamiento, se procedió a compararlos con los resultados obtenidos del cuestionario de preocupación del estado de Pensilvania. Este proceso resulta fundamental para evaluar el rendimiento del modelo de comportamiento correspondiente al trastorno de ansiedad generalizada.

Los valores para los niveles de ansiedad correspondientes al trastorno de ansiedad generalizada y utilizados en esta investigación son los siguientes: nivel de ansiedad bajo: discretizado con el número 1; nivel de ansiedad medio: discretizado con el número 2; y nivel de ansiedad alto: discretizado con el número 3. Sin embargo, ningún participante sufrió de un episodio ansioso de nivel alto que pudiera ser registrado por las reglas establecidas pertenecientes al trastorno de ansiedad generalizada, caso contrario al nivel medio, en el que un participante registró en 2 ocasiones un nivel medio de ansiedad correspondiente al trastorno de ansiedad generalizada.

Los resultados obtenidos de la comparación por día entre el cuestionario de preocupación del estado de Pensilvania, y los reportados por el modelo de comportamiento perteneciente al trastorno de ansiedad generalizada, se pueden observar en la Tabla 5.2. Estos resultados pertenecen a los 16 días monitoreados de los 3 participantes seleccionados (1000, 1001 y 1010).

Tabla 5.2. Resultados obtenidos de la comparación por día entre el cuestionario de preocupación del estado de Pensilvania y los reportados por el modelo de comportamiento perteneciente a los 3 participantes. Los valores están discretizados, nivel bajo:1, nivel medio: 2 y nivel alto: 3.

ID	Nivel de ansiedad bajo reportado por las reglas proporcionadas por el árbol de decisión	Resultado obtenido por el cuestionario de preocupación de Pensilvania
1000	1	1

ID	Nivel de ansiedad bajo reportado por las reglas proporcionadas por el árbol de decisión	Resultado obtenido por el cuestionario de preocupación de Pensilvania
1000	1	1
1000	1	1
1000	1	1
1000	2	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	1	1
1000	2	1
1000	1	1
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2
1001	1	2

comportamiento correspondiente al trastorno de ansiedad generalizada tras 16 días de monitoreo. En este caso, no es común que un participante con un nivel bajo de ansiedad registre niveles medios o altos de ansiedad; sin embargo, se pueden presentar estos niveles de ansiedad debidos a ciertos contextos o situaciones; ya sean sentimentales, sociales, laborales, académicos, económicos, o debido a algún factor que propicie o estimule un cuadro ansioso. No obstante, si estos cuadros ansiosos se vuelven más recurrentes, es cuando se detecta un trastorno de ansiedad, si estos cuadros ansiosos se presentan con recurrencia por un mínimo de 6 meses, es cuando se habla de un trastorno de ansiedad generalizada.

A pesar de que las métricas de evaluación mostraron buenos resultados en cuanto al rendimiento del modelo de comportamiento, los resultados obtenidos por las pruebas de evaluación al mismo modelo no fueron los esperados. Se observa que el modelo de comportamiento determinó 2 niveles de ansiedad medios a un participante que obtuvo un nivel bajo de ansiedad en el cuestionario de preocupación del estado de Pensilvania. Contrario a los otros dos participantes que obtuvieron un nivel medio y alto en el instrumento psicológico anteriormente mencionado, y el modelo de comportamiento determinó un nivel bajo de ansiedad.

En el caso del participante con ID 1000, el modelo de comportamiento determinó, de los 16 días monitoreados, un nivel de ansiedad bajo para 14 días; mientras que los 2 días restantes determinó un nivel medio de ansiedad. Para los participantes con ID 1001 y 1010, el modelo de comportamiento determinó para los 16 días, un nivel bajo de ansiedad. En este aspecto, es posible que los participantes con niveles medio y alto de ansiedad, reportados por el cuestionario de preocupación del estado de Pensilvania, puedan obtener tanto niveles bajo como niveles medio de ansiedad determinados por el modelo de comportamiento. No obstante, no es común que un participante con nivel bajo de ansiedad registre niveles medio de ansiedad; lo que lleva a concluir que se necesitan más pruebas de monitoreo para la construcción del modelo de comportamiento, así como para su evaluación.

En cuanto a la precisión, se puede decir que el modelo de comportamiento tiene una precisión del 75-85%; porque no está determinando bien el nivel de ansiedad

correspondiente al trastorno de ansiedad generalizada, ya que, para el participante con nivel bajo de ansiedad está determinando 2 niveles diferentes a los reportados por el instrumento psicológico; mientras que para los participantes de nivel medio y alto de ansiedad, reportados por el cuestionario de preocupación del estado de Pensilvania, no está determinando niveles medio de ansiedad. Cabe mencionar, que los resultados obtenidos por el modelo de comportamiento, permiten tener una idea del nivel de ansiedad correspondiente al trastorno de ansiedad generalizada que pudiera estar presentando algún participante. Sin embargo, no se puede asegurar completamente que sea el nivel de ansiedad que está sintiendo realmente.

Por otro lado, gracias a los 16 días definidos como periodo de monitoreo, se llegó a la conclusión de que es posible detectar el trastorno de ansiedad generalizada a través de dispositivos vestibles. No obstante, esta detección es complicada en el sentido que hay días en que las personas no presentan crisis de ansiedad; es decir, están lo suficientemente relajados y no hay algún factor que les induzca posibles cuadros ansiosos. En este sentido es que se aconseja un periodo de monitoreo mayor a los 16 días, o bien, una inducción de ansiedad bajo un ambiente controlado para poder detectar niveles altos de ansiedad correspondientes al trastorno de ansiedad generalizada.

A pesar de haber monitoreado a 8 participantes, 5 de ellos para la obtención del modelo de comportamiento y 3 de ellos para la evaluación del modelo, no se obtuvieron registros pertenecientes a niveles altos de ansiedad. En el caso de la obtención del modelo de comportamiento general del trastorno de ansiedad generalizada, el valor máximo obtenido, en contadas ocasiones, fue el nivel bajo y medio de ansiedad. No obstante, en el caso de la evaluación del modelo de comportamiento del trastorno de ansiedad generalizada, sólo se obtuvieron niveles bajos de ansiedad. Esto tal vez se deba a que, el periodo en el que se realizaron las pruebas de monitoreo para la evaluación del modelo fue en los meses de febrero y marzo de 2020, meses que, según el calendario escolar del posgrado de Ciencias Computacionales del CENIDET, no tenían una carga alta de actividades que ayudaran a propiciar posibles cuadros ansiosos. Es por esto que se recomienda inducir la ansiedad bajo pruebas de monitoreo en un ambiente controlado; ya que el modelo de comportamiento necesita de más datos

pertenecientes a niveles altos de ansiedad para mejorar en su rendimiento y detección.

Para el aspecto de la evaluación del modelo de comportamiento general del trastorno de ansiedad generalizada obtenido, se realizaron varias pruebas de predicción, asegurándose que el modelo de comportamiento trabajara de manera correcta. Sin embargo, es posible que los valores obtenidos de las métricas de evaluación se hayan visto sesgados por la cantidad mínima de datos recolectados. Asimismo, las pruebas que se realizaron para la obtención de atributos relevantes demostraron que diversas variables monitoreadas, y pertenecientes a los datos almacenados del dispositivo Fitbit, no fueron lo suficientemente relevantes para tomarse en cuenta para el modelo de comportamiento del trastorno de ansiedad generalizada, contrario a lo que se pensaba inicialmente.

Capítulo 6

Conclusiones y trabajos futuros

Este capítulo presenta las conclusiones obtenidas a lo largo del desarrollo de la investigación, así como las recomendaciones que se pueden derivar a trabajos futuros de esta investigación.

6.1 Conclusiones

Al iniciar la investigación se detectó la necesidad de conocer, desde la perspectiva psicológica, los trastornos de ansiedad, en específico, el trastorno de ansiedad generalizada, así como la comprensión de los dos marcos regulatorios establecidos en la investigación: Manual de Diagnóstico y Estadístico de Trastornos Mentales (DSM-V) (American Psychiatric Association, 2013), y la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE-10) (Organización Panamericana de la Salud, 2015). También se detectó la necesidad de conocer y comprender cómo se presenta la ansiedad biológicamente en el cuerpo, desde la perspectiva de la medicina. Esto ayudó a cumplir el objetivo general y a plantear el desarrollo de la investigación, así como definir las variables y el dispositivo vestible a ocupar para las pruebas de monitoreo a adultos jóvenes.

Para alcanzar el objetivo general de la investigación se realizaron las siguientes actividades: los instrumentos psicológicos tuvieron que ser validados para seleccionar a los participantes que formarían parte de las pruebas piloto de monitoreo, pruebas que otorgaron los datos para realizar el análisis exploratorio de los datos, la obtención de los atributos más relevantes (respuestas fisiológicas), y la obtención del modelo de comportamiento para adultos jóvenes correspondiente al trastorno de ansiedad generalizada.

La validación de los instrumentos psicológicos definidos en la investigación (cuestionario de preocupación del estado de Pensilvania y la escala de detección del trastorno de ansiedad generalizada) se llevó a cabo para cumplir el primer objetivo específico. En esta validación, se observó que la mayoría de la población encuestada no reportaba los niveles de ansiedad correspondientes al trastorno de ansiedad generalizada que se esperaban; es decir, casi la gran mayoría reportó niveles bajo y medio de ansiedad, cuando se esperaba que más de 10 reportaran un nivel alto de ansiedad. Sin embargo, sólo 4 participantes de los 34

encuestados reportaron un nivel alto de ansiedad; de los cuales se eliminó a un participante por no cumplir con los criterios de inclusión y exclusión definidos en la investigación.

Para alcanzar el segundo objetivo específico se realizaron las pruebas piloto de monitoreo, el análisis exploratorio de los datos y la obtención de los atributos más relevantes. En este aspecto, en las pruebas piloto de monitoreo se detectó una gran pérdida de datos. Inicialmente se había establecido un periodo de monitoreo de 5 días; sin embargo, al descargar los datos del dispositivo vestible se obtuvieron datos de sólo 2 días. Esto modificó la duración de las pruebas piloto, proponiendo un total de 15 días para monitorear respuestas fisiológicas relacionadas al trastorno de ansiedad generalizada. La primera prueba piloto fue descartada y se realizó nuevamente.

Por otra parte, en el análisis exploratorio de los datos se detectó una pérdida de datos correspondiente a las variables pertenecientes a la calidad del sueño. Razón por la que el periodo de monitoreo fue modificado para las pruebas de monitoreo pertenecientes a la evaluación del modelo, de los 15 días propuestos se incrementó uno; es decir, se establecieron 16 días para las pruebas de monitoreo de la evaluación del modelo de comportamiento.

En el desarrollo del modelo de comportamiento se obtuvieron sólo 5 atributos relevantes; lo que llevó a concluir que se necesitan más participantes con diferentes niveles de ansiedad para la recolección de datos. Además, al haber obtenido sólo dos niveles de ansiedad: bajo y medio correspondiente al trastorno de ansiedad generalizada, se llegó a la conclusión de que se requieren, por lo menos 3 participantes más, para la recolección de datos de un nivel alto de ansiedad. Con el incremento de datos recolectados de los participantes, el modelo de comportamiento mejorará sus resultados al aplicar el algoritmo árbol de decisión.

Para alcanzar el tercer y cuarto objetivo específico se realizó la evaluación del modelo de comportamiento obtenido, tanto con métricas de evaluación como con pruebas de monitoreo. La evaluación del rendimiento del modelo de comportamiento obtenido fue mejor de la estimada; ya que de los 74 valores obtenidos en la matriz de confusión, 37 fueron para el nivel de ansiedad bajo y

37 para el nivel de ansiedad medio, resultando en una precisión, exactitud, cobertura y valor F-1 de 1.0. No obstante, en la evaluación del modelo de comportamiento con las pruebas de monitoreo, se observó que se requieren más dispositivos vestibles a utilizar al mismo tiempo, ya que el proceso de estas pruebas ralentizó la investigación al utilizar sólo un dispositivo. Además de corroborar el incremento de participantes para las pruebas de monitoreo. Sin embargo, los resultados obtenidos tras la aplicación de las reglas definidas por el modelo de comportamiento fueron los esperados una vez se compararon con los resultados obtenidos por el cuestionario de preocupación del estado de Pensilvania.

Por último, desde la comprensión del tema de investigación, así como en el transcurso de su desarrollo, se recurrió a diversos expertos de diferentes áreas. Para la comprensión del tema, así como para la selección y validación de instrumentos psicológicos, se recurrió a especialistas en salud mental, psicología y psiquiatría. Para la selección de variables fisiológicas, se tomó como referencia los marcos regulatorios Manual de Diagnóstico y Estadístico de Trastornos Mentales (DSM-V) (American Psychiatric Association, 2013), y la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (CIE-10) (Organización Panamericana de la Salud, 2015). Sin embargo, también se recurrió a profesionales médicos para la comprensión y selección de las respuestas fisiológicas desde la perspectiva de la medicina. Para el desarrollo del modelo de comportamiento, la selección de algoritmos de aprendizaje automático y la evaluación del modelo de comportamiento se recurrió a profesionales en el área de reconocimiento de patrones y aprendizaje automático.

Esta tesis es el resultado de una investigación multidisciplinaria que propone una solución a un tema relacionado a la salud mental, y que puede ser referencia para otras investigaciones que involucren temas relacionados a los trastornos mentales.

6.2 Logros obtenidos

Los logros obtenidos a partir del desarrollo de esta investigación son listados a continuación:

1. Conferencia del artículo *Detección automática de ansiedad: Una revisión sistemática del trastorno de ansiedad generalizada*. Presentado en la Segunda Jornada de Ciencia y Tecnología Aplicada, en Cuernavaca, Morelos. Abril de 2019.
2. Artículo de revista: Cuevas Chávez P. Alejandra, Martínez R. Alicia, Sánchez G. A. Fabiola (2019), *Detección automática de ansiedad: Una revisión sistemática del trastorno de ansiedad generalizada*, *Revista de Ciencia y Tecnología Aplicada*, Vol. 2, pp. 45-49.
3. Presentación del póster *Detección automática del trastorno de ansiedad generalizada*. Presentado en el Congreso Internacional de Investigación Transdisciplinaria en Ciencias Humanas “Salud, Sociedad y Transdisciplina, en Cuernavaca, Morelos. Mayo 2019.
4. Estancia en la Universidad de Hokkaido, Japón, de septiembre de 2019 a enero de 2020. En esta estancia se desarrolló la fase 2 de la metodología de solución: Modelo de comportamiento del trastorno de ansiedad generalizada.

6.3 Trabajos futuros

Los trabajos futuros que se proponen para ampliar y/o mejorar esta investigación se listan a continuación:

- Realizar pruebas de monitoreo controladas buscando inducir la ansiedad para obtener mejores resultados, o bien, controlar aún más las pruebas de monitoreo para detectar con más facilidad los niveles de ansiedad.
- Trabajar en conjunto con organismos dedicados a la psicología clínica y a la medicina para ofrecer al usuario consejos acerca del trastorno de ansiedad generalizada, así como consejos de cómo hacer frente a las crisis de ansiedad.
- Utilizar más de un dispositivo vestible para agilizar el proceso de monitoreo y la obtención de datos de los participantes. Se recomienda monitorear a 10 participantes al mismo tiempo (Cruz, Rubin & Abreu, 2015).
- Implementar el modelo de comportamiento obtenido en una aplicación web o aplicación móvil para detectar patrones de comportamiento de los

usuarios en tiempo real, sin necesidad de descargar los datos de los servidores de Fitbit para su procesamiento y obtención de niveles de ansiedad correspondientes al trastorno de ansiedad generalizada.

- Enriquecer el modelo de comportamiento para la determinación de los niveles de ansiedad correspondiente al trastorno de ansiedad generalizada, agregando más participantes, así como aumentando los días de monitoreo. Para mejores resultados se propone aumentar a un mes de monitoreo y buscar mínimo a 60 participantes con diferentes niveles de ansiedad reportados por el cuestionario de preocupación del estado de Pensilvania.
- Trabajar en conjunto con expertos en áreas de salud especializados en investigación clínica para mejorar la precisión del modelo con el objetivo de desarrollar una estrategia en conjunto para la búsqueda de participantes con trastorno de ansiedad generalizada. De esta manera, se puede ampliar el número de participantes con niveles altos de ansiedad reportados por los instrumentos psicológicos definidos en la investigación (cuestionario de preocupación del estado de Pensilvania y escala de detección del trastorno de ansiedad generalizada), aumentando así la cantidad de datos recolectados de los dispositivos vestibles.
- Husain, Hui Yng & Abdul (Husain, Hui Yng & Abdul, 2016) reportaron buenos resultados con diferentes algoritmos para la predicción del trastorno de ansiedad generalizada, por lo que se propone trabajar con algoritmos de aprendizaje automático diferentes al utilizado en esta investigación (árbol de decisión) para obtener mejores resultados en la obtención de atributos relevantes y del modelo de comportamiento.

Referencias

- Alghamdi, N. S. (2019). Monitoring Mental Health Using Smart Devices with Text Analytical Tool. *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*.
- American Psychiatric Association. (2013). Guía de consulta de los criterios diagnósticos del DSM-V. Arlington, VA: Asociación Americana de Psiquiatría.
- American Psychiatric Association. (2005). *Let's Talk Facts About Anxiety Disorders*. Recuperado el 16 de abril de 2019, de University of South Florida Health: <http://www.fcphp.usf.edu/courses/content/rfast/Resources/anxietydisorders.pdf>
- Arnrich, B., Osmani, V. & Bardram, Jakob. (2013). Mental health and the impact of ubiquitous technologies. *Personal and Ubiquitous Computing*, 17, pp. 211-213.
- Association, A. H. (31 de julio de 2015). *All About Heart Rate (Pulse)*. Recuperado en mayo de 2020, de American Heart Association: <https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure/all-about-heart-rate-pulse>
- Bados, A. (20 de septiembre de 2017). Trastorno de Ansiedad Generalizada. Barcelona, España: Universitat de Barcelona, Facultat de Psicologia.
- Baezas, J., Balaguer, G., Belchi, I., Coronas, M., & Guillamon, N. (2008). *Higiene y prevención de la ansiedad*. España: Díaz de Santos.
- Benedetto, S., Caldato, C., Bazzan, E., Greenwood, D.C., Pensabene, V. & Actis, P. (2018). Assessment of the Fitbit Charge 2 for monitoring heart rate. *PloS one*, 13, 2, e0192691.
- Bobes, J., García-Calvo, C., Prieto, R., García-García, M., & Rico-Villademoros, F. (2006). Psychometric Properties of the Spanish Version of the Screening Scale for DSM-IV Generalized Anxiety Disorder of Carroll and Davidson. *Actas Españolas de Psiquiatría*, 34, 2, pp. 83-93.
- Cabot, S. & Wilkinson, B. (2016). Using Mobile-Based Game as a Means for the Self-treatment of Depression and Anxiety in Youth. *Marsh, T., Ma, M., Baalsrud, Hauge J., Göbel, S., Serious Games, (JCSG 2016), Lecture Notes, Computer Science, 9894*, pp. 128-133.

- Cárdenas, E., Feria, M., Palacios, L. & De la Peña, F. (2010). *Guía clínica para los trastornos de ansiedad en niños y adolescentes*. Recuperado el 10 de agosto de 2020, de Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz México. http://inprf-cd.gob.mx/guiasclinicas/trastornos_de_ansiedad.pdf
- Carlson, N. R. (2010). *Fundamentos de fisiología de la conducta*. (M. J. Platón, Trad.) Madrid, España: Pearson Educación S. A.
- Carroll, B., & Davidson, J. (2000). Screening Scale for DSM-IV GAD. *Copyright*.
- Chaitanya, M. N., Jayakkumar, S., Chong E. & Yeow C. H. (2017). A wearable, EEG-based massage headband for anxiety alleviation. *2017 39th Annual International Conference for the IEEE Engineering in Medicine and Biology Society (EMBC)*.
- Chollet, M., Massachi, T. & Scherer, S. (2017). Racing Heart and Sweaty Palms. *Beskow, J., Peters, C., Castellano, G., O'Sullivan, C., Leite, I., Kopp, S, Intelligent Virtual Agents, (IVA 2017), Lecture Notes, Computer Science, 10498, pp. 83-86.*
- Clark, D., & Beck, A. T. (2012). *Terapia Cognitiva para Trastornos de Ansiedad*. Bilbao: Decleé de Brouwer S. A.
- Cruz, L., Rubin, J., Abreu, R., Ahern, S., Eldardiry, H. & Bobrow, D. G. (2015). A wearable and mobile intervention delivery system for individuals with panic disorder. *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (MUM '15)*, pp. 175-182.
- Dávila Hernández, F., & Sánchez Corales, Y. (2012). Técnicas de minería de datos aplicadas al diagnóstico de entidades clínicas. *Revista Cubana de Informática Médica, 4, 2, pp. 174-183.*
- de Vasconcelos, E., Neto, A., dos Santos, L. & Ribeiro, P. (2019). Analysis of Biofeedback Through Heartbeat Obtained by Exposure to Phobia Through Virtual Reality. *Antona, M., Stephanidis, C., Universal Access in Human-Computer Interaction, Theory, Methods and Tools (HCI 2019), Lecture Notes, Computer Science, 11572, pp. 545-554.*
- Empatica. (s.f.). *Advanced research on human behavior*. Recuperado el 18 de abril de 2019, de Empatica: <https://www.empatica.com/research/science/>
- Elgendi, M., & Menon, C. (2019). Assessing Anxiety Disorders Using Wearable Devices: Challenges and Future Directions. *Brain Sciences, 9, 50.*
- Elsevier. (22 de agosto de 2018). *¿Qué es el 'machine learning'?* Recuperado el 16 de abril de 2019, de Elsevier: <https://www.elsevier.com/es-es/connect/ehealth/que-es-el-machine-learning-salud-digital>

- European Heart Journal. (1998). *Heart rate variability and heart rate in healthy volunteers*. European Heart Journal, pp. 1334-1341.
- Fitbit. (2019). Fitbit Inspire & Inspire HR. Recuperado el día 10 de agosto de 2020, de Fitbit: <https://www.fitbit.com/mx/inspire>
- Gaggioli, A., Pioggia, G., Tartarisco, G., Baldus, G., Corda, D., Cipresso, P. & Riva, G. A mobile data Collection platform for mental health research. *Personal and Ubiquitous Computing*, 17, pp. 241-251.
- Github. Orcasgit/python-fitbit. Recuperado en septiembre de 2019 de 2020, de Github: https://github.com/orcasgit/python-fitbit?source=post_page-----
- González, M., Ibáñez, I., García, L., & Quintero, V. (2018). El "Cuestionario de Preocupación y Ansiedad" como Instrumento de Cribado para el Trastorno de Ansiedad Generalizada: Propiedades Diagnósticas. *Behavioral Psychology*, 26, No. 1, pp. 5-22.
- Gravenhorst, F., Muaremi, A., Bardram, J., Grünerbl, A., Mayora, O., Wurzer, M. F., Osmani, V., Arnrich, B., Lukowicz, P. & Tröster, G. (2015). Mobile phones as medical devices in mental disorder treatment: an overview. *Personal and Ubiquitous Computing*, 19, pp. 335-353.
- Gray, J., & McNaughton, N. (2000). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System* (Segunda Edición ed.). New York: Oxford University Press.
- Gu, J., Gao, B., Chen, Y., Jiang, L., Gao, Z., Ma X., Ma Y. & Woo, W. L. (2017). Wearable Social Sensing and Its Application in Anxiety Assesment. *2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*.
- Gu, G. & Frasson, C. (2017). Virtual Sophrologist: A Virtual Reality Neurofeedback Relaxation Training System. *Frasson, C., Kostopoulos, G., Brain Function Assessment in Learning (BFAL 2017), Lecture Notes, Computer Science, 10512*, pp. 176-185.
- Guerra J., Gutiérrez M., Zavala, M., Singre, J., Goosdenovich, C. & Romero, E. (2017). Relación entre ansiedad y ejercicio físico. *Revista Cubana de Investigaciones Biomédicas*, 36, 2, pp. 169-177.
- Handouzi, W., Maaoui, C., Pruski, A. & Moussaoui, A. (2014). Short-term anxiety recognition from blood volume pulse signal. *IEEE 11th International Multi-Conference on Systems, Signals & Devices (SSD14)*, pp. 1-6.

- Harvard. (23 de octubre de 2018). *What your heart rate is telling you*. Recuperado en mayo de 2020, de Harvard Health Publishing. Harvard Medical School: <https://www.health.harvard.edu/heart-health/what-your-heart-rate-is-telling-you>
- Husain, W., Hui Yng, S. & Abdul, N. (2017). Prediction of Generalized Anxiety Disorder Using Particle Swarm Optimization. *Advances in Information and Communication Technology (ICTA, 2016)*. *Advances in Intelligent Systems and Computing*, 538, pp 480-489.
- Jeong, S., Logan, D. E., Goodwin, M. S., Graca, S., O'Connell, B., Goodenough, H., Anderson, L., Stenquist, N., Fitzpatrick, K., Zisook, M., Plummer, L., Breazeal, C., Weinstock, P. (2015). A Social Robot to Mitigate Stress, Anxiety, and Pain in Hospital Pediatric Care. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human Robot Interaction Extended Abstracts (HRI '15)*, pp. 103-104.
- Jung, M. F., Sirkin, D., Gür, T. M. & Steinert, M. (2015). Displayed Uncertainty Improves Driving Experience and Behavior: The Case of Range Anxiety in an Electric Car. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, pp. 2201-2210.
- Jung, Y. & Yoon, Y. I. (2017). Multi-level assessment model for wellness service based on human mental stress level. *Multimedia Tools and Applications*, 76, pp. 11305-11317.
- K. Yetisen, A., Martinez-Hurtado, J., Ünal, B., Khademhosseini, A., & Butt, H. (2018). Wearables in Medicine. *Advanced Materials*.
- Kelling, C., Pitaro, D. & Rantala, J. (2016). Good vibes: the impact of haptic patterns on stress levels. *Proceedings of the 20th International Academic Mindtrek Conference (AcademicMindtrek '16)*, pp. 130-136.
- Kennedy, L. & Henrickson Parker, S. (2019). Biofeedback as a stress management tool: a systematic review. *Cognition, Technology & Work*, 21, pp. 161-190.
- Kimani, E. (2019). A Sensor-based Framework for Real-time Detection and Alleviation of Public Speaking Anxiety. *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*. EBSE Technical Report, Software Engineering Group, School of Computer Science and Mathematics, Department of Computer Science, University of Durham.

- Kritikos, J., Tzannetos, G., Zoitaki, C., Pouloupoulou, S. & Koutsouris, D. (2019). Anxiety detection from Electrodermal Activity Sensor with movement & interaction during Virtual Reality Simulation. *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*.
- Kushki, A., Khan, A., Brian, J. & Anagnostou, E. (2015). A Kalman Filtering Framework for Physiological Detection of Anxiety-Related Arousal in Children with Autism Spectrum Disorder. *IEEE Transactions on Biomedical Engineering*, 62, 3, pp. 990-1000.
- Kryger, M., Roth, T., & Dement, W. C. (2017). *Principles and Practice of Sleep Medicine*. Philadelphia, U.S.A.: Elsevier.
- Lc, R. & Fukuoka, Y. (2019). Machine Learning and Therapeutic Strategies in VR. *Proceedings of the 9th International Conference on Digital and Interactive Arts (ARTECH 2019)*, pp. 1-6.
- Martinez S. G., Badillo-Urquiola, K. A., Leis, R. A., Chavez, J, Green, T. & Clements, T. (2016). Investigation of Multimodal Mobile Applications for Improving Mental Health. *Schmorrow, D., Difopiastis, C., Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience, Lecture Notes, Computer Science, 9744*, pp. 333-343.
- McGinnis, R., McGinnis, E., Hruschak, J., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L. & Muzik, M. (2018). Rapid Anxiety and Depression Diagnosis in Young Children Enabled by Wearable Sensors and Machine Learning. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*.
- McGinnis, R., McGinnis, E., Hruschak, J., Lopez-Duran, N. L., Fitzgerald, K., Rosenblum, K. L. & Muzik, M. (2018). Wearable sensors and machine learning diagnose anxiety and depression in young children. *2018 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*.
- McGinnis, E., McGinnis, R., Muzik, M., Hruschak, J., Lopez-Duran, N. L., Perkins, N. C., Fitzgerald, K. & Rosenblum, K. L. (2017). Movements Indicate Threat Response Phases in Children at Risk for Anxiety. *IEEE Journal of Biomedical and Health Informatics*, 21, 5, pp. 1460-1465.
- Medina-Mora, M. E., Borges, G., Lara Muñoz, C., Benjet, C., Blanco Jaimes, J., Fleiz Bautista, C., Aguilar-Gaxiola, S. (Agosto de 2003). Prevalencia de trastornos mentales y uso de servicios: resultados de la encuesta nacional de epidemiología psiquiátrica en México. *Salud Mental*, 26, 4.

- Meyer, T., Miller, M., Metzger, R., & Borkovec, T. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28, 6, pp. 487-495.
- Miranda, D., Calderón, M. & Favela, J. (2014). Anxiety detection using wearable monitoring. In *Proceedings of the 5th Mexican Conference on Human-Computer Interaction (MexIHC '14)*, pp. 34-41.
- Miranda, D., Favela, J. & Ibarra, C. (2015). Detecting State Anxiety when Caring for People with Dementia. *Bravo J., Hervás R., Villarreal V, Ambient Intelligence for Health (AmIHEALTH 2015), Lecture Notes, Computer Science*, 9456, pp. 98-109.
- Montelongo, R. V., Zaragoza, A. C., Bonilla, G. M., & Bayardo, S. J. (10 de noviembre de 2005). Los Trastorno de Ansiedad. *Revista Digital Universitaria UNAM*, 6, 11, pp. 2-3.
- Moreno-Pino F, Porrás-Segovia A, López-Esteban P, Artés A, Baca-García E. (2019). Validation of Fitbit Charge 2 and Fitbit Alta HR against polysomnography for assessing sleep in adults with obstructive sleep apnea. *Journal of Clinical Sleep Medicine*, 15, 11, pp. 1645-1653.
- Muniz, A. (2018). Intervenciones en psicología clínica. *Herramientas para la evaluación y el diagnóstico*. Recuperado en mayo de 2020, de Facultad de psicología. Universidad de la República de Uruguay.: <https://psico.edu.uy/sites/default/files/2018-11/Intervenciones%20en%20psicolog%C3%ADa%20cl%C3%ADnica.pdf>
- Myers, D. G. (2005). *Psicología*. Madrid, España: Editorial Medica Panamericana.
- NIH (Development, N. I. (9 de Julio de 2018). *¿Cuánto sueño necesito?* Recuperado en mayo de 2020, de National Institute of Child Health and Human Development: <https://espanol.nichd.nih.gov/salud/temas/sleep/informacion/necesito>
- NINDS. (9 de Julio de 2018). *¿Qué son los ritmos circadianos?* Recuperado en mayo de 2020, de Eunice Kennedy Shriver. National Institute of Child Health and Human Development: <https://espanol.nichd.nih.gov/salud/temas/sleep/informacion/circadianos>
- Nolan, K., Mooney, A. & Bergin, S. (2019). An Investigation of Gender Differences in Computer Science Using Physiological, Psychological and Behavioural Metrics. *Proceeding of the Twenty-First Australasian Computing Education Conference (ACE '19)*, pp. 47-55.

- Nuevo Benítez, R., Montorio Cerrato, I., & Ruiz Díaz, M. Á. (2002). Aplicabilidad del Inventario de Preocupación de Pensilvania (PSWQ) a la población de edad avanzada. *Ansiedad y estrés*, 8, 2-3, pp. 157-172.
- Oficina de Información Científica y Tecnológica para el Congreso de la Unión (INCyTU). (Enero de 2018). Salud mental en México. *Foro Consultivo*, 007, 1-6. Ciudad de México, México: INCyTU.
- Organización Mundial de la Salud (OMS). (13 de abril de 2016). *La inversión en el tratamiento de la depresión y la ansiedad tiene un rendimiento del 400%*. Recuperado el 18 de abril de 2019, de Organización Mundial de la Salud: <http://www.who.int/es/news-room/detail/13-04-2016-investing-in-treatment-for-depression-and-anxiety-leads-to-fourfold-return>
- Organización Panamericana de la Salud. (2015). *Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud. Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud, Volumen 1, Edición de 2015*. (C. C. Centro Venezolano de Clasificación de Enfermedades, Trad.) Washington D.C., E.U.A: Organización Panamericana de la Salud.
- Oti, O., Azimi, I., Anzanpour, A., Rahmani, A., Axelin, A. & Liljeberg. (2018). IoT-Based Healthcare System for Real-Time Maternal Stress Monitoring. *2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*.
- Padros-Blazquez, F., Gonzalez-Betanzos, F., Martinez-Medina, M., & Wagner, F. (2018). Propiedades Psicométricas del Cuestionario de Preocupación Pensilvania (PSWQ) de las Versiones Original y Reducida en Muestras Mexicanas. *Actas Esp Psiquiatra*, 46, 4, pp. 116-24.
- Paliokas, I., Tsakiris, A., Vidalis, A. & Tzovaras, D. (2014). Sense of Presence and Metacognition Enhancement in Virtual Reality Exposure Therapy in the Treatment of Social Phobias and the Fear of Flying. *Shumaker, R., Lackey S. Virtual, Augmented and Mixed Reality, Applications of Virtual and Augmented Reality (VAMR 2014). Lecture Notes, Computer Science*, 8526, pp. 316-328.
- Parnandi, A. & Gutierrez-Osuna, R. (2015). A comparative study of game mechanics and control laws for an adaptive physiological game. *Journal on Multimodal User Interfaces*, 9, pp. 31-42.
- Peake, J., Kerr, G. & Sullivan J. (2018). A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring

- Stress, and Sleep in Physically Active Populations. *Frontiers in Physiology*, 9, 743.
- Pérez de la Mora, M. (Abril-junio de 2003). Dónde y cómo se produce la ansiedad: sus bases biológicas. *Ciencia*, pp. 16-28.
- Pintelas, E. G., Kotsilieris, T., Livieris, I. E., & Pintelas, P. (2018). A review of machine learning prediction methods for anxiety disorders. 2018 *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2018)*, pp. 8-15.
- Price, K., Bird, S. R., Lythgo, N., Raj, I. S., Wong, J. Y. L. & Lynch C. (2017). Validations of the Fitbit One, Garmin Vivofit and Jawbone UP activity tracker in estimation of energy expenditure during treadmill walking and running. *Journal of Medical Engineering & Technology*, 41:3, pp. 208-215.
- Psicosalud. (12 de enero de 2015). *¿Cuándo y por qué sufrimos Ansiedad?* Recuperado el 18 de abril de 2019, de Psicosalud, gabinete de psicología: <https://psicosaludtenerife.com/2015/01/12/cuando-y-por-que-sufrimos-ansiedad/>
- Ramos, K., Stanley, M. Anxiety Disorders in Late Life. (2018). *The Psychiatric clinics of North America*, 41,1, pp. 55–64.
- RedHat. (s.f.). *¿Qué son las API?* Recuperado el 17 de abril de 2019, de RedHat: <https://www.redhat.com/es/topics/api/what-are-application-programming-interfaces>
- Repetto, C., Gaggioli, A., Pallavicini, F., Cipresso, P., Raspelli, S. & Riva, G. (2013). Virtual reality and mobile phones in the treatment of generalized anxiety disorders: a phase-2 clinical trial. *Personal and Ubiquitous Computing* 17, pp. 253-260.
- Rivera, L. G. (10 de octubre de 2017). *Hablemos de la ansiedad en el Día Mundial de la Salud Mental*. Recuperado el 16 de abril de 2019, de TecReview: <https://tecreview.tec.mx/hablemos-de-ansiedad-en-el-dia-mundial-de-la-salud-mental/>
- Rosa, Bruno M. G. & Yang, G. Z. (2019). A Flexible Wearable Device for Measurement of Cardiac, Electrodermal, and Motion Parameters in Mental Healthcare Applications. *IEEE Journal of Biomedical and Health Informatics*, 23, 6, pp. 2276-2285.

- Rother, R., Sun, Y. & Lo, B. (2019). Internet of Things based pervasive sensing of psychological anxiety via wearable devices under naturalistic settings. *Living in the Internet of Things (IoT 2019)*.
- Rubin, J., Eldardiry, H., Abreu, R., Ahern, S., Du, H., Pattekar, A. & Bobrow, A. G. (2015). Towards a mobile and wearable system for predicting panic attacks. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*, pp. 529-533.
- Ruiz, M. M., Autran, R. R., Balderas, E. C., & Huerta, G. C. (2016). La tecnología al servicio de la salud mental en comunidades indígenas. *Frontera Norte*, 28, 556, pp. 135-163.
- Salmerón Enciso, T. A. (2016). *Data Scientist. Software R. De las técnicas clásicas a las avanzadas*.
- Sanches, P., Janson, A., Karpashevich, P., Nadal, C., Qu, C., Daudén Roquet, C., Umair, M., Windlin, C., Doherty, G., Höök, K. & Sas, C. (2019). HCI and Affective Health: Taking stock of a decade of studies and charting future research directions. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*, pp. 1-17.
- Senaaratne, H. (2019). Detecting Temporal Phases of Anxiety in The Wild: Toward Continuously Adaptive Self-Regulation Technologies. *2019 International Conference on Multimodal Interaction (ICMI '19)*, pp. 446-452.
- Sotelo Narváez, V. (2016). *Fundamentos de la Psicología. Manual Autoformativo*. Huancayo-Perú: Universidad Continental.
- Tichon, J. G., Wallis, G., Riek S. & Mavin, T. (2014). Physiological measurement of anxiety to evaluate performance in simulation training. *Cognition, Technology & Work*, 16, pp. 203-210.
- Tiwari, A., Cassani, R., Narayanan, S. & Falk, T. H. (2019). A Comparative Study of Stress and Anxiety Estimation in Ecological Settings Using a Smart-shirt and a Smart-bracelet. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*.
- Tomita, A., Harada, E., Ando, S., Miyashiro, K., Ohmori, M. & Yano, H. (2017). On-time Measurement of Subjective Anxiety of a Passenger in an Autonomous Vehicle: Gradually Changing Sounds Decreases Anxiety of Passenger. *Harris, D., Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness (EPCE 2017), Lecture Notes, Computer Science, 10275*, pp. 209-219.

- UPMC. (2015 de julio de 2015). *What is a normal heart rate? Understanding your health rate*. Recuperado en mayo de 2020, de UPMC. Health Beat: <https://share.upmc.com/2015/07/what-is-a-normal-heart-rate/>
- Vega Calcines, A. (2014). Algoritmos de Aprendizaje Automático: Aplicación en la Solución a Problemas Medio Ambientales. *Cuadernos de Educación y Desarrollo*, 49.
- Watson, D., Clark, L. A., & Stasik, S. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, 11, 3, pp. 429-422.
- Xiao, Y., & Watson, M. (28 de Agosto de 2017). Guidance on Conducting a Systematic Literature Review. *Planning Education and Research*, 39, 1, pp. 93-112.
- Yadav, M., Behzadan, A. & Chaspari, T. (2019). Speak Up! Studying the Interplay of Individual and Contextual Factors to Physiological-Based Models of Public Speaking Anxiety. *2019 First International Conference on Transdisciplinary AI (TransAI)*.
- Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I. M. & Baker, F. C. (2018). A validation study of Fitbit Charge 2 compared with polysomnography in adults. *Chronobiology International: The Journal of Biological and Medical Rhythm Research*, 35:4, pp. 465-476.
- Zheng, Y., C. H. Wong, T., H. K. Leung, B., & C. Y. Poon, C. (2016). Unobtrusive and Multimodal Wearable Sensing to Quantify Anxiety. *Sensors*, 16, 10, pp. 3689-3696.
- Zhou, S., Ogihara, A., Nishimura, S. & Jin Q. (2018). Analyzing the changes of health condition and social capital of elderly people using wearable devices. *Health Information Science and Systems*, 6, 4.

Anexos

Anexo I

Inventario de Preocupación del estado de Pensilvania

Inventario de Preocupación del estado de Pensilvania

Nombre:

Fecha:

Indique hasta qué punto se identifica con cada una de las situaciones que vamos a presentarle a continuación, referidas al modo que tienen las personas de preocuparse.

	1-Nada	2-Algo	3-Regular	4-Bastante	5-Mucho
1. Se preocupa si no tiene suficiente tiempo para hacerlo todo					
2. Sus preocupaciones le agobian					
3. Suele preocuparse por las cosas					
4. Hay muchas circunstancias que hacen que se preocupe					
5. Sabe que no debería preocuparse por las cosas, pero no puede evitarlo					
6. Cuando está bajo tensión tiende a preocuparse mucho					
7. Siempre está preocupándose por algo					
8. Le resulta difícil dejar de lado las preocupaciones					
9. Tan pronto como termina una tarea, en seguida empieza a preocuparse por alguna otra cosa que debe hacer					
10. Se preocupa por todo					
11. Aunque no haya nada más que se pueda hacer por algo, sigue preocupándose por ello					
12. Ha estado preocupado toda su vida					
13. Se da cuenta de que siempre está preocupándose por las cosas					
14. Una vez que comienza a preocuparse por algo, ya no puede parar					
15. Está todo el tiempo preocupándose por algo					
16. Se preocupa por un proyecto hasta que está acabado					

(Nuevo, Montorio y Ruiz, 2002; Ansiedad y Estrés, 8, 157-172; versión original de Meyer, Miller, Metzger y Borkovec, 1990)

Anexo II

Escala de detección del trastorno de ansiedad generalizada

Escala de Detección del Trastorno de Ansiedad Generalizada

Nombre:

Fecha:

Estas cuestiones son para preguntarle sobre cosas que puede haber sentido la mayoría de los días en los últimos 6 meses.

1. La mayoría de los días me siento nervioso/a
 Sí No
2. La mayoría de los días me preocupo por muchas cosas
 Sí No
3. La mayoría de los días no puedo parar de preocuparme
 Sí No
4. La mayoría de los días me resulta difícil controlar mis preocupaciones
 Sí No
5. Me siento inquieto/a, intranquilo/a, o con los nervios de punta
 Sí No
6. Me siento cansado/a fácilmente
 Sí No
7. Tengo problemas para concentrarme
 Sí No
8. Me enfado o irrito fácilmente
 Sí No
9. Mis músculos están tensos y agarrotados
 Sí No
10. Tengo problemas de sueño
 Sí No
11. Las cosas que ha señalado anteriormente, ¿afectaron su vida diaria (en el hogar, trabajo o en su tiempo libre) o le causaron mucho malestar?
 Sí No
12. Las cosas que ha señalado anteriormente, ¿fueron suficientemente molestas como para pensar en buscar ayuda para ellas?
 Sí No

(Validada al español por Bobes, García-Calvo, Prieto, García-García y Rico. Villademoros, versión original de Carroll y Davidson, 1990)

Anexo III

Inventario de Preocupación del estado de Pensilvania (15 ítems)

Inventario de Preocupación del estado de Pensilvania

Nombre:

Fecha:

Indique hasta qué punto se identifica con cada una de las situaciones que vamos a presentarle a continuación, referidas al modo que tienen las personas de preocuparse.

	1-Nada	2-Algo	3-Regular	4-Bastante	5-Mucho
Sus preocupaciones le agobian					
Suele preocuparse por las cosas					
Hay muchas circunstancias que hacen que se preocupe					
Sabe que no debería preocuparse por las cosas, pero no puede evitarlo					
Cuando está bajo tensión tiende a preocuparse mucho					
Siempre está preocupándose por algo					
Le resulta difícil dejar de lado las preocupaciones					
Tan pronto como termina una tarea, en seguida empieza a preocuparse por alguna otra cosa que debe hacer					
Se preocupa por todo					
Aunque no haya nada más que se pueda hacer por algo, sigue preocupándose por ello					
Ha estado preocupado toda su vida					
Se da cuenta de que siempre está preocupándose por las cosas					
Una vez que comienza a preocuparse por algo, ya no puede parar					
Está todo el tiempo preocupándose por algo					
Se preocupa por un proyecto hasta que está acabado					

(Nuevo, Montorio y Ruiz, 2002; Ansiedad y Estrés, 8, 157-172; versión original de Meyer, Miller, Metzger y Borkovec, 1990)

Anexo IV

Escala de detección del trastorno de ansiedad generalizada (6 ítems)

Escala de Detección del Trastorno de Ansiedad Generalizada

Nombre:

Fecha:

Estas cuestiones son para preguntarle sobre cosas que puede haber sentido la mayoría de los días en los últimos 6 meses.

1. La mayoría de los días me siento nervioso/a
 Sí No
2. La mayoría de los días me preocupo por muchas cosas
 Sí No
3. La mayoría de los días no puedo parar de preocuparme
 Sí No
4. La mayoría de los días me resulta difícil controlar mis preocupaciones
 Sí No
5. Me siento inquieto/a, intranquilo/a, o con los nervios de punta
 Sí No
6. Tengo problemas de sueño
 Sí No

(Validada al español por Bobes, García-Calvo, Prieto, García-García y Rico.Villademoros, versión original de Carroll y Davidson, 1990)

Anexo V

Consentimiento informado



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México
Centro Nacional de Investigación y Desarrollo Tecnológico

cenidet[®]
Centro Nacional de Investigación
y Desarrollo Tecnológico

CONSENTIMIENTO INFORMADO PARA PARTICIPAR EN UN ESTUDIO DE INVESTIGACIÓN

Tiempo aproximado de lectura: 5 minutos.

Introducción

El Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) es un centro de investigación educativo que forma parte del Tecnológico Nacional de México (TecNM). El CENIDET tiene como objetivo impulsar la ciencia, la tecnología y la innovación, buscando contribuir a la transformación de México en una sociedad del conocimiento, así como consolidar la vinculación con los sectores público, social y privado.

La investigación que pertenece a este estudio lleva por nombre: *Detección automática de ansiedad a través del monitoreo de sensores vestibles*. Tiene como objetivo detectar el trastorno de ansiedad generalizada de forma automática a través de herramientas tecnológicas, como los sensores vestibles, determinando un patrón de comportamiento de ansiedad. Para poder verificar el correcto funcionamiento de este patrón de comportamiento, es necesario contar con la participación de personas en un monitoreo para poder detectar la ansiedad.

Descripción del estudio

De manera inicial, se te procederá a monitorear, con la ayuda de pulseras de actividad inteligentes, los síntomas fisiológicos que presentes en el periodo que duren las pruebas de estudio. Esto tiene como finalidad el monitorear tu frecuencia cardíaca, tu calidad de sueño, tu movimiento o actividad en relación con tus actividades diarias. De esta manera, se te hará un monitoreo en tiempo real sin que te sientas incómodo y sin que afecte a tus actividades diarias.

Descripción de las pruebas del estudio

El estudio se divide una sola etapa, a continuación, se te explica en qué consiste a detalle.

Etapa I

En esta etapa se te realizará un monitoreo en tus actividades diarias con una pulsera de actividad que usarás por un periodo de 16 días hábiles.



TECNOLÓGICO
NACIONAL DE MÉXICO

Tecnológico Nacional de México
Centro Nacional de Investigación y Desarrollo Tecnológico

cenidet[®]
Centro Nacional de Investigación
y Desarrollo Tecnológico

Con los resultados obtenidos de la pulsera de actividad, se podrá detectar el patrón de comportamiento de tu frecuencia cardíaca, tu calidad de sueño, tu movimiento o actividad en relación con tus actividades diarias y con el trastorno de ansiedad generalizada.

Riesgos y beneficios de la prueba de estudio

No existen riesgos asociados a este estudio en cuanto a la invasión de la tecnología a utilizar.

Derechos

La participación es totalmente voluntaria. Tienes todo el derecho a abstenerte de participar o retirarte del estudio en cualquier momento.

Confidencialidad

La información recolectada será usada únicamente con fines propios a la investigación. Se mantendrá plena confidencialidad de tus datos personales como participante.

Para cualquier aclaración, queja o sugerencia, puedes comunicarte con Alejandra Cuevas al (XXX) XXXXXXX, al e-mail: XXXX@cenidet.edu.mx, con la especialista en el trastorno de ansiedad generalizada y co-directora de tesis, la maestra Fabiola Sánchez, al e-mail: XXXX@uaem.mx o con la dra. Alicia Martínez Rebollar, profesora-investigadora del CENIDET y directora de la tesis, al e-mail: XXXX@cenidet.edu.mx. Sus comentarios son importantes para nuestro estudio.

Nombre del participante	Firma	Fecha
Documento de identificación oficial:		

Nombre del responsable de la evaluación	Firma	Fecha
---	-------	-------