



TECNOLÓGICO NACIONAL DE MÉXICO
TECNOLÓGICO DE ESTUDIOS SUPERIORES TIANGUISTENCO
DIVISIÓN DE INGENIERIA INDUSTRIAL



TECNOLÓGICO DE ESTUDIOS SUPERIORES DE TIANGUISTENCO
DIVISIÓN DE INGENIERIA INDUSTRIAL

“Machine Learning en Control de Procesos”

TESIS

QUE PARA OBTENER EL TITULO DE:

INGENIERO INDUSTRIAL

P R E S E N T A :

Sharon Madai Zetina Almazán

DIRIGIDA POR:

Dra. Cristina Arely de León Condes

Dra. Maribel González Torres

Tianguistenco, Estado de México, mayo 2023

Agradecimientos

Le agradezco muy profundamente a la Dra. Cristina Arely de León Condes, por todo el asesoramiento, conocimientos y consejos compartidos, así como también, por su paciencia y constancia que demostró hacia mi persona, sin sus palabras y correcciones precisas, nada de esto habría sido posible.

Agradezco por su dedicación, perseverancia y tolerancia a la Dra. Maribel González Torres, su guía y apoyo dieron pauta a la motivación para mejorar día con día, gracias porque me hizo entender que siempre seré capaz de mucho más.

A mi padre, por siempre brindarme el apoyo incondicional y nunca rendirse, porque con cada acción demuestra que todo es posible, sin sus consejos y enseñanzas nunca habría llegado hasta este lugar, espero algún día se sienta tan orgulloso de mi, como yo lo estoy de él.

A mi madre que sin duda siempre ha sido la principal promotora de mis sueños, gracias por demostrarme su fortaleza ante todos los obstáculos que se anteponen en el camino de la vida y por el amor ilimitado e incondicional que siempre han sido mi mayor inspiración, gracias por acompañarme en las noches agotadoras de estudio, y por confiar y creer en mí.

A mis hermanos por todos los consejos y enseñanzas que aportaron durante el transcurso de esta etapa, por ser partícipes en gran parte de mi felicidad y por estar siempre cuando más los necesito. Porque, aunque pareciera que siempre estamos en una batalla por cualquier cuestión, cada momento a su lado ha sido la clave perfecta para alcanzar todos nuestros objetivos.

A la Lic. Nayelly Osorio Hernández por ser parte fundamental de todas las acciones que han surgido durante esta etapa, por haberme guiado a lo largo del periodo de mis estudios de inicio a fin, así como también por haberme escuchado en los momentos más difíciles de mi carrera profesional.

A la institución del Tecnológico de Estudios Superiores por abrirme las puertas y ponerme al pie de todos los docentes que aportaron los conocimientos necesarios para crecer profesionalmente.

A todos, muchas gracias.

Resumen

La creación de fábricas inteligentes ha dado un gran paso a la posibilidad de resolver problemas y tomar decisiones de manera eficiente con ayuda de las nuevas tecnologías, las cuales son efecto de la cuarta revolución industrial o industria 4.0. En las últimas décadas, el incremento en las implementaciones de Machine Learning (ML) como técnica de la inteligencia artificial, en el ámbito industrial, ha desencadenado optimizaciones exitosas en el control de procesos, estas han surgido como resultado de las predicciones alcanzadas a través de los algoritmos informáticos que se implementan en máquinas a través de softwares para que la información sea clasificada por patrones y tendencias encontradas de manera autónoma.

Este documento, registra un análisis breve de las implementaciones más importantes y con resultados positivos de la toma de decisiones en áreas organizacionales. Así mismo se enlistan y describen los conceptos mayormente sustanciales que conllevan al entendimiento de esta tecnología, la cual visualiza un mayor impacto en su desarrollo a corto o mediano plazo, representando un pilar importante dentro de las revoluciones industriales futuras. Esta descripción se respalda por la exploración y revisión analítica de casos en diferentes industrias con mayor énfasis en las aplicaciones de control de procesos, artículos pertenecientes a diferentes revistas científicas de renombre. Se describe la sucesión de pasos que se llevaron a cabo en el desarrollo de la propuesta generalizada, basada en el análisis del funcionamiento de cada implementación, para ser ejecutada en implementaciones de cualquier área industrial con recomendaciones de lo más utilizado.

Palabras clave: Machine Learning, control de procesos industriales, toma de decisiones, industria 4.0, inteligencia artificial.

Abstract

The creation of smart factories has taken a big step towards the possibility of solving problems and making decisions efficiently with the help of new technologies, which are the effect of the fourth industrial revolution or industry 4.0. In recent decades, the increase in the implementations of Machine Learning (ML) as an artificial intelligence technique, in the industrial field, has triggered successful optimizations in process control, these have emerged as a result of the predictions achieved through computer algorithms. that are implemented in machines through software so that the information is classified by patterns and trends found autonomously.

This document records a brief analysis of the most important implementations and with positive results of decision making in organizational areas. Likewise, the most substantial concepts that lead to the understanding of this technology are listed and described, which visualizes a greater impact on its development in the short or medium term, representing an important pillar within future industrial revolutions. This description is supported by the exploration and analytical review of cases in different industries with greater emphasis on process control applications, articles belonging to different renowned scientific journals. The succession of steps that were carried out in the development of the generalized proposal is described, based on the analysis of the operation of each implementation, to be executed in implementations of any industrial area with recommendations of the most used.

Keywords: Machine Learning, industrial process control, decision making, industry 4.0, artificial intelligence.

Índice	
Agradecimientos	3
Resumen.....	5
Abstract	6
Índice.....	7
Índice de figuras.....	9
Índice de tablas	12
CAPÍTULO I. GENERALIDADES	13
1.1 Introducción	14
1.2 Justificación.....	16
1.3 Planteamiento del problema	18
1.4 Objetivos	19
1.4.1 Objetivo general	19
1.4.2 Objetivos específicos.....	19
CAPÍTULO II. MARCO TEÓRICO	20
2.1 Industria 4.0.....	21
2.1.1 Inteligencia artificial	27
2.2 Machine Learning	30
2.2.1 Tipos de algoritmos Machine Learning	33
2.2.2 Lenguajes y software de programación para Machine Learning	47

2.2.3 Sistemas de recolección de datos	50
CAPÍTULO III. REVISIÓN DE LA LITERATURA	53
3.1 Artículos de revisión	54
3.2 Análisis de la literatura.....	80
CAPÍTULO IV. METODOLOGÍA Y DESARROLLO.....	88
4.1 Descripción de las actividades	89
CAPÍTULO V. RESULTADOS	94
5.1 Resultados y discusión	95
5.1.1 Planificar	96
5.1.2 Hacer/implementar.....	110
5.1.3 Verificar	112
5.1.4 Actuar/Estandarizar.....	113
5.2 Conclusiones	115
5.3 Recomendaciones.....	117
Referencias.....	120

Índice de figuras

Figura 1. Elementos de la industria (Elaboración propia con información de Stentoft, J., et al., 2019)	21
Figura 2. Características de la industria 4.0 (Smit, J., et al., 2016) (Elaboración propia)	24
Figura 3. Factores importantes de las 4 revoluciones industriales (Elaboración propia con información de Stearns, P., 2020).....	26
Figura 4. División de los campos de la inteligencia artificial (Elaboración propia con información de (Elaboración propia con información de Cioffi, R., et al., 2020).....	29
Figura 5. Relación de tecnologías de la industria 4.0 con la inteligencia artificial y sus ramas (Elaboración propia con información de Woschank, M., et al., 2020).....	30
Figura 6. Patrón que rige al funcionamiento de Machine Learning. (Elaboración propia con información de (Elaboración propia con información de Sarkar, D., et al., 2018).....	31
Figura 7. Principales disciplinas que rigen en los procesos de creación de programas ML (Sarkar, D., et al., 2018).....	33
Figura 8. Métodos de aprendizaje automático y sus diferencias (Elaboración propia con información de Sarker, I., 2021)	35
Figura 9. Proceso del modelo de aprendizaje supervisado (Elaboración propia con información de Sarker, 2021)	36
Figura 10. Principales algoritmos del aprendizaje supervisado (Elaboración propia con información del Misilmani., et al., 2019).....	37
Figura 11. Tipología de la regresión. (Elaboración propia con información de Sarkar et al., 2018)	38

Figura 12. Comportamiento de los datos en función a tres tipos de regresión (Elaboración propia con información de Christodoulou, E., et al., 2019)	39
Figura 13. Estructura general de una red neuronal artificial (Sarkar et al., 2018).....	40
Figura 14. Funcionamiento de las maquinas vector soporte (Elaboración propia con información de Sarkar, D., et al., 2018)	41
Figura 15. Formación de grupos de elementos en aprendizaje no supervisado (Elaboración propia con información de Sarkar, D., et al., 2018).....	42
Figura 16. Ramas del aprendizaje no supervisado (Naous, T., 2019)	42
Figura 17. Agrupamiento de elementos de manera hibrida del método semi-supervisado (Elaboración propia con información de Sarker, I., 2021)	43
Figura 18. Sucesión de pasos del aprendizaje por refuerzo (Elaboración propia con información de Sarker, I., 2021).....	45
Figura 19. Flujo general (elaboración propia con información de Mahesh, B., 2020).....	46
Figura 20. Proceso para llegar al aprendizaje por transferencia (Elaboración propia con información de Weiss, K., et al., 2016)	47
Figura 21. Principales librerías de Python enfocadas a Machine Learning (Elaboración propia con información de Sarkar, D., et al., 2018).....	48
Figura 22. Matlab y Toolbox (Elaboración propia con información de Sarkar et al., 2018).....	49
Figura 23. Elementos de aprendizaje automático a partir de Julia (Gao et al., 2020)	50
Figura 24. Relaciones de SCADA (Elaboración propia con información de Yadav, G., et al., 2021)	51
Figura 25. Áreas funcionales de una empresa con implementación Machine Learning (Elaboración propia).	80

Figura 26. Especificaciones (Elaboración propia).....	82
Figura 27. Herramientas de lenguaje y programación (Elaboración propia).....	84
Figura 28. Modelos Machine Learning Utilizados (Elaboración propia).....	85
Figura 29. Número de artículos por revista y su factor de impacto. (Elaboración propia).....	86
Figura 30. País de procedencia (Elaboración propia).....	87
Figura 31. Bosquejo del método (Elaboración propia).....	89
Figura 32. Documentos para indagar en el buscador académico (elaboración propia).....	90
Figura 33. Palabras clave (Elaboración propia).....	91
Figura 34. Ciclo Deming (Elaboración propia con información de Isniah S., et al., 2020).....	95
Figura 35. Dependencias (Elaboración propia).....	98
Figura 36. Artículos por dominio de implementación (Elaboración propia).....	99
Figura 37. Algoritmos en producción (Elaboración propia).....	101
Figura 38. Actividades generales de cada responsable para la implementación (Elaboración propia).....	102
Figura 39. Proceso general de las actividades del jefe de proyecto (Elaboración propia).....	105
Figura 40. Proceso general del aprendizaje por refuerzo (Elaboración propia).....	107
Figura 41. Aprendizaje no supervisado (Elaboración propia).....	107
Figura 42. Proceso de datos etiquetados y no etiquetados para obtener elementos que permitan la implementación del aprendizaje supervisado. (Elaboración propia).....	108
Figura 43. Técnicas y algoritmos de aprendizaje supervisado mayormente utilizados en las implementaciones (Elaboración propia).....	109
Figura 44. Pasos de verificación.....	113
Figura 45 Composición de las áreas (Elaboración propia).....	114

Índice de tablas

Tabla 1. Clasificación de los datos (Elaboración propia con información de Erraissi, A., et al., 2018)	22
Tabla 2. Principales tecnologías de la industria 4.0 (Elaboración propia con información de Javaid, M., et al., 2020).....	27
Tabla 3. Tipología de SCADA (Elaboración propia con información de Yadav & Paul, 2021)..	52
Tabla 4. Resumen de artículos (Elaboración propia).....	54
Tabla 5. Desarrollo de información (Elaboración propia)	92
Tabla 6. Herramientas para la medición de la calidad (Elaboración propia con información de Isniah S., et al., 2020)	96
Tabla 7. Designación e variables (Elaboración propia)	100
Tabla 8. Costos por lenguaje de programación (Elaboración propia).	110
Tabla 9. Competencias adquiridas (Elaboración propia).....	118

CAPÍTULO I.

GENERALIDADES

1.1 Introducción

El Machine Learning es una derivación de las tecnologías de la industria 4.0, el cual, tiene sus inicios antes de 1950, con la creación del “Test de Turing”. En 1952 el profesor e informático teórico Arthur Samuel presento un software apto para jugar damas chinas a través de información almacenada sobre las diferentes técnicas utilizadas, de esta manera, controlaba las mejoras del nivel del juego, gracias a esto era capaz de mejorar las técnicas con ayuda de todos los datos almacenados en la base de datos. Sin embargo, estas propuestas e innovaciones que se presentaron estaban en riesgo de ser ignoradas debido a que se consideraba que su desarrollo era altamente costoso. Fue hasta el año 1979 que estudiantes de Stanford con ayuda de la herramienta principal de la inteligencia artificial Nearest Neighbor, crearon a “Stanford car” detonando con ello un incremento de la creación de modelos Machine Learning, los cuales fueron aceptados positivamente por diferentes organizaciones teniendo mayor tendencia en sus aplicaciones exitosas a partir del año 2000 con los descubrimientos de su eficacia dentro de áreas médicas e industriales.

El análisis de datos dentro de la industria es una herramienta importante y óptima para la toma de decisiones. Dado que Machine Learning es una rama de la Inteligencia artificial la cual tiene como función principal, relacionarse con algoritmos computarizados capaces de encontrar y analizar antecedentes de datos obteniendo con ello un estándar con el fin de mejorar las previsiones y predicciones de procesos contrarrestando al máximo la intervención humana. Así mismo se ha utilizado actualmente como una técnica fundamental en el sector corporativo. Dentro del sector industrial, la intervención de las nuevas tecnologías en los procesos para dar respuesta a las nuevas necesidades es indispensable. La globalización permite que las organizaciones industriales optimicen todo tipo de problemas relacionados con la

interconectividad de acuerdo con la generación de nuevas técnicas de mejora. Las cuales han sido derivado de un análisis exhaustivo de datos que han permitido la toma de decisiones considerando todas las variables que competen en el incremento o decremento de resultados exitosos en la optimización industrial a partir de las predicciones impactando en el control de procesos industriales. Este documento, cumple el principal objetivo de abordar una investigación de las tendencias potenciales del tema en cuestión y sus aplicaciones en la industria, sin eludir la delimitación de la información actualizada con la finalidad de obtener instrumentos teóricos para determinar las exigencias de los requerimientos que representan la evolución del buen funcionamiento del software y así verificar que los resultados propuestos tengan consecuencias funcionales para la toma de decisiones en todas las áreas que componen una industria. Así mismo el documento consta de cinco capítulos a saber:

Capitulo I: se presentan las generalidades que representa su elaboración, tomando en cuenta el objeto de la investigación y los aspectos formales.

Capitulo II: Se desarrolla un marco teórico que contiene los conceptos y planteamientos teóricos registrados sobre el tema de investigación

Capitulo III: Se registra la revisión de la literatura, los acontecimientos y resultados más importantes con respecto a las implementaciones de Machine Learning en control de Procesos industriales

Capitulo IV: Se enlista un conjunto de métodos y etapas que se siguen en la presente investigación y se describe el desarrollo ejecutado para cumplir con el objetivo general.

Capitulo V: se presenta los resultados como propuesta de etapas a realizar para una posible implementación en la industria a partir de la metodología PDCA.

1.2 Justificación

El Machine Learning resulta ser una herramienta importante dentro del sector industrial, debido a su capacidad de brindar una visión en las tendencias del comportamiento de todos los procesos.

Además de respaldar el desarrollo de nuevos productos, este proceso se da mediante softwares especializados en crear modelos que permiten que la información sea precisa en la predicción de resultados sin estar programadas explícitamente, utilizando datos históricos como entrada para predecir nuevos valores de salida. Esta nueva herramienta, permite que las organizaciones obtengan información útil para evitar riesgos que aún se desconocen a partir de una toma de decisiones previamente analizada

Machine Learning se ha venido aplicando en años recientes abarcando muchos tipos de programas con los que se encontrará en el análisis de big data y la minería de datos. Existe ya una gran variedad de registros del éxito en las aplicaciones más relevantes que se han obtenido, hoy en día es de vital importancia que las organizaciones actualicen sus procesos con las nuevas tecnologías que las revoluciones industriales ofrecen, para que los procesos se optimicen de acuerdo a las necesidades económicas, sociales y ambientales.

Al aplicar machine Learning es posible analizar grandes conjuntos de datos sin errores. Esto ayuda a concentrarse mejor en el trabajo relevante para el campo de diferentes grupos multidisciplinarios.

Las búsquedas de tendencias actuales para aplicarlas en la ingeniería Industrial incluyen la reducción del consumo de energía, la minimización del impacto ambiental y un mayor enfoque en la automatización, este es el incentivo de introducir temas como este trabajo de

investigación. Por tal motivo este trabajo se enfoca en adquirir información y datos más importantes de los registros para emitir un informe de las delimitaciones y tendencias, es decir, se dirige al estudio de las inclinaciones encontradas en cuanto a la estadística de las áreas industriales en las que mayor impacto se tiene.

Machine Learning será una de las tecnologías que ayude a maximizar los procesos después de ser analizados y posteriormente predecir los resultados que mejor satisfagan el proceso en la solución de problemas.

El impacto que cualquier organización puede obtener al aplicar inteligencia artificial tiene relación con un impacto tecnológico, que a su vez deriva en el económico en un mediano o largo plazo y que puede contribuir en la optimización de procesos que se alinean en la mayoría de casos a la reducción de costos.

1.3 Planteamiento del problema

Machine Learning, derivación de la inteligencia artificial que permite a las maquinas aprender a partir de una base de datos, hacer predicciones, sin ser programadas por el hombre, además que es una herramienta principal que servirá para atender a la necesidad de crear técnicas innovadoras y proponer mejoras en el control de procesos industriales y obtener resultados exitosos a partir de la toma de decisiones. Debido a que los procedimientos industriales cambian de acuerdo a las nuevas necesidades del cliente, en la actualidad es indispensable fomentar el cuidado del medio ambiente y optimizar el tiempo al máximo posible, es así que Machine Learning será la herramienta que permita analizar la información para tomar decisiones, evitando errores en su impacto. Dado que esta técnica ha sido muy poco implementada dentro del sector industrial en comparación con el sector médico debido a que en sus inicios el principal objetivo estaba alineado a la resolución de problemas del área médica, por lo que se pretende que los análisis arrojados nos lleven a la creación de técnicas eficientes para crear modelos que impacten positivamente en los cambios del proceso. Los datos con los que se pretende realizar el análisis son de características industriales que han influenciado en el rendimiento y eficiencias de las mejoras en el desarrollo industrial, es decir, se analiza el desenvolvimiento en áreas como almacén, logística, cadena de suministro y producción, un área en la que el impacto de las predicciones registradas por modelos Machine Learning han resultado exitosas según (Min et al., 2019). En dichas áreas, las predicciones que se realizan, crean la capacidad en el proceso de evitar mayormente la intervención humana y de esta manera es posible reducir tiempos.

1.4 Objetivos

1.4.1 Objetivo general

Determinar las aplicaciones de Machine Learning, tipos de algoritmos y lenguajes de programación de mayor utilización dentro de una sucesión de pasos de las implementaciones en áreas industriales.

1.4.2 Objetivos específicos

- Investigar tendencias potenciales de Machine Learning y clasificar artículos con información detallada de las implantaciones delimitando el tema de lo más aplicado en la actualidad.
- Determinar las áreas en que frecuentemente se ha implementado Machine Learning
- Comprender el tipo de algoritmo según los datos a analizar
- Investigar los lenguajes de programación dependiendo de su aplicación y precio.

CAPÍTULO II.

MARCO TEÓRICO

2.1 Industria 4.0

Industria 4.0 o Cuarta revolución industrial, es un término que puntualiza la constitución organizacional específicamente de los procesos de fabricación y producción que están basados en las nuevas tecnologías y dispositivos que poseen autonomía en el establecimiento de la comunicación entre ellos, es decir, dispositivos que no necesitan de intervención humana directa para compartir información durante el plazo de la cadena de valor en patrones informáticos. Este concepto abarca una estructurada implantación de los medios informáticos con la que cuentan las empresas manufactureras o el enriquecimiento de redes de información que integran la totalidad de las características de los objetos físicos que se encuentran dentro del entorno propio. En la figura 1 se muestra el objetivo principal de esta revolución, el cual esta direccionado a la creación de fábricas inteligentes y sus 4 elementos principales con las que es capaz de lograrlo (Stentoft, J., et al., 2019).

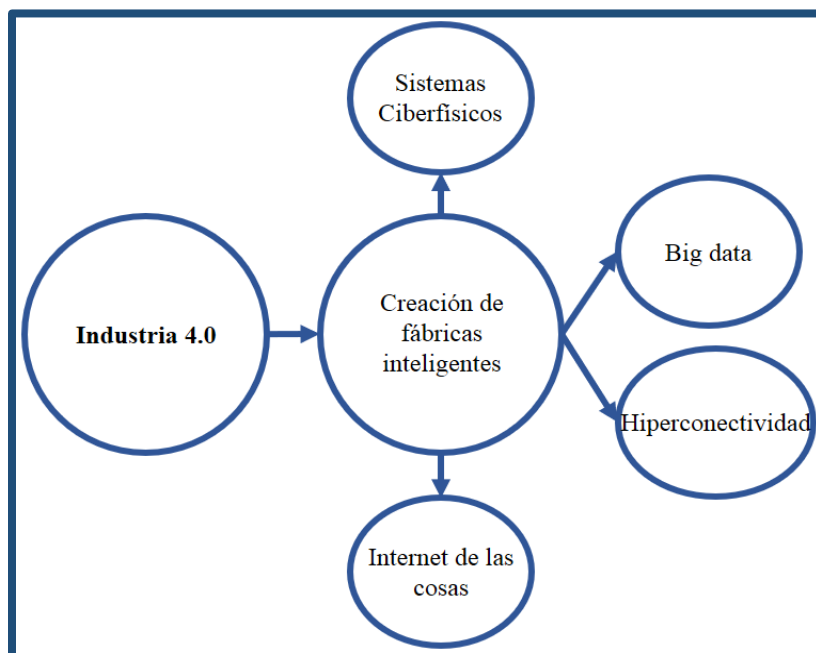


Figura 1. Elementos de la industria (Elaboración propia con información de Stentoft, J., et al., 2019)

Los sistemas ciber-físicos que otorgan a un entorno manufacturero la posibilidad de comunicarse entre los operarios, almacenamiento masivo, y la posibilidad de un seguimiento óptimo para mantener un control del mundo tangible a través de un mundo virtual. IOT o Internet de las cosas es actualmente una de las tecnologías más importantes como técnica para conectar cualquier objeto y con la cual se mantiene comunicación en tiempo real y a cualquier distancia entre procesos industriales, objetos importantes dentro del sistema y el personal, lo que facilita la prevención de fallas y errores. La hiperconectividad, que no es más que todos los medios que mantienen a la humanidad en constante comunicación a través del internet, la telefonía y mensajería, correo electrónico y lo más utilizado en la actualidad, las redes sociales, esto, dentro de las empresas facilita el intercambio rápido de información y en tiempos precisos. Y por último Big Data, referente al análisis de la gran cantidad de datos que a su vez son clasificables como etiquetados y no etiquetados (Tabla 1), con el fin de obtener ideas claras que conducen a una mejor toma de decisiones para dar solución a problemas o simplemente crear mejora continua dentro del proceso (Stentoft, J., et al., 2019).

Tabla 1. Clasificación de los datos (Elaboración propia con información de Erraissi, A., et al., 2018)

Etiquetados	No etiquetados
<p>Conjunto de información bien definida, que ya se conoce. Normalmente es muy fácil de interpretar debido a que se encuentra bien organizada, y en formatos ya estandarizados.</p>	<p>Información con organización escasa, no existen formatos estandarizado y carecen de características específicas. No cuentan con estructura ni clasificación.</p>

Actualmente, el desarrollo y transformación digital han cambiado drásticamente la cotidianidad humana. Industria 4.0 resulta ser una combinación de técnicas y nuevas tecnologías para mejorar el control de procesos industriales. Es tarea de cada organización identificar las virtudes de cada tecnología para elegir la que mejor satisfaga la demanda a la solución de problemas. Una de las consecuencias que esta revolución provoca, es el cambio en el comportamiento de la sociedad, por lo tanto, los requerimientos del cliente se mantienen en constante modificación y se vuelve crucial la integración de las nuevas tecnologías al proceso, para darle mejoras a los productos y servicios, apegados a la actualización de las necesidades y así mantener activo el funcionamiento organizacional. La industria 4.0 trae consigo una serie de cualidades que la distinguen de las demás revoluciones (Figura 2). La virtualización otorga la oportunidad de crear simulaciones de entornos físicos de manera digital limitándose al conjunto de componentes o materiales del sistema informático que se utiliza, para establecer el panorama visual de lo que posiblemente resultara en el futuro. Su principal competencia se dirige a la orientación al servicio, dotar de aptitudes serviciales para las personas priorizando las necesidades del cliente antes que las necesidades empresariales, con la finalidad de aumentar la demanda de lo producido. Las empresas obtienen productividad y competitividad a partir del fomento en las innovaciones del sistema para el desarrollo del crecimiento en la diversidad de productos, adaptándose a las necesidades en tendencia como estrategia de modularidad. A través de dispositivos que captan magnitudes físicas o alteraciones en el entorno, las máquinas tienen el poder de tomar decisiones con mayor rapidez sin depender de los humanos a esto se le llama descentralización, aun cuando este proceso es automatizado, la industria 4.0 promueve la comunicación precisa entre todos los elementos que integran la organización sean máquinas o

humanos en tiempo real exhortado por las características capacidad en tiempo real e interoperabilidad. (Smit, J., et al., 2016).

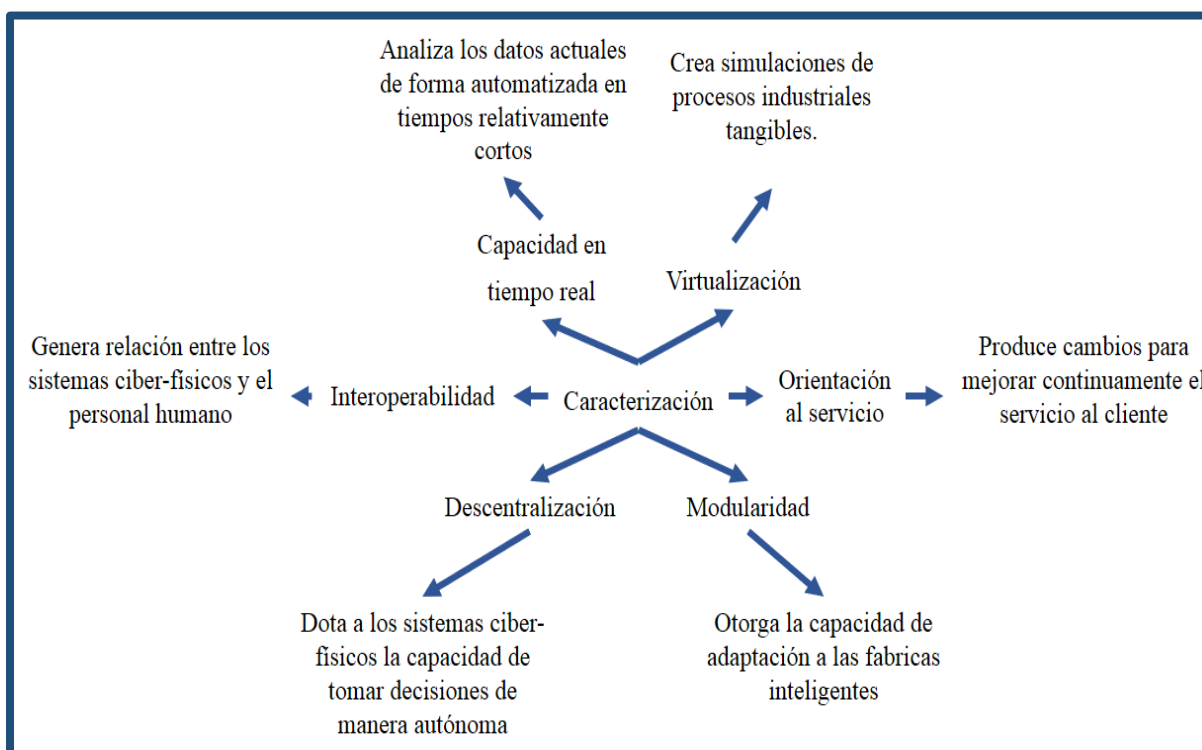


Figura 2. Características de la industria 4.0 (Smit, J., et al., 2016) (Elaboración propia)

Cada una de estas características, son fragmentos que impactan en diferentes ámbitos entre sí. Es importante conocer las características de cada revolución industrial iniciada en Inglaterra en el siglo XVIII se trata de cambios que se dieron en los procesos de productividad de la humanidad en un periodo determinado de tiempo, pasando de la ganadería y agricultura, a la llegada de procesos industrializados y su mecanización para la transformación económica. En la figura 3, se presentan los rasgos de cada una de las revoluciones, que conducen a la creación de los modelos de automatización. Durante la primera revolución industrial que permitió una transformación de la sociedad basada en la agricultura a los procesos de fabricación que

funcionaban con carbón como principal energía y el principal medio de transporte eran los trenes.

Para la segunda revolución industrial comenzó cuando se inventó el motor de combustión interna, ahora utilizaban petróleo y electricidad para la producción de grandes cantidades del mismo producto en menor tiempo.

La tercera revolución industrial representa el uso de las tecnologías de la información y la ejecución y puesta en marcha de la electrónica para la automatización en los procesos, lo cual provocó una disminución importante de la mano de obra para el sistema.

Y la cuarta revolución industrial que a diferencia de la tercera en la que algunos procesos implicaban que la mano de obra tuviera que ensamblar varios componentes de una sola pieza con ayuda de diferentes técnicas, está permitiría la llegada de las nuevas tecnologías con lo que sería posible el diseño de piezas a computadora para poder ser materializadas a partir de la impresión tridimensional.

Con la existencia de los inéditos retos empresariales y la gran disponibilidad de nuevas herramientas tecnológicas para afrontarlos, es posible el crecimiento de manera exponencial de las posibles soluciones para la creación de valor en el proceso (Stearns, P., 2020)

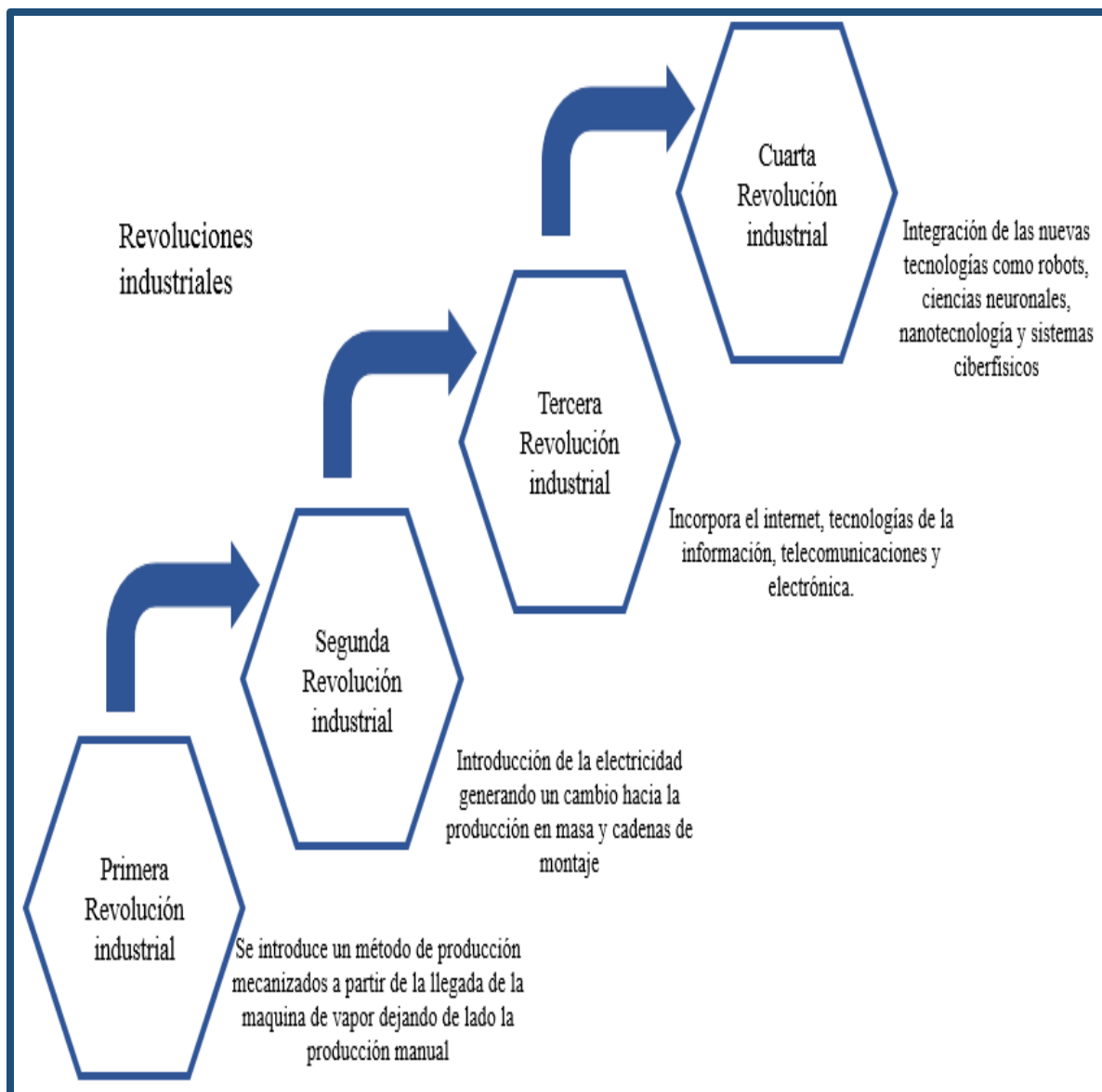


Figura 3. Factores importantes de las 4 revoluciones industriales (Elaboración propia con información de Stearns, P., 2020)

La tabla 2, enlista las tecnologías que impactan en el cambio de los procesos industriales para satisfacer los conceptos de automatización en procesos productivos, y que dan pauta a las herramientas necesarias para la creación de fábricas inteligentes, son tecnologías con mayor

incidencia de aplicación en áreas industriales, sin embargo, existen muchas más que pueden ser implementadas en otros entornos organizacionales (Javaid, M., et al., 2020).

Tabla 2. Principales tecnologías de la industria 4.0 (Elaboración propia con información de Javaid, M., et al., 2020)

Tecnologías presentes en la industria 4.0		
Big data y análisis	Internet de las cosas	Biología sintética
Fabricación aditiva	Inteligencia artificial	CPS o sistemas ciber-físicos
Robot colaborativo	Computación en la nube	Ciberseguridad
Realidad Virtual	Biosensores	Impresión 3D

2.1.1 Inteligencia artificial

Una de las tecnologías más flexibles de la industria 4.0 es la inteligencia artificial, básicamente, el termino, hace referencia a la capacidad que tienen las máquinas u ordenadores de imitar la inteligencia humana. Las máquinas poseen facultades basadas en algoritmos y datos informativos de todo lo sucedido históricamente para poder crear informes y permitir la toma de decisiones autónoma sin demandar acciones humanas.

Esta tecnología permite que todos los procesos que solo se tenían reservados para ser ejecutados por los humanos, ahora sean realizados por los sistemas de inteligencia artificial, que, a diferencia de los humanos, ventajosamente no tienen la necesidad de descansar y es posible analizar grandes cantidades de datos en tiempos relativamente cortos para dar respuesta a

soluciones óptimas. Superando así, el desempeño humano dentro de los sectores industriales (Arrieta, A., et al., 2020)

Durante las últimas décadas, la inteligencia artificial se ha convertido en un tema tendencial consecuente a las implementaciones exitosas que se han tenido en campos como; ingeniería, industria, ciencia, educación, finanzas; y que provocaron un impacto importante en la operatividad de la sociedad.

La inteligencia artificial integra distintos instrumentos o sistemas informáticos creados por la inteligencia humana. Se trata de 16 campos que se desarrollan para ser implementados cuando se consideren útiles, cada uno de ellos están direccionados a la satisfacción de diversas necesidades. Sin embargo, en conjunto se dirigen a dotar de conocimientos a cualquier máquina para que estas tengan comportamientos que serían propios de acciones exclusivamente humanas. Dentro de estos campos se encuentra Machine Learning, considerada la más importante ya que busca el desarrollo propio de la inteligencia artificial para que las maquinas aprendan por si solas a partir de programas con información que no siempre se encuentra estructurada o etiquetada (Figura 4.) (Cioffi, R., et al., 2020)

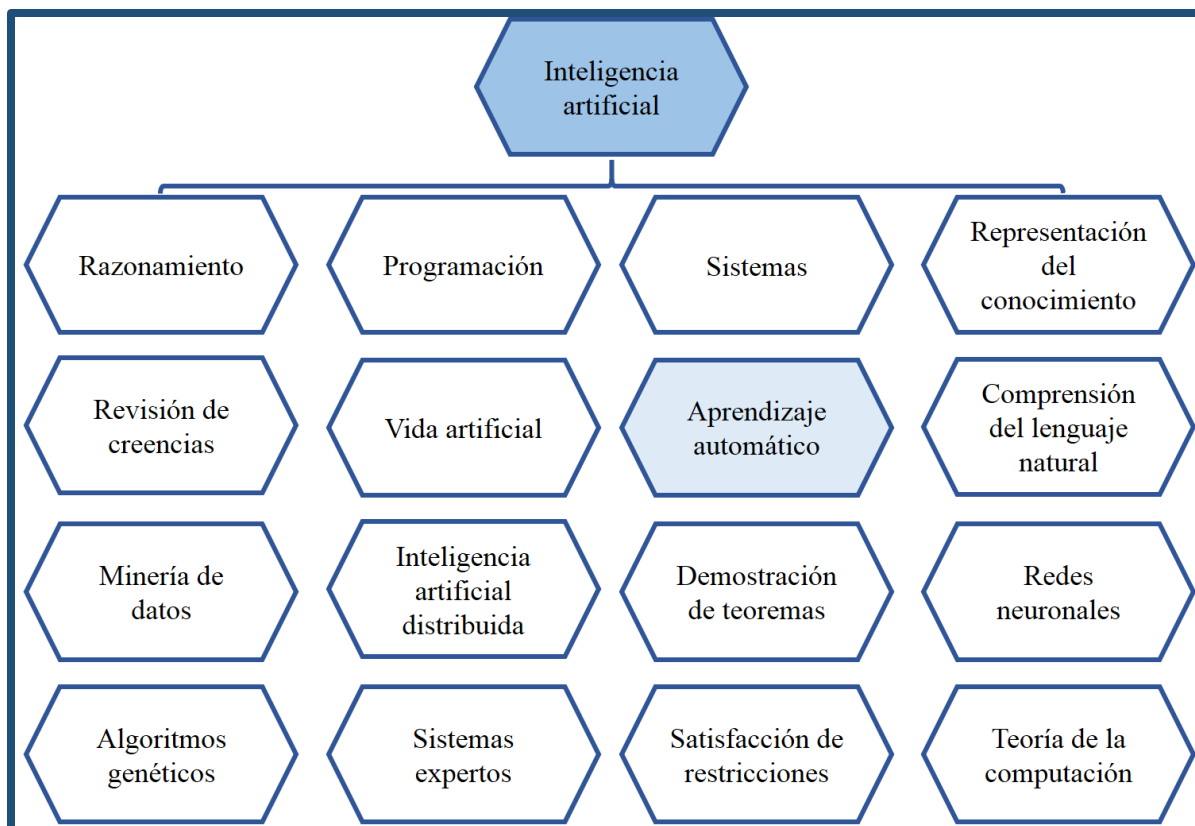


Figura 4. División de los campos de la inteligencia artificial (Elaboración propia con información de (Elaboración propia con información de Cioffi, R., et al., 2020)

La inteligencia artificial dedicada al desarrollo y aprendizaje de máquinas utiliza una de las tecnologías de la industria 4.0 para su funcionamiento (Figura 5)., se trata del big data o análisis de volúmenes grandes de datos en conjunto con su campo propio, Data Mining o minería de datos para poner organizadamente la base de datos, se encargan de buscar y encontrar relaciones en la información y analizarla para arrojar respuestas a preguntas no necesariamente existentes. Machine Learning es una rama de la inteligencia artificial que se encarga del aprendizaje de la maquina en particular a través de modelos y Deep Learning por su parte es la derivación de Machine Learning como modelo o algoritmo empleado para el aprendizaje

supervisado en el que se utilizan redes neuronales funcionales para crear aspectos similares a los del cerebro (Woschank, M., et al., 2020).

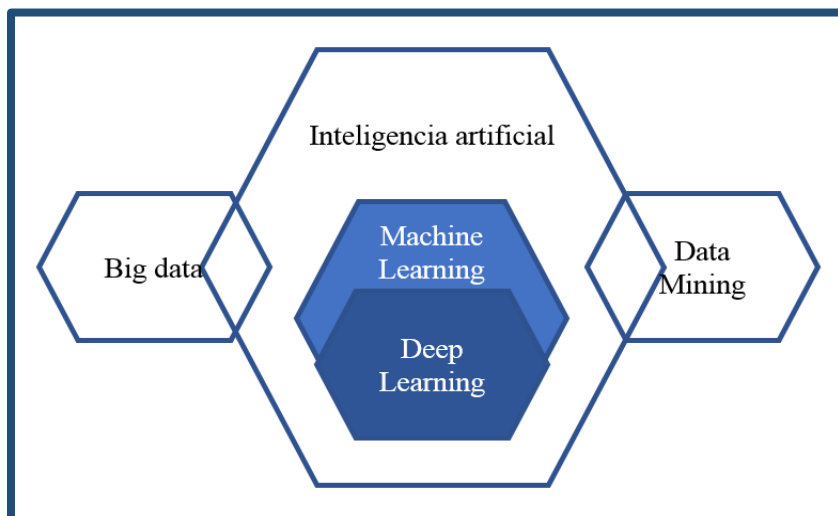


Figura 5. Relación de tecnologías de la industria 4.0 con la inteligencia artificial y sus ramas (Elaboración propia con información de Woschank, M., et al., 2020)

2.2 Machine Learning

Machine Learning se presenta cuando un sistema aprende a ejecutar actividades con la ayuda de ejemplos o a partir de prueba y error. En la actualidad existen elevados registros de prosperidad o avance en cuanto a su implementación en diversos sectores empresariales. Se trata de un modelo o programa que permite a las máquinas la capacidad de identificar conceptos o sucesos que se repitan dentro de grandes cantidades de datos y de esta manera puedan ejecutar predicciones en poco tiempo. Permite la automatización en los procesos debido a que supera la capacidad humana y tiene la facultad de resolver diligencias o actividades que antes solo estaban destinadas a resolverse por humanos. (Janiesch, C., et al., 2021).

Dentro de su sistema de funcionamiento, Machine Learning utiliza como entradas o inicio, una base de datos, y posibles salidas que ya se esperan como resultado de su

implementación, esta información se traslada a hardware o interfaces que ayuden a almacenar grandes cantidades de datos para que próximamente esta, sea la herramienta principal de la creación del modelo o programa ML, dotado de capacidades eficientes para resolver problemas o emitir informes que ayuden con la toma de decisiones. Este proceso se divide en tres etapas, recopilación de datos, almacenamiento y administración de la información, y por último creación del programa capaz de analizar los datos de manera rápida y eficiente.

Por otro lado, el funcionamiento del programa que cuenta con tres factores elementales;

1. Rendimiento, donde se evalúa el impacto de las predicciones que se emiten por el programa. 2. Actividades, que sirve para analizar qué acciones se deben tomar para realizar el análisis y verificar que la ruta sea correcta para no evadir ningún tipo de dato que sea importante. finalmente, el 3. programa puesto en marcha, en el que ya se conoce el flujo de pasos que se llevaran a cabo para evitar la omisión de técnicas que ayuden a optimizar los procesos. (figura 6).

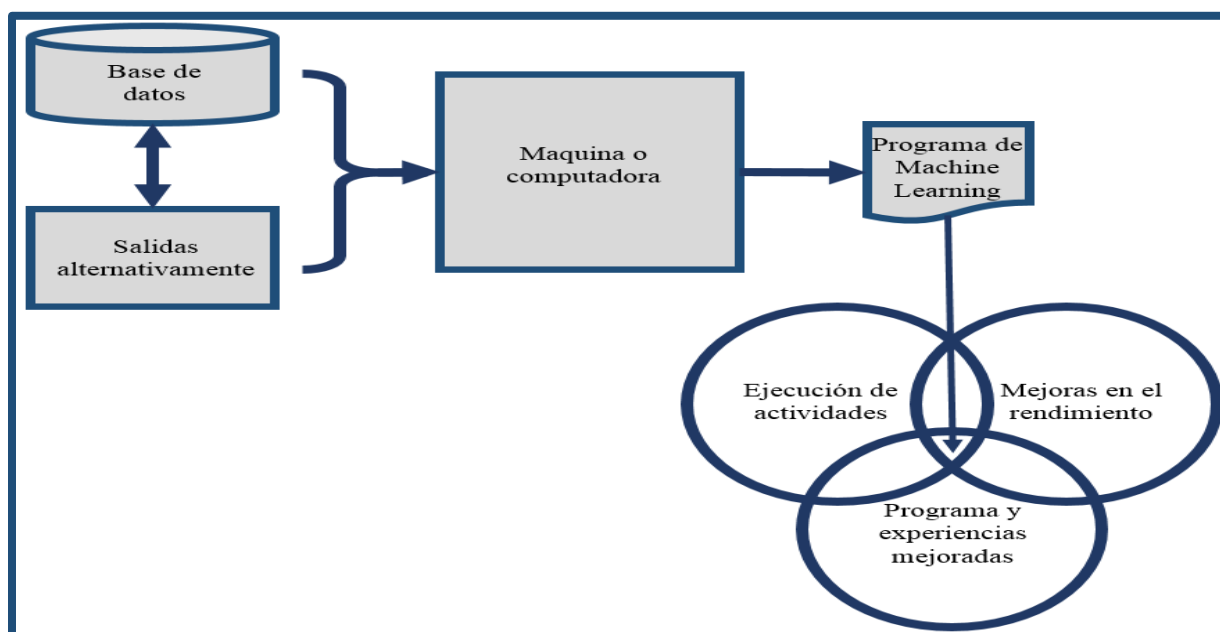


Figura 6. Patrón que rige al funcionamiento de Machine Learning. (Elaboración propia con información de (Elaboración propia con información de Sarkar, D., et al., 2018)

Y aunque parece fácil la creación de un modelo o programa Machine Learning, resulta que este se basa en diferentes disciplinas científicas e informáticas para su programación (figura 7.), ya que utiliza varias herramientas para generar métodos, sistemas y procedimientos, que dirijan a una mejor solución. Dentro de un mundo nuevo direccionado a la ciencia de datos, se encuentra la disciplina dominante, Inteligencia artificial, se trata de la cabeza de Machine Learning, el concepto reflejado en una especialización de inteligencia artificial. Deep Learning, el subcampo de Machine Learning que, de forma generalizada, se dota de experiencias a través de la adquisición de más datos, es decir, genera técnicas y conocimiento a partir de una introducción profunda a todos los posibles conjuntos de datos existentes.

Las matemáticas por su parte será el campo que se ocupa del estudio de las cantidades, es decir, determinara el funcionamiento de los sistemas numéricos y lógicos. las estadísticas, el factor más importante dentro de los análisis, pues está destinado a organizar la información y conseguir que esta sea clara concisa para que se pueda interpretar claramente y por ultimo las ciencias de la computación, es la herramienta que ayudara a entender el punto principal para la implementación del modelo, es decir, comprender el comportamiento de las computadoras, pues no siempre es posible ejecutar todas las expectativas que se tienen en un solo software y con un solo algoritmo, a veces es imprescindible la creación de nuevos sistemas para resolver aspectos que no solo un sistema contiene.

Esto conlleva el conocimiento avanzado acerca del entorno del procesamiento natural del lenguaje y el procesamiento de los datos. Son técnicas que visualizan un funcionamiento en conjunto para generar métodos que ayuden a extraer conocimiento a partir de patrones encontrados dentro del análisis.

Quizás, para la creación de un solo modelo o programa no sea necesario utilizar todas estas materias, sin embargo, dentro del proceso principal, son herramientas importantes para facilitar su diseño y construcción.

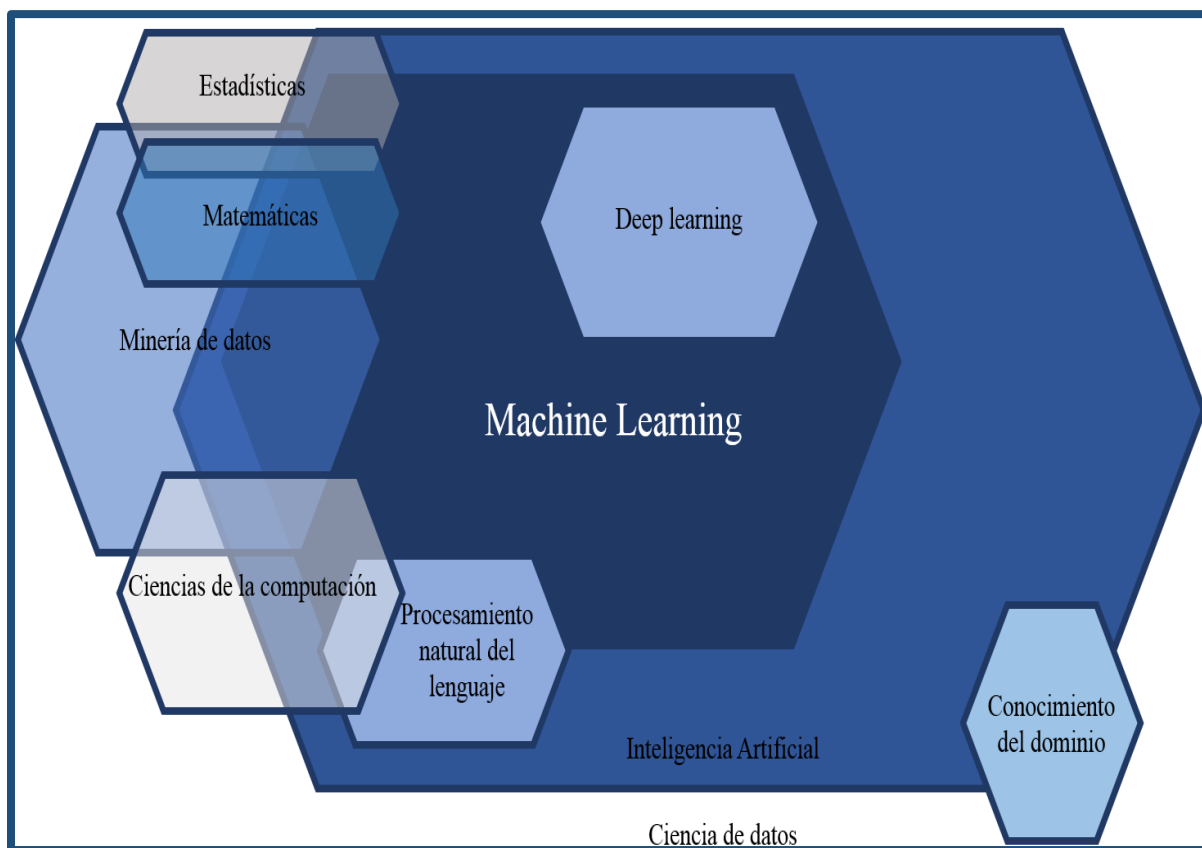


Figura 7. Principales disciplinas que rigen en los procesos de creación de programas ML (Sarkar, D., et al., 2018)

2.2.1 Tipos de algoritmos Machine Learning

Esta disciplina se muestra como una herramienta eficiente para la optimización de procesos, que se divide en diferentes tipos de algoritmos que ayudan en la resolución de problemas con propiedades específicas. (Figura 8.). Es decir, dependiendo de las características o limitantes del problema presentado, habrá un método que se apegue a la mayoría de estas para

dar predicción a las posibles soluciones. Existen 4 tipos más recurrentes en su aplicación, sin embargo, se actualizan según las necesidades del problema y se crean nuevos continuamente. Cada uno con peculiaridades que los hacen diferentes y precisos para ser utilizados en diferentes sistemas. El aprendizaje supervisado es un método que se utiliza cuando existen consecuencias previstas y los efectos son muy específicos en cuanto a la resolución del problema principal. Se encarga de resolver problemas conocidos a través de la experiencia, de acuerdo con el conjunto de datos de entrada que ya han sido etiquetados entrena los modelos o métodos para que se ejecuten actividades específicas. Este método dentro de su proceso tiene la característica de hacer uso de herramientas semejantes a los árboles de decisión para conseguir que el modelo aprenda de datos históricos y los resuelva como entradas que no se conocen para generar salidas. El aprendizaje no supervisado en general, trabaja con datos o información no estructurada y sin etiquetar, así entonces, debe encontrar un amañera de entender y aprender de esta información por sí mismo. Un híbrido de estos dos es el aprendizaje semi-supervisado, el cual contiene tanto datos no etiquetados como ya etiquetados, sin embargo, este se debe encargar de etiquetar los datos sin etiquetar a partir de los ya etiquetados para generar mayor conocimiento encontrando relaciones y patrones entre ellos y por último el aprendizaje por refuerzo en el que se adquiere conocimiento a partir de experiencias.

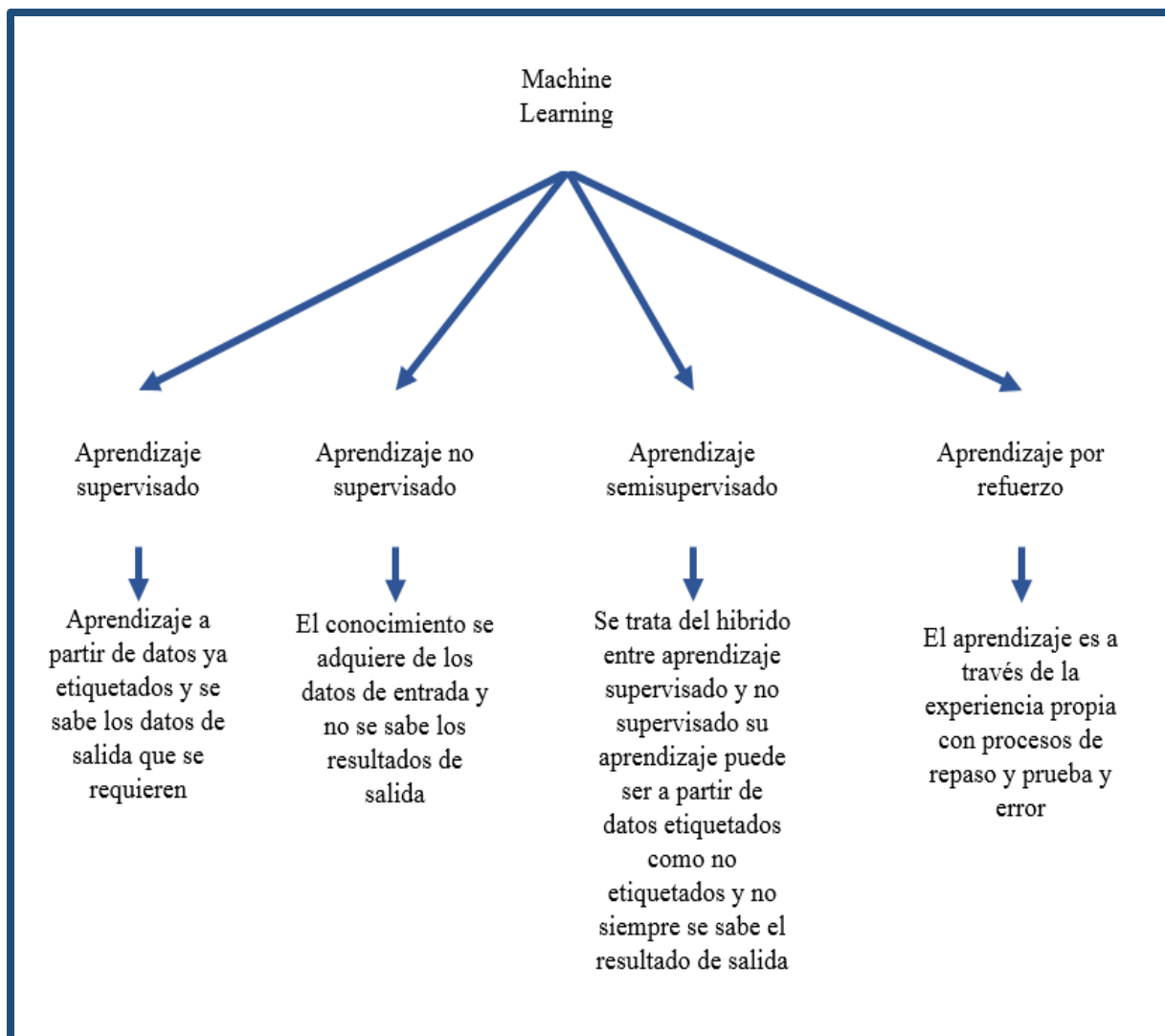


Figura 8. Métodos de aprendizaje automático y sus diferencias (Elaboración propia con información de Sarker, I., 2021)

Ahora bien, el aprendizaje supervisado trabaja con datos que ya se conocen o que están etiquetados que funcionan como entrenamiento, toma muestras de estos para aprender y en el futuro poder emitir predicciones. Este proceso (Figura 9.) se da para saber si el algoritmo que se creó es rentable en la resolución del problema y poder ser implementado y si no es así, se vuelve a crear otro modelo de algoritmo para ser probado. El entrenamiento, no es más que la combinación de

datos históricos que funcionan como conocimiento previo, y su transición a un modelo en el que los datos se relacionaron como conjunto similar.

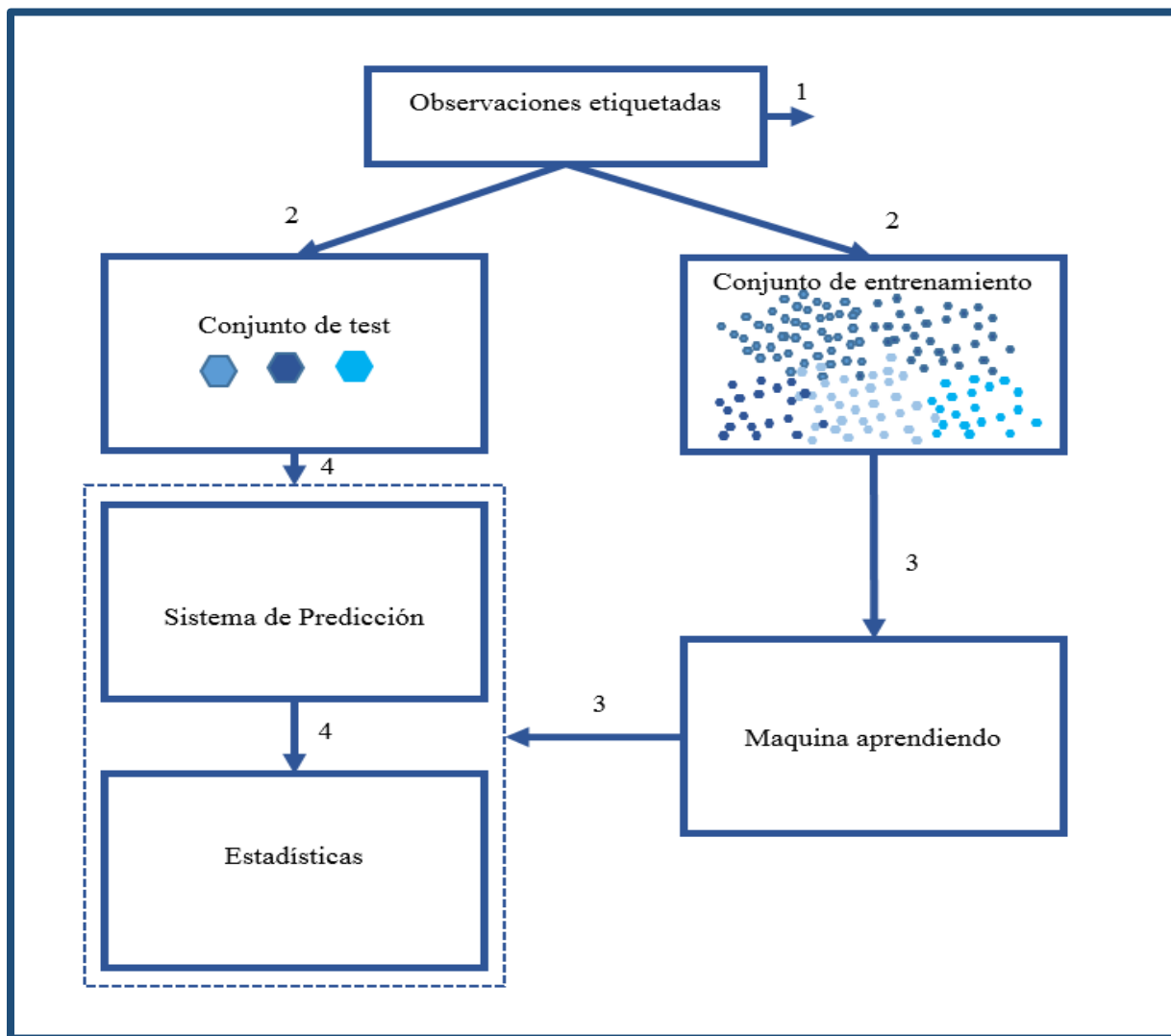


Figura 9. Proceso del modelo de aprendizaje supervisado (Elaboración propia con información de Sarker, 2021)

El aprendizaje supervisado se divide en diferentes categorías, existen 5 más utilizadas y cada una se dirige a una solución diferente dependientes de la necesidad que se tenga, se muestran cada una de estas y sus características más importantes (Figura 10.)

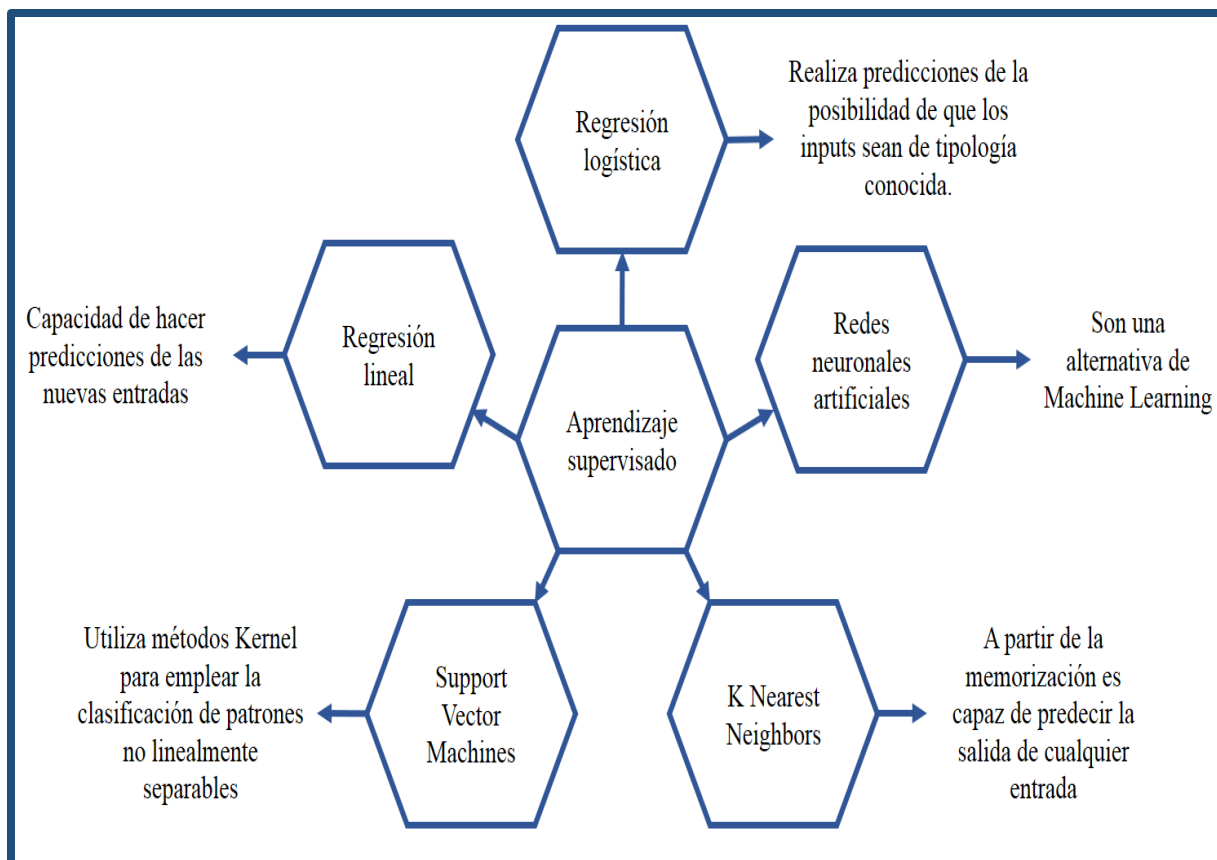


Figura 10. Principales algoritmos del aprendizaje supervisado (Elaboración propia con información del Misilmani., et al., 2019)

Cada Algoritmo contiene sucesiones propias para especificar los procesos de previsiones futuras y en algunos casos estos se pueden utilizar en conjunto para obtener más salidas de información. durante la clasificación se encuentra la regresión que sirve en la estimación de valores numéricos. Existen varios tipos de regresión (figura 11.), los principales son:

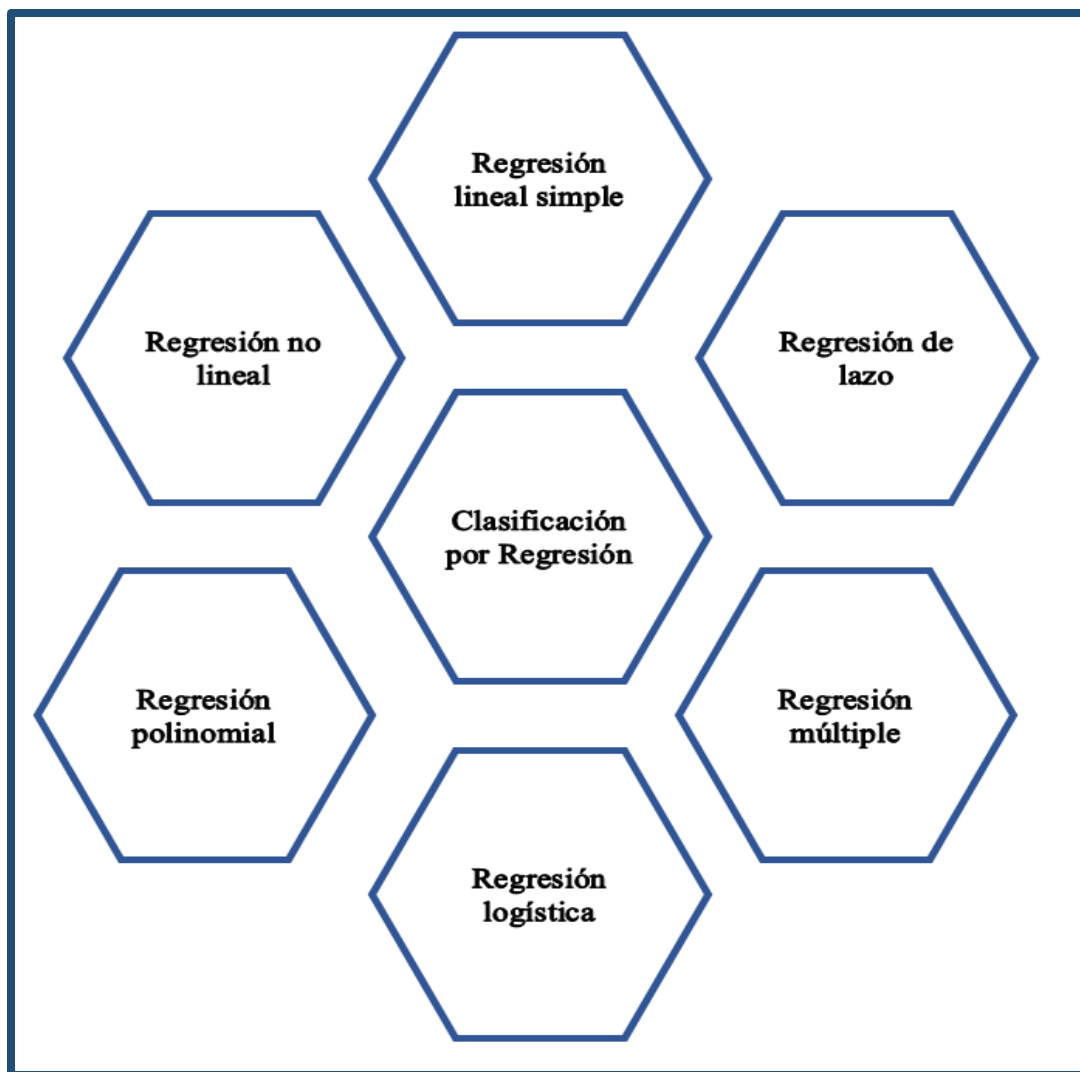


Figura 11. Tipología de la regresión. (Elaboración propia con información de Sarkar et al., 2018)

En la figura 12 se muestra el funcionamiento de 3 tipos de regresión, la regresión logística que emite predicciones a partir de una variable etiquetada a partir de una o más predicciones pasadas. La regresión lineal simple que emite datos cuantitativos a partir de datos cualitativos y la regresión polinomial que emite respuestas cualitativas a partir de datos cuantitativos encontrando relaciones semejantes entre ellos para poder clasificarlos.

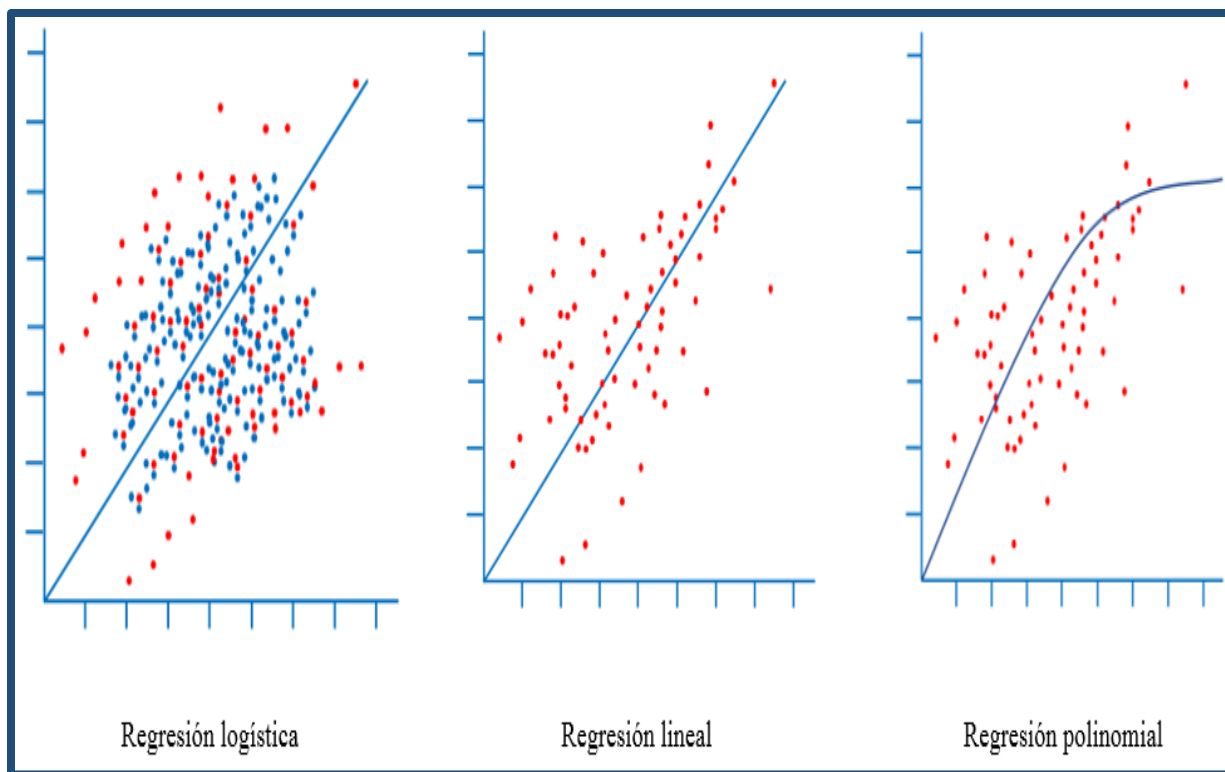


Figura 12. Comportamiento de los datos en función a tres tipos de regresión (Elaboración propia con información de Christodoulou, E., et al., 2019)

Otra forma de clasificación de aprendizaje supervisado son las redes neuronales artificiales, se trata de programas informáticos con estructuras y funcionamiento similares a las neuronas humanas o biológicas, están compuestas por capas iniciales o de entradas, finales o salidas, y medias u ocultas, estas a su vez se relacionan por conexiones como se muestra en la figura 13.

Existe una forma de estructura principal entre las que se encuentran 3 partes del diagrama, el inicio está constituido por las entradas o los posibles resultados que se desean obtener para la salida. En este se introducen datos que ya se conocen o son etiquetados. La segunda parte es el medio o los datos ocultos en los que se pueden encontrar relaciones, normalmente una red neuronal, contiene esta parte en su constitución ya que son atributos que, si bien no sirven como

salida, obstruyen el flujo del análisis. Y el final en donde se crea un umbral para la constitución del perceptrón. El cual nos ayudara con la toma de decisiones que a su vez es funcional como otro dato de entrada para otra red neuronal en solución a la respuesta de un posible problema nuevo.

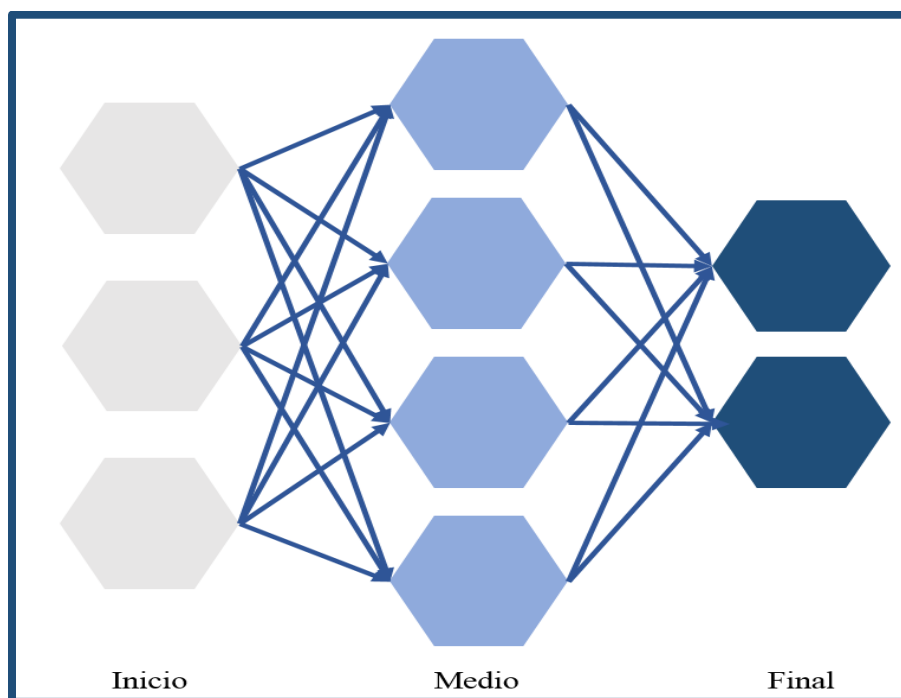


Figura 13. Estructura general de una red neuronal artificial (Sarkar et al., 2018)

Las máquinas de vector soporte son otro tipo de algoritmos entre los que se buscan métodos de relación en un conjunto de datos de entrenamiento (Figura 14).

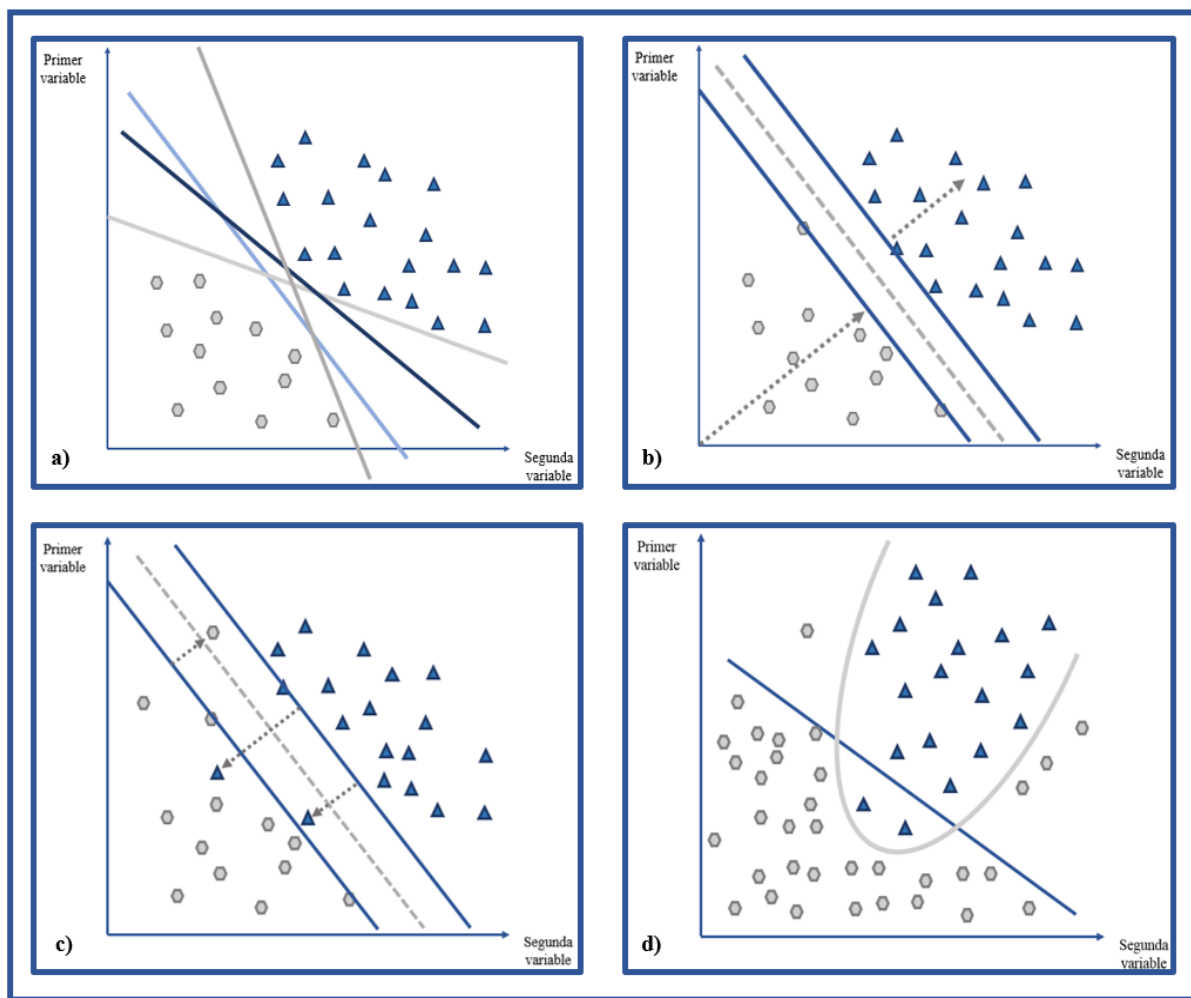


Figura 14. Funcionamiento de las maquinas vector soporte (Elaboración propia con información de Sarkar, D., et al., 2018)

El aprendizaje no supervisado es aplicable cuando no hay concurrencia en la existencia de datos ya etiquetados, y la posibilidad de etiquetarlos de forma manual, se limita a partir de los recursos económicos que esto genera. La forma de funcionamiento de este método se representa en la figura 15. Se basa en el suministro de elementos de entrada al algoritmo, estos simplemente son clasificados y agrupados por el reconocimiento de datos en común de las características generales de cada uno, las cuales deben ser aportadas de manera manual, ya que este agrupamiento, no es del todo automático o intuitivo.

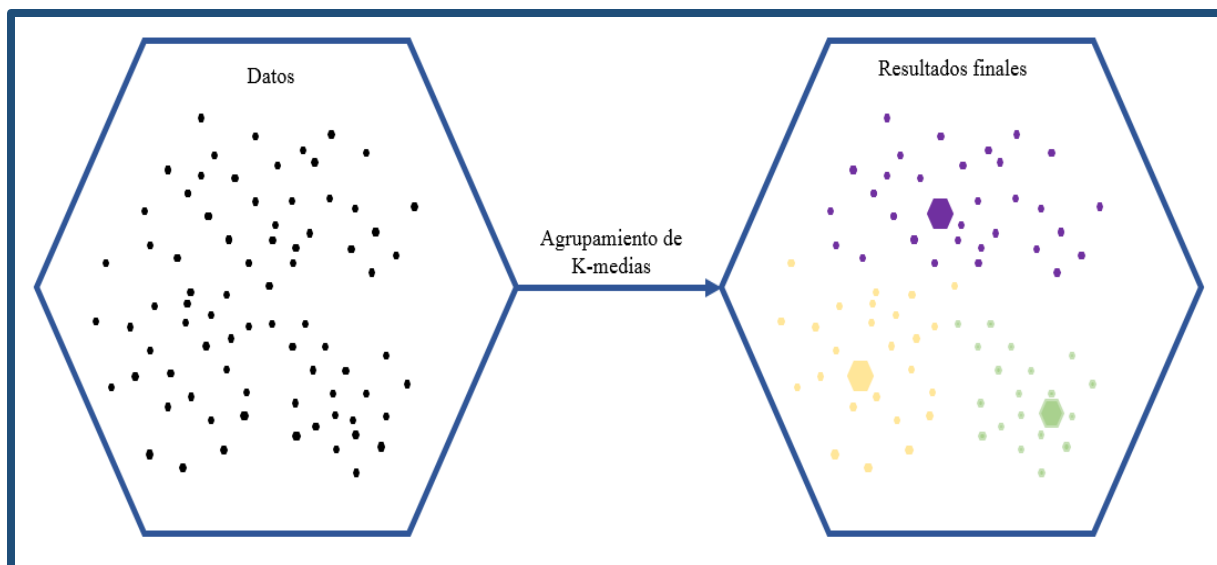


Figura 15. Formación de grupos de elementos en aprendizaje no supervisado (Elaboración propia con información de Sarkar, D., et al., 2018)

Por su parte, el aprendizaje no supervisado se clasifica en dos subdivisiones las cuales ayudan a que el proceso de agrupamiento sea concreto y omita grandes cantidades de datos muy precisos, ya que, para este método, no es necesario el estudio profundizado de cada elemento.

(figura 16).

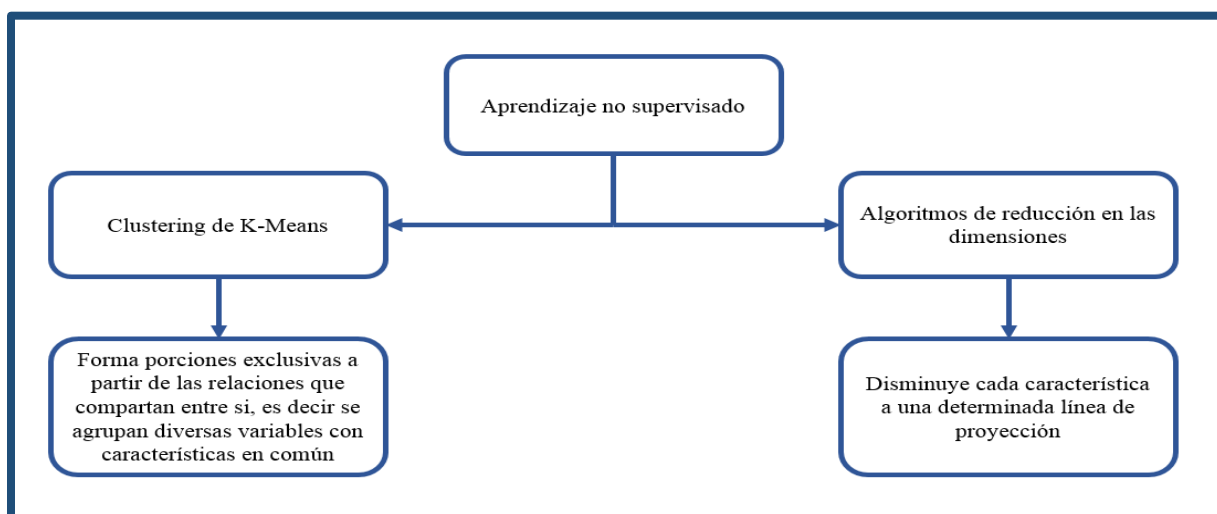


Figura 16. Ramas del aprendizaje no supervisado (Naous, T., 2019)

El aprendizaje semi-supervisado funciona como mezcla de los aprendizajes supervisado y no supervisado, es decir, existe la capacidad de etiquetar elementos no etiquetados con ayuda de pequeñas cantidades de elementos ya etiquetados como se muestra en la figura 17.

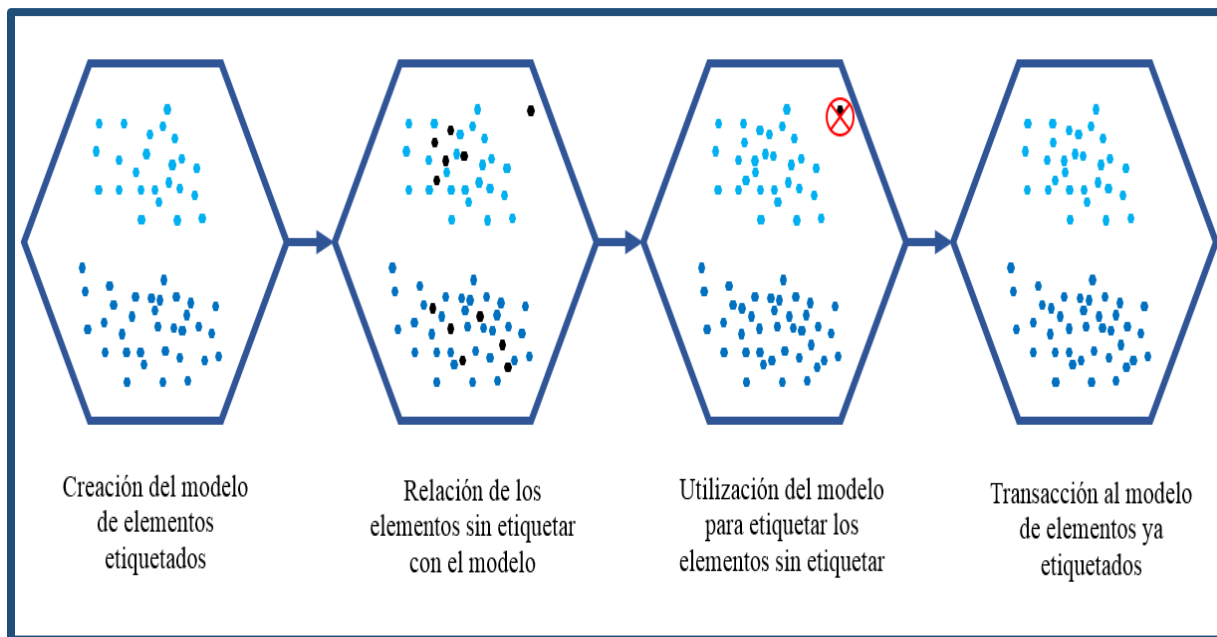


Figura 17. Agrupamiento de elementos de manera híbrida del método semi-supervisado (Elaboración propia con información de Sarker, I., 2021)

Los métodos del aprendizaje semi-supervisado funcionan de manera parcial, con información poco etiquetada o datos generalizados del entorno. De esta manera crea la capacidad de entrenarse autónomamente. Cuando dominan los elementos ya etiquetados, este método puede ser ejecutado como aprendizaje supervisado. Existen dos tipos de algoritmos en los que se basa; las redes adversarias generativas

El aprendizaje por refuerzo genera estrategias basadas en la experimentación de los datos de forma automática, las máquinas aprenden a partir de los datos sin recibir órdenes

estructuradas y es así como se genera un entorno ideal para la futura optimización de fallas (Naous, T., 2019)

Este tipo de aprendizaje se aísla de los aprendizajes supervisado y no supervisado, debido a que no cuenta con datos etiquetados en su entrada para considerarse supervisado, la manera en que se alimenta de información se encuentra en la recompensa basada en la experimentación. En la figura 18., se muestra el comportamiento que tiene la máquina para aprender, es decir, se encuentra en un entorno donde existen cuatro agentes diferentes; el primer paso es la observación la máquina es capaz de tomar la decisión de clasificar los agentes, el siguiente paso se basa en la experimentación a través del acto que ejecuto la máquina entiende que existen cuatro agentes con características diferentes y su sanción está en que no clasifico los agentes por completo y le faltan muchos por clasificar, ahora la máquina actualiza la información en su algoritmo y aprende que existen 5 componentes de cada tipo de agente, sin embargo, debe experimentar sanciones o recompensas para dar la mejor propuesta a la solución del problema. (Sarker, I., 2021)

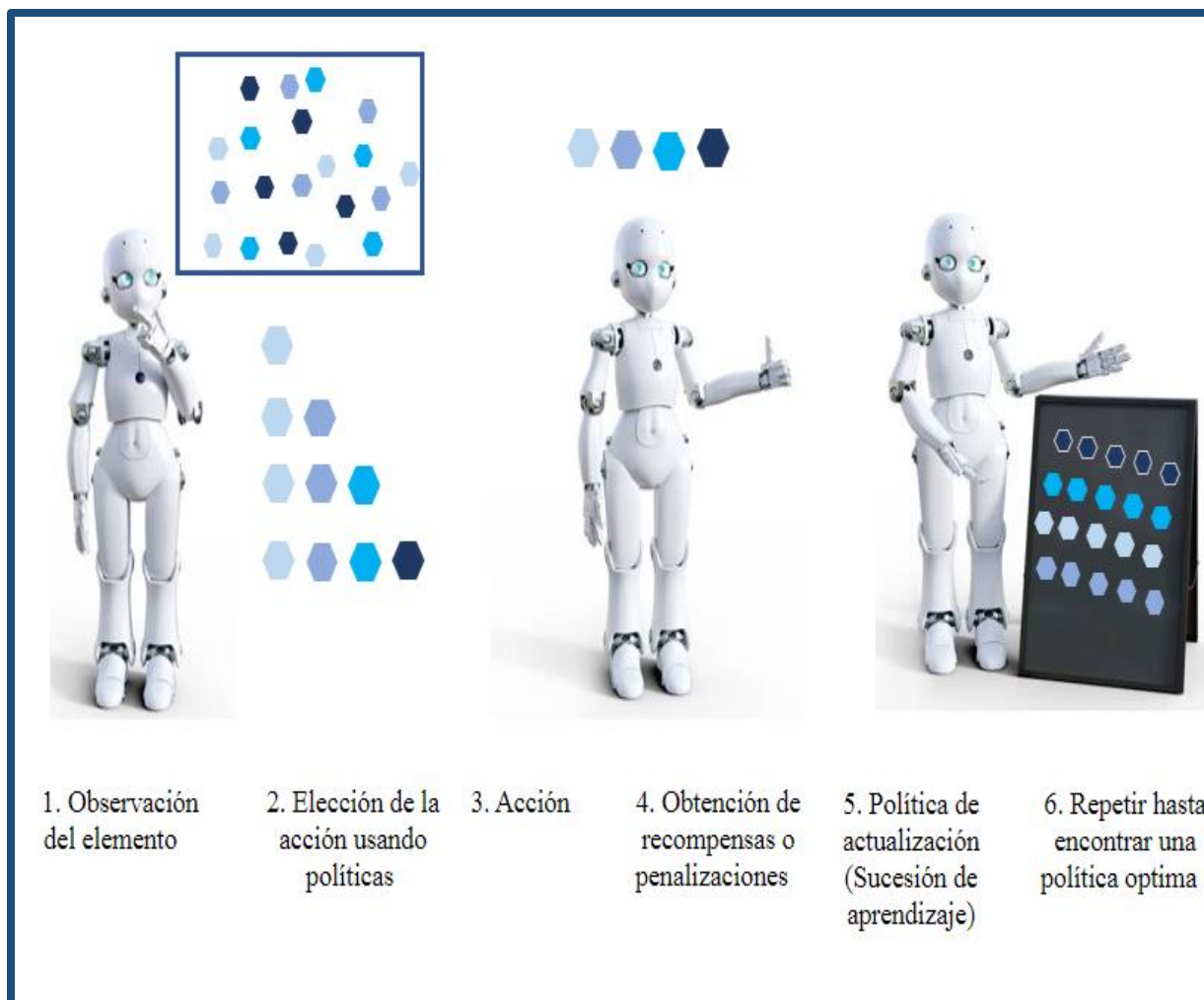


Figura 18. Sucesión de pasos del aprendizaje por refuerzo (Elaboración propia con información de Sarker, I., 2021)

El aprendizaje por refuerzo se convierte en un ciclo que se ocupa de indagar como es que los agentes deben funcionar para ejecutar actos automáticamente en cualquier entorno con el fin de que exista mayor incidencia en los premios o recompensas que se presenten como se muestra en la figura 19. La sucesión consta de 5 elementos por los que pasa la máquina para obtener información propia y ejecutar acciones de manera autónoma. (Mahesh, B., 2020)

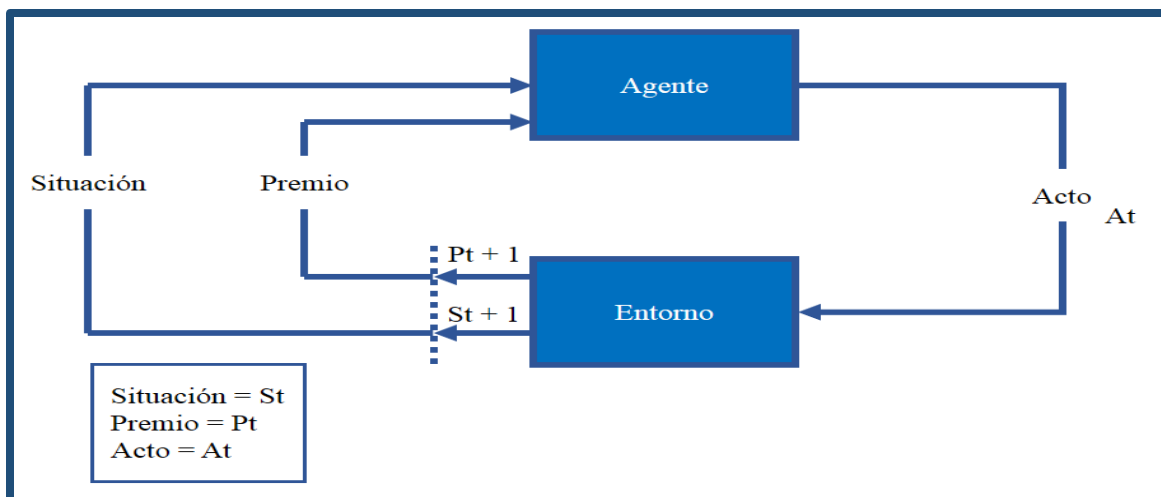


Figura 19. Flujo general (elaboración propia con información de Mahesh, B., 2020)

Existe un método llamado aprendizaje por transferencia para este último modelo de aprendizaje, se muestra una combinación de algunas características del funcionamiento de los modelos anteriores, su objetivo es relacionar elementos de métodos que anteriormente ya tenía aprendidos, con elementos de métodos nuevos, así entonces crea un modelo con gran cantidad de elementos a partir de métodos con pocos elementos. Es decir, aprende a partir de la relación encontrada en datos guardados y datos nuevos. En la figura 20 se muestra la manera generalizada del funcionamiento del aprendizaje por transferencia, relaciona elementos encontrados en 2 métodos, para formar conceptos que contengan características de ambos conjuntos. Reutiliza un método propio o previamente memorizado y un método nuevo pre entrenado (Weiss, K., et al., 2016)

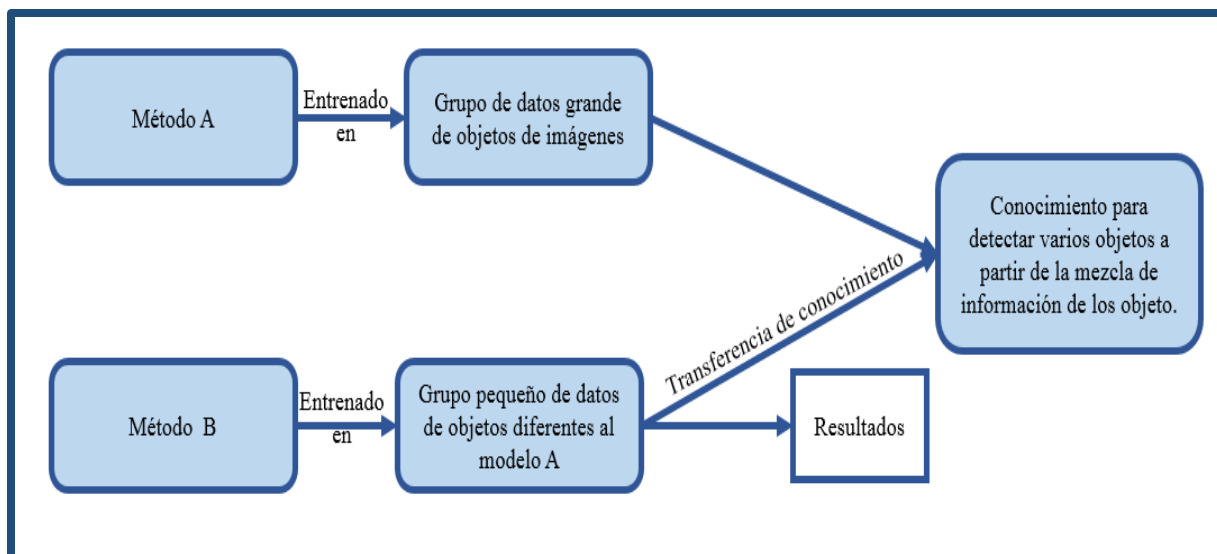


Figura 20. Proceso para llegar al aprendizaje por transferencia (Elaboración propia con información de Weiss, K., et al., 2016)

2.2.2 Lenguajes y software de programación para Machine Learning

En la actualidad, existe una gran variedad de lenguajes de programación que cuentan con la capacidad de ejecutar implementaciones de los estándares de Machine Learning, se trata de lenguajes de código abierto que cuentan con librerías destinadas al desarrollo de la ciencia de los datos, entre las que destacan, Python, R y R studio, MATLAB, Julia, etc. Sin embargo, con el paso del tiempo, se crean nuevos softwares que proporcionan las mismas herramientas pero que posibilitan el aumento de las capacidades. Por su parte, Python es una plataforma de lenguaje de programación, que destaca por ser muy fácil en su desarrollo y aprendizaje para la elaboración de programas y modelos, actualmente, ha sido muy utilizada para desarrollar implementaciones de la ciencia de los datos, cuenta con una serie de librerías como SCIKIT-LEARN y SCIPY en las que se encuentran PANDAS y NUMPY (Figura 21.).

SCIKIT-LEARN encontrada como la principal librería enfocada a generar sistemas de programación a partir de los datos, con el fin de emitir predicciones, trabaja con áreas como clasificación por regresiones, arboles de decisión, máquinas de vector soporte.

Numpy que, aunque también se trata de la ciencia de datos que se enfoca en los cálculos de los valores numéricos dentro del conjunto de datos y por último Pandas que a partir de su aplicación se obtiene la manipulación y análisis de datos.

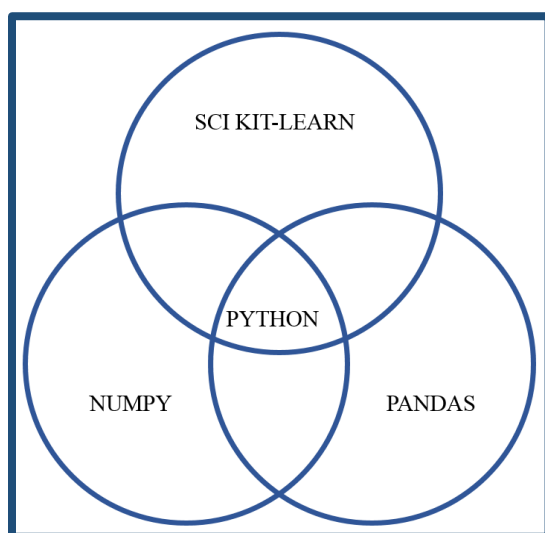


Figura 21. Principales librerías de Python enfocadas a Machine Learning (Elaboración propia con información de Sarkar, D., et al., 2018)

Le sigue el lenguaje de programación R, cuenta con herramientas para desarrollar algoritmos a partir de análisis estadísticos. Se trata de otro tipo de lenguaje que cuenta con servicios encaminados a Machine Learning como XG boost que administra los datos en árboles de decisión y los clasifica de acuerdo a las expectativas del usuario. R studio es considerado como una derivación de R, generado para facilitar el entendimiento de los códigos.

Por otro lado, MATLAB conocido por ser un software integrado por complementos costosos pero eficaces en su funcionamiento, pues está diseñado para ser un algoritmo certero y duradero. Matlab contiene una librería llamada Toolbox, (Figura 22.) con objetivos direccionados a Machine Learning en la cual se puede ejecutar procedimientos para la construcción de clústeres, estudios Anova, clasificaciones por regresión, etc.

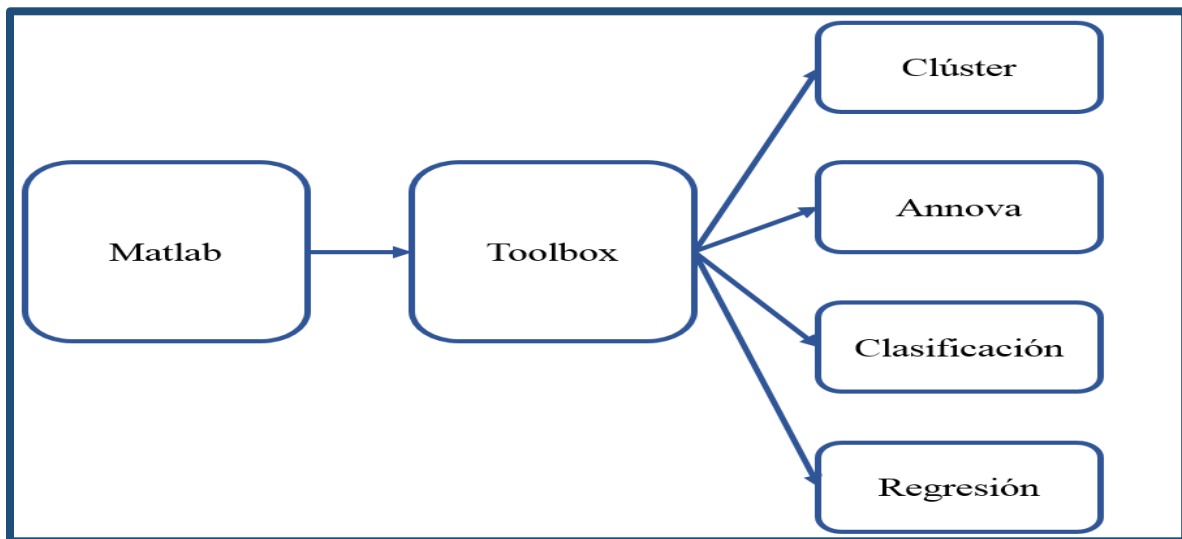


Figura 22. Matlab y Toolbox (Elaboración propia con información de Sarkar et al., 2018)

Y por último Julia, que se dirige al análisis numérico y en la actualidad, se encuentra desarrollando implementos encaminados al desarrollo de la ciencia de los datos, buscando una estabilidad en las principales aplicaciones de Machine Learning (Figura 23.) (Gao et al., 2020)

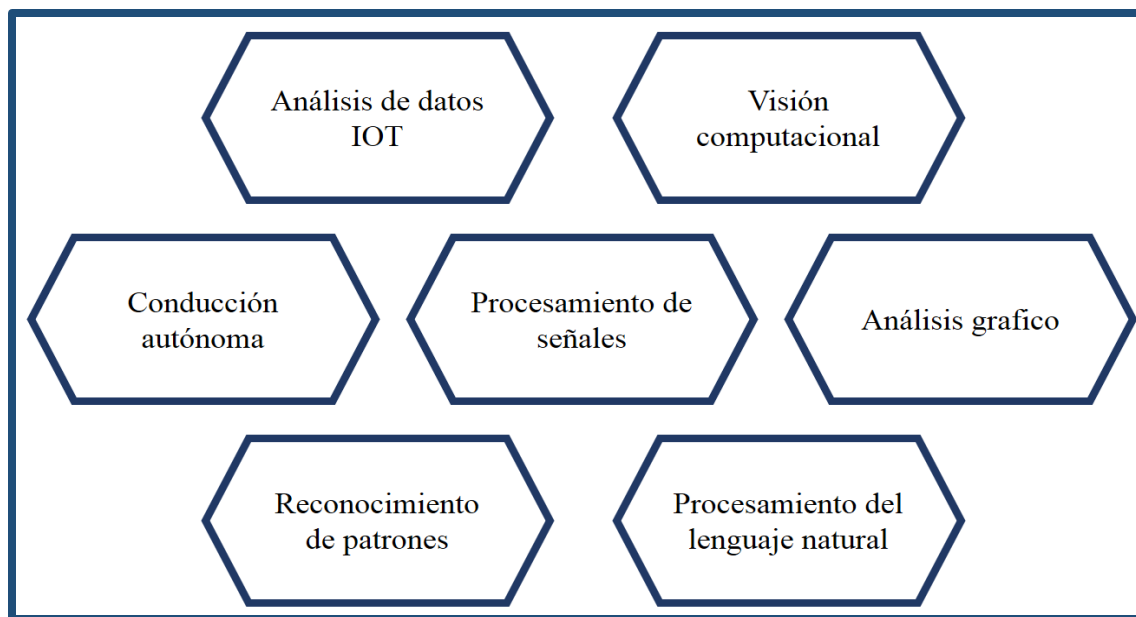


Figura 23. Elementos de aprendizaje automático a partir de Julia (Gao et al., 2020)

2.2.3 Sistemas de recolección de datos

Debido a que el desarrollo de modelos Machine Learning está orientado al tratamiento de grandes cantidades de datos, es de vital importancia la recolección exitosa de estos, por ello se crean métodos a través de diversas fuentes, con el objetivo de obtener un programa integro para poder resolver esta actividad. Los más utilizados dentro de las aplicaciones en la industria se encuentran SCADA como sistema de recolección específicamente en tiempos reales y MES como sistemas de digitalización de la información dentro de un sistema de producción. Por su parte SCADA se trata de un sistema que integra las siguientes 3 actividades primordialmente supervisión, control y adquisición de datos, en la que se encuentra una mezcla del uso entre un

hardware y un software (Figura 24.) en la que el hardware presenta cuatro componentes

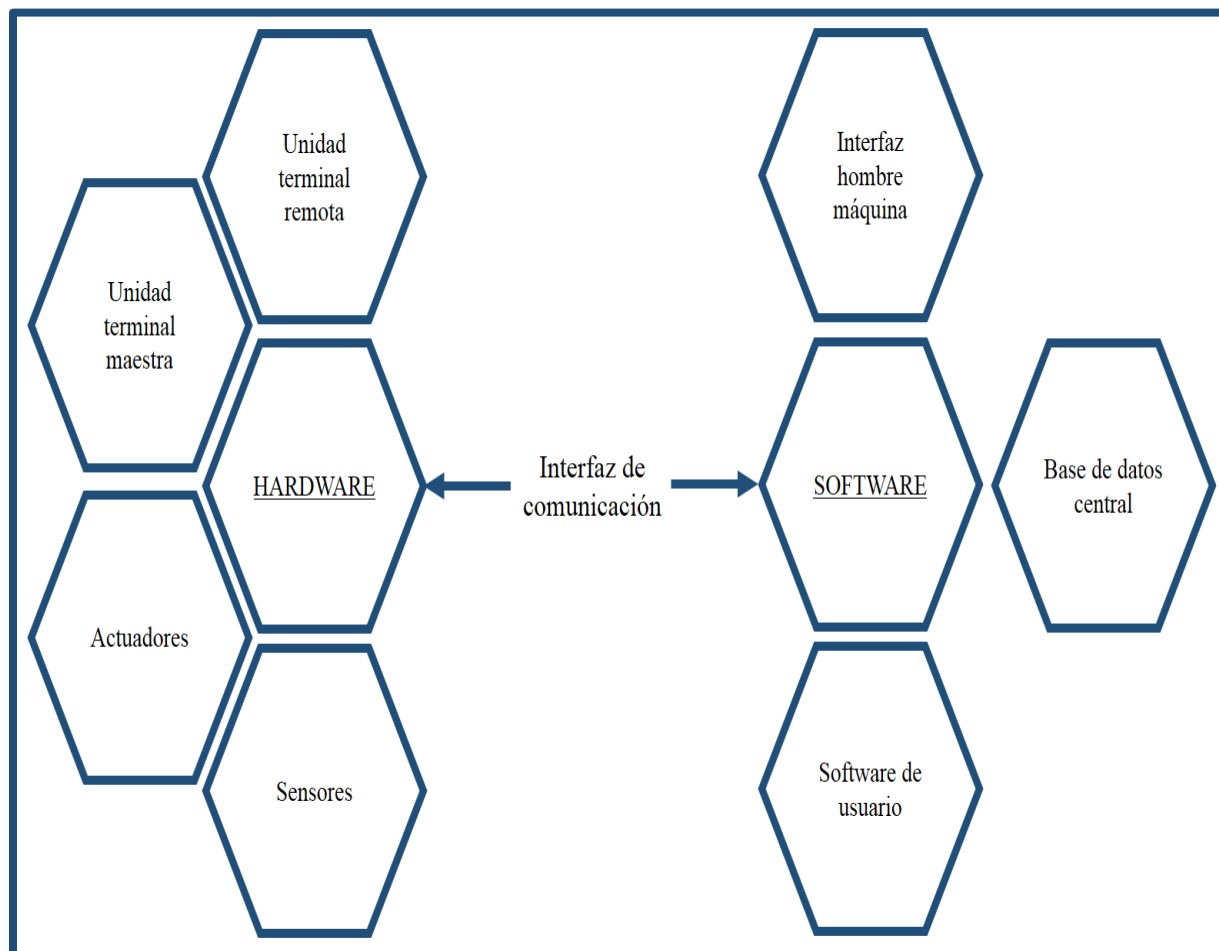


Figura 24. Relaciones de SCADA (Elaboración propia con información de Yaday, G., et al., 2021)

Existe un comportamiento muy parecido a la revolución industrial con respecto al desarrollo de la clasificación de los sistemas SCADA, pues con el paso del tiempo se desarrollaron diferentes sistemas con el fin darles solución a los problemas en las actividades de recolección (Tabla 3.)

Tabla 3. Tipología de SCADA (Elaboración propia con información de Yadav & Paul, 2021)

Categoría	Descripción
Sistemas SCADA monolíticos	Se trata de ciertos sistemas en los que su forma de mantenimiento está basada en la ejecución de actividades de manera autónoma y solitaria
Sistemas SCADA distribuidos	Están conectadas con redes de área local, es decir su conexión se genera a partir de pequeñas distancias para su alcance.
Sistemas SCADA en red	Hacen uso de la web ya que se encargan de proponer resoluciones a problemas que se presentan en vigorosos sistemas.
Cuarta generación	Sistema que incluye tecnologías en funcionamiento similar a la industria 4.0, y se encargan del manejo y tratamiento de la información.

Por último, Sistema de Ejecución de Fabricación (MES), que más que un sistema para la recolección de datos, funciona como evaluador de la calidad del sistema de la producción y ayuda a detectar las deficiencias que se presenta en tiempos reales. La información que este sistema emite, ayuda a los modelos de Machine Learning a obtener datos direccionados a la información futura que aún no ha sido prevista. (Chen & Voigt, 2020)

CAPÍTULO III.
REVISIÓN DE LA
LITERATURA

3.1 Artículos de revisión

En la tabla 4 se muestra un resumen de 30 artículos de divulgación científica, seleccionados y analizados sistemáticamente con el fin de examinar la información más actualizada de las implementaciones de modelos de Machine Learning desarrollados dentro del sector industrial manufacturero, como resultado de una herramienta eficaz para emitir predicciones en solución a una amplia variedad de problemas emitiendo información valiosa para una mejor toma de decisiones.

Tabla 4. Resumen de artículos (Elaboración propia)

Nombre	Proceso y herramientas	Resultados
<p>“Métodos de aprendizaje automático para una mejor predicción de la calidad del agua”</p> <p>País: Malasia</p> <p>(Ahmed, A., et al., 2019)</p>	<p>Se midió la calidad del agua en 7 pasos. MLP-ANN: red neuronal de perceptrón multicapa. ANFIS: sistema de inferencia neurodifuso adaptativo. La eliminación de las ondas de ruido. Evaluación del desempeño del modelo.</p> <p>Variables de entrada y procesamiento de datos.</p> <p>Criterios de parada. Diferentes escenarios. En los que se introdujeron 4 diferentes</p>	<p>A través del método prueba y error se determinó el uso óptimo de la red neuronal de perceptrón multicapa (MLP-ANN) ya que no existe un enfoque que indique cual o cuales modelados se deben usar. La propuesta del modelo dedicado a técnicas de eliminación de ruido con el fin de que los estudios realizados para la predicción fueran suficientemente precisos se</p>

Nombre	Proceso y herramientas	Resultados
	<p>enfoques de modelado (ANFIS, RBF-ANN, MLP-ANN, WDT-ANFIS) para el periodo de entrenamiento. Basándose en doce parámetros de entrada, se presentaron dos escenarios; el primero buscando predicción de la calidad del agua en todas las estaciones y el segundo basado en el valor del mismo parámetro existente en la estación anterior.</p>	<p>puso en marcha con datos de 2009 a 2010 y se concluyó que el escenario dos tuvo mejoras considerables de 0.5% al 5% superando al escenario 1 en todos los parámetros de la calidad del agua en todas las estaciones. Los datos fueron analizados con software estadístico, ya que disminuyeron la cantidad de datos con los modelados.</p>
<p>“Un estudio de caso industrial utilizando datos de vibración y Aprendizaje automático para predecir el estado de los activos” País: Alemania (Amihai, I., et al., 2018)</p>	<p>Se analizaron datos reales del funcionamiento de los activos de una industria con el objetivo de disminuir fallas en el monitoreo de las vibraciones de los activos. Utilizaron algoritmos de Random Forest para clasificar los datos provenientes del software Wimon Data Manager que se</p>	<p>Con ayuda de lenguajes de programación R y Python, se logró la construcción de un modelo contenedor de los algoritmos de Machine Learning para deliberar los KCI observados en los equipos. Todo esto con el fin de anticipar la predicción de las métricas enfocadas a las fallas.</p>

Nombre	Proceso y herramientas	Resultados
	<p>transfirieron a través de una puerta de enlace inalámbrica HART1 por los datos extraídos del sensor.</p>	
<p>“Sistema de mantenimiento predictivo para líneas de producción en manufactura: un enfoque de aprendizaje automático utilizando datos de IoT en tiempo real” País: Turquía (Ayvaz & Alpay., 2021)</p>	<p>Con ayuda de una interconexión creada para mantener comunicación entre todas las máquinas, los sensores envían todos los datos recopilados en tiempo real a un sistema y estos son analizados junto con datos históricos por modelos de Machine Learning para que exista una recopilación efectiva de señales provenientes de los fallos que existan en el funcionamiento de manera automática. Se utilizó Python y bibliotecas de Machine Learning, KERAS y SCI-KIT LEARN para</p>	<p>Integración de un sistema que informe sobre las posibles fallas que se tengan en un proceso antes de que ocurra, con el objetivo principal de que la producción no tenga paros, y por lo tanto se eviten los costos de tiempos perdidos. Es así que el sistema es capaz de generar informes sobre el estado de los equipos de los que depende que la línea o área funcione de manera correcta, y este informe es enviado a las interfaces WEB FLASK y WEB REST para que puedan ser analizadas como protocolos http.</p>

Nombre	Proceso y herramientas	Resultados
	desarrollar y evaluar los modelos de Machine Learning	
<p>“Estrategia de mantenimiento predictivo basada en aprendizaje automático: un enfoque de súper aprendizaje con redes neuronales profundas”</p> <p>País: Túnez</p> <p>(Butte, S., et al., 2018)</p>	<p>Disminución en pérdidas de tiempo y costos para los eventos de mantenimiento no planificados en las fábricas con ayuda de métodos de aprendizaje automático como modelos lineales generalizados, aprendizaje profundo utilizando redes neuronales, bosque aleatorio, y máquina de aumento de gradiente para regresión y clasificación evaluados con los datos del equipo.</p>	<p>Programación adecuada a partir del mantenimiento predictivo basado en Machine Learning, enfocado en la maquinaria en la que una de sus piezas debe ser remplazada antes de que comience a degradar. El modelo se basará en los datos destacables del ciclo de vida útil de dicha pieza y datos del equipo en general.</p>
<p>“Espectrómetro portátil de infrarrojo cercano y métodos de aprendizaje automático</p>	<p>El estudio se desarrolló en el software MATLAB en TOOLBOX con el objetivo principal de evaluar el potencial de un espectrómetro</p>	<p>Con ayuda de algoritmos de Machine Learning se determinó que el espectrómetro NIR portátil es una herramienta que ayuda ventajosamente en el</p>

Nombre	Proceso y herramientas	Resultados
<p>aplicados al monitoreo de múltiples etapas del proceso en la producción industrial de azúcar”</p> <p>País: Brasil (da Silva, B., et al., 2018)</p>	<p>NIR portátil para la producción de azúcar desde la recepción de la caña hasta el refinado del azúcar. Se realizó un estudio preliminar para definir el mejor procedimiento de adquisición espectral para el análisis de diferentes muestras industriales. Debido a las grandes diferencias de estas muestras, se utilizaron la transmitancia y la reluctancia difusa.</p>	<p>desarrollo de sensores para calcular la calidad en los procesos de fabricación de azúcar. El monitoreo puede ejecutarse en maneras más económicas y fáciles de entender, lo que genera toma de decisiones en cambios de la ejecución y el control del proceso.</p>
<p>“Supervisión de procesos automatizados en impresión 3D mediante aprendizaje automático supervisado”</p>	<p>Se construyó un código para automatizar el monitoreo en tiempo real del proceso de impresión con el lenguaje de programación Python en conjunto con la librería SCI-KIT-LEARN para poder de entrenar y accionar el algoritmo de aprendizaje supervisado que</p>	<p>Creación de un modelo de Machine Learning para medir la calidad de las impresiones generadas específicamente en la impresora Lulzbot Mini y el proceso de impresión 3D aplicado es Fused Deposition Modeling (FDM), con la propuesta del SVM que tuvo el</p>

Nombre	Proceso y herramientas	Resultados
País: Estados Unidos (Delli & Chang, 2018)	se utilizó para la clasificación y regresión de los datos que se propuso.	objetivo de etiquetar los parámetros de las características de las piezas durante el proceso de impresión. Este método cumple con la capacidad de detectar defectos en el momento adecuado para tomar la decisión de que se siga con dicho proceso o se detenga.
“Métodos guiados por aprendizaje automático en la construcción de un modelo de templabilidad de composición química para acero resistente al desgaste” País: China	En un estudio de la resistencia al desgaste del acero, se adquirieron datos de los componentes de los materiales (Si, Cr, Ni y Mo). Los cuales fueron validados en el software J-Mat-Pro, se utilizaron para entrenar los modelos de Machine Learning (SVR, regresión lineal, K-vecinos más cercanos, RF, GBR Y ANN con el fin de medir la dureza o	La librería SCI-KIT LEARN de Python con la técnica de validación cruzada fue la plataforma utilizada para analizar los datos de prueba y entrenamiento. Se revisó el ajuste del modelo de Machine Learning para verificar el balanceo para poder comparar los grupos de datos de prueba y entrenamiento. Resultando que la red neuronal artificial, fue el

Nombre	Proceso y herramientas	Resultados
(Dong, G., et al., 2020)	resistencia a la penetración de los metales.	método apropiado para la predicción debido a su obtención del mejor rendimiento en la precisión de las predicciones de los parámetros resistentes al desgaste.
<p>“Enfoque de aprendizaje automático para la predicción de la rugosidad de la superficie utilizando características estadísticas de la señal de vibración adquirida en el torneado”</p> <p>País: India</p> <p>(Elangovan, M., et al., 2015)</p>	<p>Se pretende realizar un estudio sobre la rugosidad de la superficie, los datos informan los estándares recurrentes en cortes y la vibración que surgen en el proceso, la rugosidad de la superficie también es medible y de suma importancia la toma de muestras para generar el modelo, dicho modelo se enfoca en reducir las dimensiones de la información para crear reportes enfocados al proceso.</p>	<p>Se realizaron predicciones de acuerdo con los parámetros de la rugosidad de la superficie de torneado, a través de algoritmos Machine Learning (regresión múltiple). Los resultados estadísticos permiten visualizar el tiempo de las señales retomadas de las muestras. La manera en que Machine Learning ejecuta las señales es canalizando los datos a una clasificación que permite un estudio de los</p>

Nombre	Proceso y herramientas	Resultados
		patrones que se generan en las incidencias de las señales.
<p>“Aplicación de Machine Learning para apoyar la planificación de la producción de una industria alimentaria en el contexto de generación de residuos bajo incertidumbre”</p> <p>País: España (Garre, A., et al., 2020)</p>	<p>Con ayuda del software SCADA se registró información de lotes en un periodo de 10 meses que fueron analizados por un modelo de regresión implementado en el software R. Buscando establecer estándares que permitan conocer cuáles son los parámetros en los que los alimentos comienzan a perder propiedades para su consumo, como la caducidad, traducida en el tiempo que dura el producto para ser consumido.</p>	<p>Se generaron modelo implementados en R para predecir el ciclo de vida útil o servible de los alimentos, ya que representa uno de los mayores problemas dentro de las industrias alimenticias. Para esto se analizan datos de una industria en específico sobre las características con las que sus productos se distribuyen. En este caso los líquidos a base de frutas presentan mayor desequilibrio en los requisitos, ya que varían dependiendo de los requerimientos del cliente.</p>
<p>“Predicción de ventas basada en Machine Learning”</p>	<p>Se comparó el rendimiento de dos modelos lineales, tres modelos de aprendizaje</p>	<p>Dentro de la comparación con ayuda de métricas RMSE y tiempo de capacitación, se</p>

Nombre	Proceso y herramientas	Resultados
País: China (Huo, Z. et al., 2021)	automático y dos modelos de aprendizaje profundo, para obtener predicciones de las ventas a corto plazo de los próximos 28 o 30 días, de tiendas comerciales, los datos fueron obtenidos de la tienda Wal-Mart, con productos de las áreas; pasatiempos, alimentos y hogar. Dichos productos son vendidos en 3 tiendas ubicadas en diferentes estados, además de los datos históricos utilizaron datos de precios pronosticados para información más puntual	encontró que existe mayor efectividad en los modelos de regresión lineal simple, ejecutadas en la librería de Python, STATS MODELS que permitió el análisis con técnicas de modelado estadístico. Los modelos ML en la librería SCIKIT-LEARN y aprendizaje profundo en TENSORFLOW por su parte, no fueron deficientes, sin embargo se sugirió contar con mayor cantidad de datos relacionados al calendario y precio para tener mayor impacto en las predicciones
“Predicción de posibles escenarios de pedidos pendientes en la cadena de	El deseo por una toma de decisiones precisa en el momento adecuado crea la necesidad de analizar si los estudios arrojados por modelos	Se consideró el pronóstico utilizando DRF y GBM que informa escenarios probables de decisión de pedidos. Los escenarios presentados se

Nombre	Proceso y herramientas	Resultados
<p>suministro mediante el uso de técnicas de aprendizaje automático</p> <p>Distributed Random Forest y Gradient Boosting”</p> <p>País: Canadá (Islam, S., et al., 2020)</p>	<p>de Machine Learning son realmente correctos, para esto se hace un estudio sobre las predicciones de cadena de suministro destinada únicamente en las predicciones de los pedidos atrasados. Se utilizan dos técnicas, bosques aleatorios o Distributed random forest y gradient boosting machine. Los cuales son elegidos en la plataforma H2O</p>	<p>dieron del resultado del estudio únicamente de características generales de los pedidos basado en el inventario histórico, los tiempos de entrega, las ventas históricas y las ventas que ya fueron pronosticadas, evitando datos de riesgo más específicos. Así se identificó que el rendimiento de los modelos ML es mayor cuando aumentan los datos a distancia, y el modelo es flexible a la variación de las características de las demandas.</p>
<p>“Pronóstico del riesgo de obsolescencia y el ciclo de vida del producto con aprendizaje automático”</p>	<p>El objetivo se concentra en predecir el ciclo de vida del producto (mercado de teléfonos celulares) argumentando y evidenciando que los enfoques de Machine Learning resultan eficientes para dicho estudio.</p>	<p>El modelo ML arrojó datos de las características de las piezas, clasificándolas como activas u obsoletas con un 98 % de precisión y predijo las fechas de obsolescencia en próximos meses. Dicho modelo predice</p>

Nombre	Proceso y herramientas	Resultados
<p>País: Estados Unidos (Jennings, C., et al., 2016)</p>	<p>Se utilizaron algoritmos ML como Random Forest, Maquinas de vectores de soporte y redes neuronales artificiales, para la propuesta, sin embargo, el algoritmo Random Forest presento mayor efectividad en el estudio a partir de su ejecución en el entorno y lenguaje de programación R</p>	<p>la fecha de en que el producto quedara obsoleto y se evaluó con la técnica MSE. La principal aplicación se presenta durante el desarrollo y mejoras de nuevos productos, es decir, a partir de los datos, se crea un análisis de cambios de materiales en la construcción del diseño.</p>
<p>“Una comparación de métodos de aprendizaje automático para la predicción de parámetros de corte en procesos de torneado de alta velocidad” País: Croacia</p>	<p>Los datos representan las características del proceso de torneado de alta velocidad (ciclo de vida de la herramienta, estabilidad del terreno, y la fuerza de corte, esta información es analizada en el software MATLAB a partir del modelado de Machine Learning (redes neuronales artificiales y regresión de</p>	<p>El estudio pertenece al proceso de torneado de piezas de trabajo elaboradas con materia prima como hierro fundido, acero y aluminio. Se encontró la maquina adecuada y se concluye en que el algoritmo que más se acerca a los resultados contenedores de los requerimientos del cliente en cuanto al producto, es la</p>

Nombre	Proceso y herramientas	Resultados
(Jurkovic, Z., et al., 2018)	vectores de soporte), con el fin de comparar que algoritmo resulta más eficiente para predeterminar el estado de los parámetros, y la calidad del torneado.	regresión, que demostró el mejor rendimiento dentro de la propia evaluación de acuerdo a la predicción de la calidad.
<p>“Aprendizaje automático para el mantenimiento predictivo de máquinas industriales utilizando datos de sensores IoT”</p> <p>País: India</p> <p>(Kanawaday & Sane, 2017)</p>	<p>El estudio de caso se realiza para generar pronósticos con el fin de prevenir fallas y generar un mantenimiento predictivo para potenciar la producción al máximo posible.</p> <p>Específicamente se requiere generar la predicción de la rugosidad de la superficie, ya que representa un parámetro importante para que se defina la calidad tanto del proceso como del producto y su estética respectivamente.</p>	<p>Los datos a analizar pertenecen al proceso de una cortadora que rebobina y corta películas de embalaje, y se retoman las características de su proceso (presión, ancho, diámetro, tensión y tiempo). Las fases se clasifican en aprendizaje supervisado y modelos predictivos. Lo cual proporciona herramientas para resolver las limitantes en la productividad y costos de mantenimiento.</p>

Nombre	Proceso y herramientas	Resultados
<p>“Pronóstico de fallas de equipos industriales utilizando clasificadores de aprendizaje automático”</p> <p>País: Grecia</p> <p>(Kolokas, N., et al., 2018)</p>	<p>El principal objetivo de esta implementación fue la predicción de posibles paros o fallas existentes en un equipo.</p> <p>Los datos con los que se trabajó, son resultados de observaciones ópticas que se hicieron durante el proceso de funcionamiento para determinar el ciclo de ocurrencia de la falla. Los algoritmos utilizados fueron; perceptrón multicapa, regresión logística, árbol de decisión, bosque aleatorio, y estos fueron implementados con un lenguaje de programación Python y las librerías SCIKIT- LEARN, KERAS Y PYRENN.</p>	<p>Se obtuvo un estudio del proceso con el que funcionan los algoritmos Machine Learning en el que se encontró que este, solo es capaz de predecir los paros de 5 a 10 minutos antes de que falle el equipo dentro de la industria, tiempo que resulta insuficiente para solucionar el problema, por lo cual se sugiere dotar de mayor información al modelo para generar la capacidad de generar informes con mayor tiempo de anticipación.</p>
<p>“Selección y aplicación de algoritmos de</p>	<p>La necesidad de tomar decisiones optimas en cuanto a la elección de un algoritmo de</p>	<p>Se presentaron los algoritmos aplicables en un entorno de calidad de producción, en la</p>

Nombre	Proceso y herramientas	Resultados
<p>aprendizaje automático en calidad de producción”</p> <p>País: Alemania (Krauß, J., et al., 2019)</p>	<p>Machine Learning, con el fin de que las decisiones no afecten de manera positiva en la resolución de un problema de manera específica, para este estudio, se seleccionan 2 software de aprendizaje automático (PYTHON y R), sin embargo, el software principal es Python con la librería SCIKIT-LEARN para desarrollar el modelo.</p>	<p>fabricación compuesta específicamente por seis procesos, el estudio arroja instrucciones de los parámetros de cada requerimiento con los que debe cumplir el producto final. Y el algoritmo resultante como eficiente para este tipo de problema fue árbol de decisión.</p>
<p>“Previsión de ventas de tiendas minoristas utilizando técnicas de aprendizaje automático”</p> <p>País: India (Krishna, A., et al., 2018)</p>	<p>Se proponen técnicas para la predicción de ventas en una tienda minorista</p> <p>Con ayuda del lenguaje PYTHON y la librería SCIKIT-LEARN se administraron los datos para su ejecución, como los datos son etiquetados, se consideró un estudio en aprendizaje de tipo</p>	<p>Se evaluó y comparo el rendimiento de varios modelos de Machine Learning utilizados para las predicciones de venta, y se calificó al mejor. Esta evaluación se ejecutó a partir del error cuadrático medio. Se realizó el ajuste de la predicción de las ventas, con la técnica de Machine Learning,</p>

Nombre	Proceso y herramientas	Resultados
	<p>supervisado. La información obtenida pertenece a las características descriptivas tanto del producto (marca, utilidad, ofertas, empaque, etc.) como de la tienda (ubicación, ambiente, clientes, competidores, etc.)</p>	<p>potenciación del gradiente, con la que se estudiaron los datos y se clasificaron de acuerdo a los tiempos.</p>
<p>“Análisis de componentes principales e índice de calidad de vía: un enfoque de aprendizaje automático” País: Estados Unidos (Lasisi & Attoh-Okine, 2018)</p>	<p>Se busca encontrar los parámetros de estabilidad en la calidad de vías férreas, los estudios se ejecutan en programación R y R studio. Los datos, fueron extraídos de muestras de un ferrocarril de clase 1 y sus recorridos en una milla de vía. Existían 3 secciones con por lo menos un defecto por lo cual fueron especificaciones clave para determinar los componentes e</p>	<p>Se propone la reducción de análisis de la información en cuanto a la geometría de las vías, para generar consecuencias de programación de mantenimiento en tiempos totalmente alejados. Por lo cual se optimizan los procesos predictivos bajo enfoques de Machine Learning, únicamente se ejecuta el proceso cuando las técnicas arrojan alarmas de necesidad.</p>

Nombre	Proceso y herramientas	Resultados
	<p>índices de la calidad para futuras construcciones de vía.</p>	
<p>“Predicción del tiempo de entrega utilizando algoritmos de aprendizaje automático: un estudio de caso de un fabricante de semiconductores” País: Austria (Lingitz, L., et al., 2018)</p>	<p>El estudio de caso se llevó a cabo en una industria de semiconductores, en el que durante el proceso del producto final es necesario la producción de múltiples capas, los datos analizados fueron proporcionados por la empresa en cuestión, la información de un periodo de 2 años es referente a los históricos de las máquinas y cuestiones relacionadas con los clientes. El análisis se ejecutó en el software R y R studio para el manejo mayormente facilitado.</p>	<p>Durante la implementación fue posible evaluar diferentes métodos de aprendizaje, modelos lineales y de regresión, validados con la técnica cruzada, los datos utilizados fueron numéricos exclusivamente, y el resultado de la evaluación de los modelos condujo a la superación cualitativa producida por los conjuntos de árboles a los demás modelos. Así es posible determinar el impacto en el comportamiento de los tiempos de entrega.</p>
<p>“Marco de gemelos digitales basado en aprendizaje</p>	<p>Uno de los principales componentes a implementar en la industria petroquímica es el</p>	<p>Las simulaciones están destinadas a resolver predicciones en tiempos</p>

Nombre	Proceso y herramientas	Resultados
<p>automático para la optimización de la producción en la industria petroquímica”</p> <p>País: China</p> <p>(Min, Q., et al., 2019)</p>	<p>marco de gemelos digitales con el fin de darle atención a la toma de decisiones de manera efectiva, en un entorno mayormente cambiante, el enfoque que se le da a este marco estará basado en modelos de Machine Learning y big data para dar solución al análisis de grandes cantidades de datos lo más rápido posible.</p> <p>Los datos son analizados en línea con el sistema MES para desarrollar la información virtualmente.</p>	<p>específicos con el fin de prever los estándares de la producción real, así sería posible dar respuestas al control del proceso y su mejora continua de acuerdo a una mayor optimización. Los modelos son entrenados con cuatro algoritmos diferentes, sin embargo, el de mayor rendimiento resulta ser Light BBM, en el marco de mejoras del gradiente mejorando los impactos de sus resultados de predicción.</p>
<p>“Aprendizaje automático para procesos de fabricación basados en sensores”</p>	<p>Se pretende hacer un estudio de un conjunto de datos generado mediante un proceso de fabricación moderna basada en sensores de semiconductores (SECOM).</p>	<p>Se estudiaron técnicas de Machine Learning para datos que se caracterizan por el ruido. Dichos datos extraídos de plataformas en línea fueron muestreados en el software WEKA. Los algoritmos de</p>

Nombre	Proceso y herramientas	Resultados
País: Rumania (Moldovan, D., et al., 2017)	El algoritmo se ejecutó en el software R y R studio con la prioridad de evaluar los algoritmos como técnicas que ayuden a mejorar dichos procesos a partir de la clasificación de las características obtenidas, antes de que el producto llegue a la fase final.	clasificación con los que se ejecuta el análisis principalmente fueron la regresión adaptativa multivariada y Boruta.
“Técnica de detección basada en aprendizaje automático para END en fabricación industrial” País: Italia (Niccolai, A., et al., 2021)	Machine Learning se implementó para realizar identificaciones del estado de la falla de ciertos parámetros, específicamente en este estudio, se ejecutaron redes neuronales artificiales en el sistema de cómputo MATLAB y su librería NEURAL NETWORK TOOLBOX para el desarrollo del modelo propuesto. Las imágenes	El modelo permitió identificar la clasificación de los puntos en las imágenes. Se realizó un análisis numérico de los datos obtenidos de la clasificación, directo e indirecto. Debido a que los datos fueron etiquetados de manera manual, persiste la posibilidad de información errónea, sin embargo, el modelo cumple con la posibilidad de ejecutar

Nombre	Proceso y herramientas	Resultados
	<p>analizadas se sometieron a la reducción del ruido a través del algoritmo de umbral. Los datos describen el estado de la calidad encontrada en las imágenes.</p>	<p>modificaciones en el sistema para un estudio más concreto a partir de la suma de datos. Las pruebas no destructivas realizadas se aplicaron a pruebas de calidad industrial. Y se detectaron los parámetros defectuosos a partir de clasificaciones de las partes más importantes.</p>
<p>“Análisis de aprendizaje automático sobre la estabilidad de las células solares de perovskita” País: Turquía (Odabaşı & Yıldırım, 2020)</p>	<p>Se analizan datos de publicaciones históricas de las características a largo plazo de células solares de perovskita de haluro de organoplomo. Los datos se recopilan con ayuda del software DIGIZELT, para ser ejecutados en el software R y R studio. Con el fin de definir las condiciones de almacén que se debe considerar en las estancias de las células para</p>	<p>Si bien, los resultados arrojan que el estado del entorno ambiental, tienen importante impacto en la estabilidad de las células, un factor que también impacta sustancialmente está basado en las características de las células, los componentes que contienen. Este estudio finalizó con la construcción de un manual de manejo frente a las condiciones ambientales,</p>

Nombre	Proceso y herramientas	Resultados
	que estas mantengan parámetros estables.	basado en predicciones Machine Learning del estado de las células estudiadas.
<p>“Detección automática de fallas para la fusión de lecho de polvo láser mediante aprendizaje automático semi-supervisado”</p> <p>País: Reino Unido (Okaro, L., et al., 2019)</p>	<p>Se utilizó un modelo de Machine Learning para comprobar la calidad de las piezas producidas en una impresora tipo 3D. La herramienta principal se basó en el aprendizaje semi-supervisado debido a que no existían cantidades suficientes de datos con especificaciones claras. Con la ayuda de métodos se comprobó el porcentaje de funcionamiento del modelo para saber si la cantidad de datos que se etiquetaron a partir de los ya etiquetados era real o factible.</p>	<p>Se construyó un estudio de viabilidad del uso de Machine Learning para predecir las fallas a partir de un modelo de algoritmo de aprendizaje semi-supervisado con ayuda de datos tomados de muestras producidas por la descomposición aleatoria de variables singulares. El modelo a partir de dos tipos de datos que como resultado funciono como aprendizaje supervisado con la totalidad de datos etiquetados. Para la estrategia establecida direccionada a controlar o eliminar el riesgo de fallas fue exitosa.</p>

Nombre	Proceso y herramientas	Resultados
<p>“Aprendizaje automático para detección de anomalías y clasificación de fases de procesos para mejorar las actividades de seguridad y mantenimiento”</p> <p>País: Italia</p> <p>(Quatrini, E., et al., 2020)</p>	<p>El conjunto de datos pertenece a una empresa farmacéutica dedicada a la realización de granulación en lecho fluido, el proceso está basado en dos tipos de métodos, húmedos y secos. Y este estudio se generó en la producción por el método húmedo. Los problemas a solucionar se relacionan con peligros, resultado de las altas temperaturas y la contaminación. Se utilizaron los softwares, Python y Azure Machine Learning Studio para el desarrollo del modelo Machine Learning. Y los algoritmos utilizados están direccionados a la clasificación de tipo no supervisada o algoritmos de aprendizaje no supervisado.</p>	<p>El modelo que se generó permitió la detección satisfactoria de anomalías, obteniendo datos en tiempo real, el análisis se divide en dos fases. La primera fase identifica todos los acontecimientos ocurridos durante el proceso que se está ejecutando y la segunda fase está basada en la información encontrada en la fase 1. El entrenamiento del modelo se ejecutó a partir de los datos proporcionados por el propio personal de la empresa y los resultados son satisfactorios, pues las pruebas reflejan un etiquetamiento correcto de datos, y la predicción de fallas y anomalías puntual</p>

Nombre	Proceso y herramientas	Resultados
<p>“Automatización robótica de procesos con aumento de la productividad y mejora de la calidad del producto utilizando inteligencia artificial y aprendizaje automático”</p> <p>País: India (Rajawat, A., et al., 2021)</p>	<p>El principal objetivo de esta propuesta, está basado en promover la flexibilidad en inteligencia de un dispositivo destinado al micro posicionamiento, los datos son entrenados por el algoritmo YOLO a través de las imágenes partiendo de su implementación en JAVA, Se representan los puntos de parada de robots en la industria, y se genera un entrenamiento a partir de los conocimientos de las actividades ejecutadas, los modelos se derivan de cinco algoritmos entrenados, redes neuronales artificiales y convolucional, k-medias. K-vecino más cercano y máquinas de vectores de soporte, para la cual, resulta de mayor</p>	<p>Como resultado, se propone un sistema en el que los robots y los humanos sean capaces de colaborar en espacios y tiempos iguales e interactuar en conjunto para optimizar al máximo los procesos, si bien, la automatización robótica de procesos, inicialmente resultaba optima únicamente para industrias aeroespaciales, ahora se promueve la incorporación al sector manufacturero para restringir elevados costos de producción y de energía y mejorar el ambiente laboral, para que la producción no dependa en altos niveles de intervención humana.</p>

Nombre	Proceso y herramientas	Resultados
	rendimiento de precisión la red neuronal convencional.	
<p>“Enfoques de aprendizaje automático para estimar el consumo de energía de edificios comerciales”</p> <p>País: Atlanta</p> <p>(Robinson et al., 2017)</p>	<p>Utilizando modelos de aprendizaje automático implementados en PYTHON con la librería SCIKIT-LEARN ejecutados para evitar costos esfuerzos de modelado físico y poder obtener estimaciones razonables y crear mapas integrales de consumo de energía comercial metropolitana aplicándolos a 73388 edificios comerciales de Atlanta, estos mapas ayudaran a que los urbanistas entiendan mejor las relaciones entre forma urbana y consumo de energía y plan para el futuro.</p>	<p>Predicción y consumo de energía principal anual de edificios comerciales usando solo algunas características de los edificios</p> <p>Expresan este objetivo en un formato de regresión de aprendizaje automático</p> <p>X: características de todos los edificios en el conjunto de datos MFBTU para el edificio i y así, a partir de esto generar la predicción del valor MFBTU de un edificio con algoritmos de regresión lineal.</p>

Nombre	Proceso y herramientas	Resultados
<p>“Inspección de calidad basada en modelos predictivos utilizando Machine Learning y Edge Cloud Computing”</p> <p>País: Alemania (Schmitt, J., et al., 2020)</p>	<p>La implementación se desarrolló en una industria electrónica, los modelos de Machine Learning se ejecutan en el clúster Spark con librerías de ML. Su principal objetivo es dar soluciones basadas en algoritmos predictivos de la calidad del producto terminado y con ayuda de datos etiquetados, es decir, los modelos serán ejecutados de tipo supervisado. Por otro lado, se propone una evaluación de los modelos para que surjan cambios positivos y mejoras de la calidad a partir del proceso, y no hasta la parte final de la fabricación.</p>	<p>Se dota al sistema la capacidad de generar reportes sobre la calidad del producto con ayuda de técnicas de Machine Learning y Edge Cloud Computing para el análisis de los datos, características tanto del proceso como del producto final. Los datos con los que se entrenan los modelos, surgen de datos históricos de la industria y se despliega una serie de cambio a partir de las predicciones resultado del análisis.</p>
<p>“Un procedimiento para rastrear cadenas de</p>	<p>Se propuso la plataforma de ETHEREUM VIRTUAL MACHINE, y el lenguaje</p>	<p>Las técnicas utilizadas del contexto Machine Learning cumplieron con su principal</p>

Nombre	Proceso y herramientas	Resultados
<p>suministro de alimentos perecederos basado en Blockchain, aprendizaje automático y lógica difusa”</p> <p>País: Corea (Shahbazi & Byun, 2020)</p>	<p>SOLIDITY con el fin de mejorar el sistema de trazabilidad de los alimentos, el principal mecanismo tecnológico que se uso fue Blockchain Machine Learning, con el cual se administró toda la información de los alimentos que se requiere con el fin de que los clientes tengan el conocimiento y la seguridad de todas las características que este contiene, así se tendría una mejor toma de decisiones en el consumo.</p>	<p>objetivo de informar al consumidor la mayor cantidad de información de los alimentos. Por otro lado, los datos utilizados fueron analizados para obtener la trazabilidad en el mecanismo de base de datos que se comparte dentro de las empresas. Se obtuvo la calidad y el ciclo de vida de los productos en tiempos cortos y técnicas confiables.</p>
<p>“Diseño de producción de baterías utilizando modelos de aprendizaje automático de múltiples salidas”</p>	<p>El caso de estudio se realizó en un laboratorio de baterías, la necesidad de diseñar efectivamente un modelo de producción que permita disminuir los costos de la línea permitió la implementación de</p>	<p>El algoritmo de red neuronal artificial cumplió con el enfoque para el diseño y producción de baterías, concluyó en la obtención de un modelo de previsión de un sistema de producción ciber-</p>

Nombre	Proceso y herramientas	Resultados
País: Alemania (Turetsky, A., et al., 2021)	modelos de Machine Learning con el principal fin de pronosticar ciertas características clave de los parámetros del producto final, realizando análisis que inician desde la predicción del rendimiento de los procesos hasta la calidad de lo producido. El software utilizado fue PYTHON 3.8.5 con las librerías NUMPY, SCIKIT-LEARN Y PANDAS.	físico flexible a las posibles actualizaciones de los parámetros de las piezas. Estas actualizaciones se darían según las predicciones de las propiedades que el mismo sistema arrojaría. El modelo mostro la capacidad suficiente para predecir parámetros en la producción, sin embargo, existe la recomendación de un entrenamiento con la mayor cantidad de datos posible.

Machine Learning, ha sido desarrollado en los últimos años de manera práctica y con gran rapidez dentro de la industria manufacturera, esto, como herramienta de predicción a partir de los datos o información para dar respuesta a una gran diversidad de preguntas, o simplemente mejorar y controlar procesos. Es por eso que es de vital importancia explorar acerca de las implementaciones de lo ya demostrado de Machine Learning, con el fin de respaldar las teorías y enfoques a los que esta direccionado el estudio, conocer los avances más destacados, generar propuestas de mejoras en su ejecución o funcionamiento y experimentar el uso de las nuevas tecnologías para optimizar al máximo los procesos productivos.

3.2 Análisis de la literatura

Se muestra el análisis de los resultados obtenidos por cada artículo selecto, con el fin de comparar y resaltar características específicas de cada uno.

De los 30 artículos, en la Figura 25. se pueden hacer las siguientes observaciones por área; uno se relaciona a la logística productiva, ya que va direccionado a las predicciones del tiempo de entrega, dos son aplicaciones para la detección temprana de fallas, dos se dirigen al pronóstico de ventas a partir de clasificadores de datos históricos, tres son implementaciones en diferentes ámbitos relacionados con la cadena de suministro, es decir pedidos o clientes en cuanto al producto, cuatro son mejoras a partir de modelos en áreas de mantenimiento predictivo específicamente, 7 artículos de mejoras en el área de calidad del servicio, proceso o productos y por ultimo 11 de estos artículos se desarrollan dentro de áreas productivas en general.

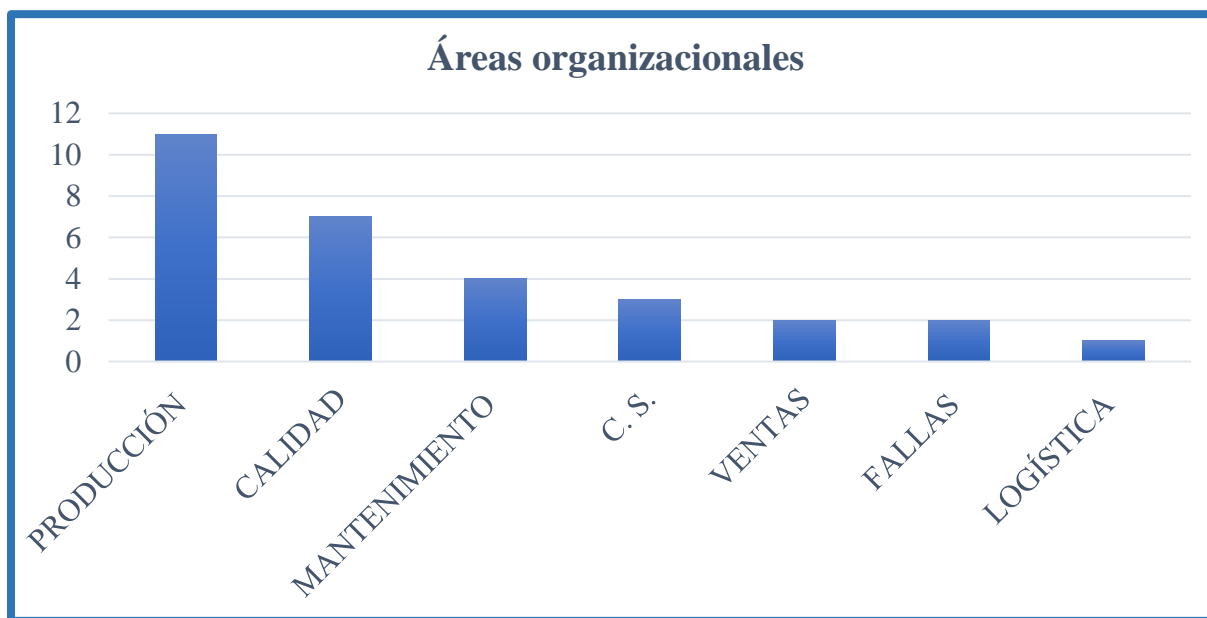


Figura 25. Áreas funcionales de una empresa con implementación Machine Learning (Elaboración propia).

En la figura 26 se registran las especificaciones de la implementación de acuerdo a las áreas utilizando cuatro gráficos. La figura 26a representa el desarrollo en áreas de producción, de acuerdo con las once implementaciones se visualiza que cuatro se direccionan al monitoreo del proceso, es decir a partir de los algoritmos se supervisa el buen funcionamiento de las áreas. Tres se desarrollan para predecir ciertas características o parámetros que se deben seguir, evitando así posibles errores o paros con ayuda de métodos guiados por ML. Dos son consistentes en implementar planes que permitan construir un entorno favorable tanto del proceso como del producto. Y por último dos de estos se enfocan en optimizar el proceso con ayuda de tecnologías de la inteligencia artificial y principalmente ML. En la figura 26b se presentan dos características en las que se dividen los 4 enfoques de mantenimiento por un lado existen tres pertenecientes a los mantenimientos predictivos hacia el proceso en el que se presentan la detección de anomalías directa y el análisis de datos en tiempo real para emitir la orden de mantenimientos en tiempos que permitan evitar al máximo los paros y por otro lado el mantenimiento predictivo orientado a las máquinas. La figura 26c representa el enfoque de áreas de calidad, de los siete artículos, tres se focalizan en la calidad del funcionamiento únicamente de las máquinas, donde se encuentran estudios de muestras obtenidas a través de registros de fallas ocurridas en las maquinas o las herramientas de las que dependen los procesos, y el impacto que el modelo tiene para resolver los problemas funcionales. Dos son parámetros enfocados al producto en el cual encontraremos predicciones de los requerimientos, riesgos y ciclo de vida, con el fin de evitar errores, mejorar el producto, o simplemente emitir información que sirva de utilidad para hacer comparaciones. Y dos se centran en la calidad del proceso, con el fin de inspeccionar que las actividades ejecutadas sumen valor al producto final, evitando pérdidas de tiempo y costos. Y por último los pronósticos, figura 26d muestra 4 de las áreas, ventas en la que se registran dos artículos,

aplicados a tiendas minoristas y tiendas convencionales en las que la predicción resulta una herramienta efectiva para decidir los estándares de producción dependiendo de las nuevas necesidades del cliente, dos representan el pronóstico de fallas tanto en las maquinas como en equipos o dentro del proceso productivo, solo uno se relaciona con los tiempos de entrega y modos de envío del producto. Finalmente, los pronósticos en el área de cadena de suministro en el que se mejora la trazabilidad de los productos, se evitan grandes cantidades de pedidos que aún no han sido entregados y cuestiones evidentes de los proveedores, o insumos importantes durante el proceso.

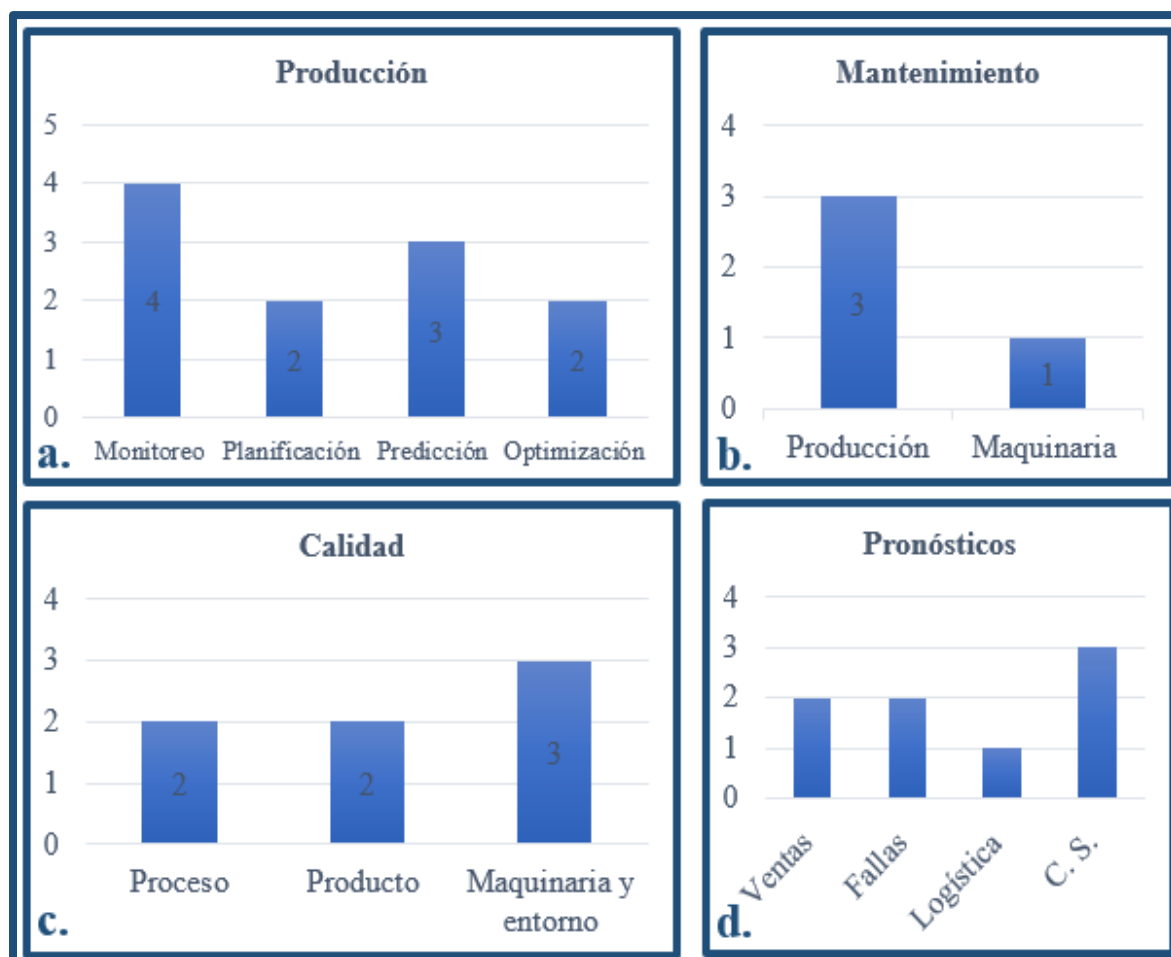


Figura 26. Especificaciones (Elaboración propia).

Debido a que Machine Learning es una herramienta basada en el análisis constante de grandes cantidades de datos, su implementación debe ser ejecutada en software destinados a realizar dicha tarea (Figura 27.) Diez de los artículos afirman que la plataforma utilizada fue PYTHON como lenguaje de programación, utilizando diferentes librerías consideradas como herramientas básicas para desarrollar ciencia de los datos entre las que se encuentran SCIKIT-LEARN, NUMPY, PANDA, etc. Por otro lado, R un entorno de programación en el que se desarrollaron cinco de los artículos, en compañía con R studio, un programa derivado de R, pero con características que facilitan su entendimiento para el usuario. MATLAB software o lenguaje producto de un laboratorio de matrices, con ayuda de su librería TOOLBOX en el que se desarrollaron tres artículos. El resto de artículos utilizaron un software específico cada uno; SCADA el sistema de software diseñadas para la supervisión control y optimización de los procesos. “H2O.ai” una plataforma de inteligencia artificial que emite interfaces para el desarrollo de Machine Learning. Azure Machine Learning Studio, fungió en un artículo de mantenimiento para acelerar los procesos del desarrollo del código en Python. Un artículo de cadena de suministro utilizó la Ethereum virtual machine para el análisis de Blockchain, dicha maquina ejecuta una cadena de bloques para controlar los datos. Apache Spark un software que ayudo en la generación de un clúster de aprendizaje automático y WEKA (Waikato Environment for Knowledge Analysis), un software o librería de JAVA ML desarrollada para clasificar datos. Por último, seis de los artículos omiten especificar con claridad el software utilizado, sin embargo, informan características del hardware en algunos casos.

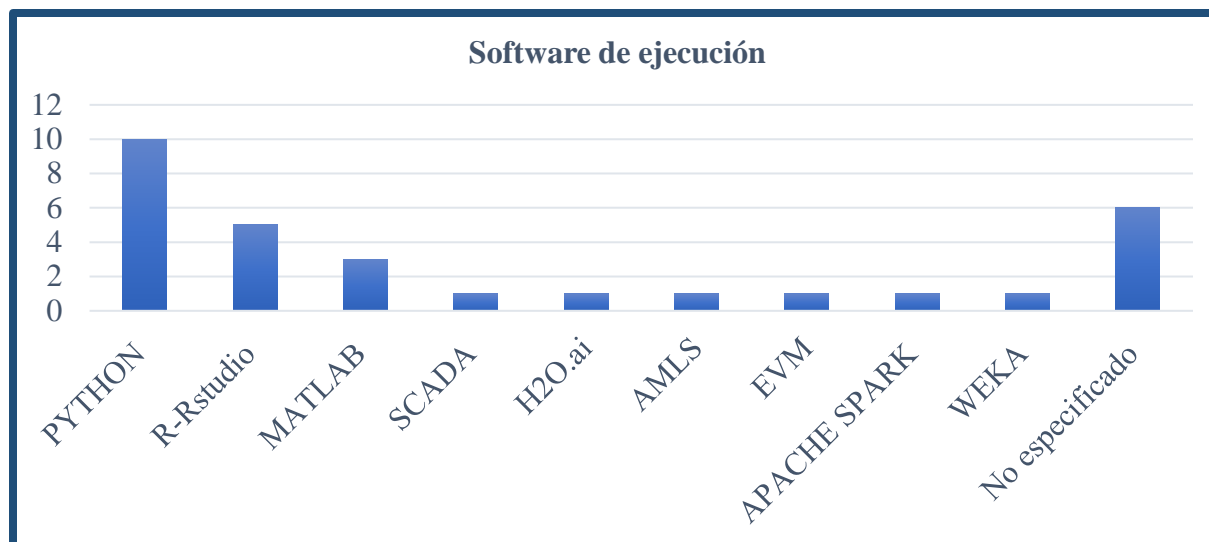


Figura 27. Herramientas de lenguaje y programación (Elaboración propia).

Los artículos analizados en su totalidad señalan que el tipo de modelo de Machine Learning que implementaron fue aprendizaje supervisado, sin embargo, en algunos casos, existe evidencia de que los datos sin etiquetar, se clasificaron de acuerdo a sus características y comparaciones con los ya etiquetados, por lo tanto, se considera que se utilizó el aprendizaje semi-supervisado. En la figura 28. Se muestran los algoritmos de aprendizaje supervisado y el número de artículos en los que fueron desarrollados o evaluados.

Es importante mencionar la revista de publicación de los artículos, y es que el factor de impacto tiene mucho que ver con la importancia que tiene cada artículo y el efecto que ejerce para saber si el análisis de determinado artículo es importante.

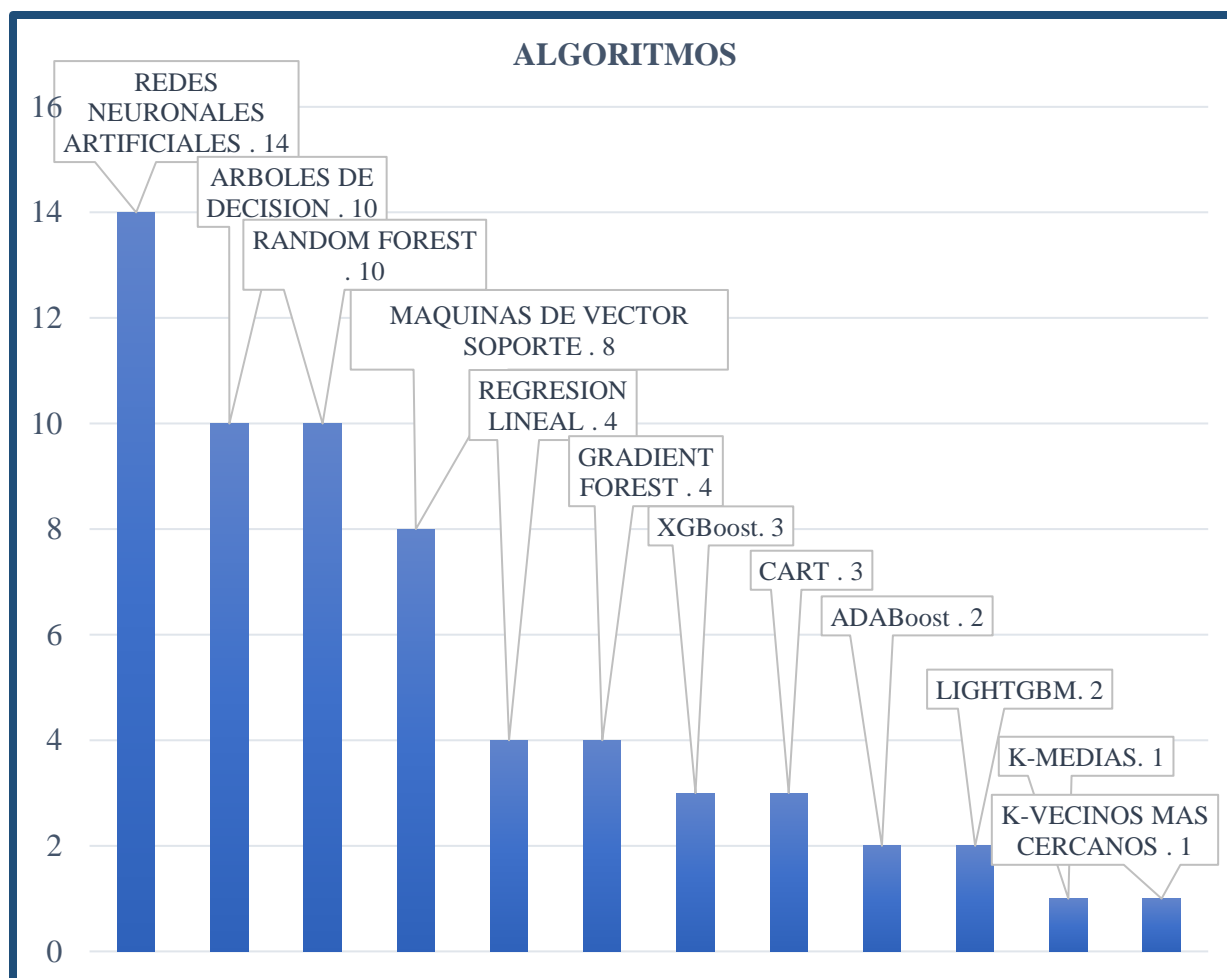


Figura 28. Modelos Machine Learning Utilizados (Elaboración propia).

La figura 29 muestra las revistas de las que fueron extraídos los artículos, entre las que podremos encontrar Elsevier, unas revistas científicas en la que se publican artículos de tipo informativo para la cual, muchos investigadores pueden generar avances, cuenta con índice de factor de impacto de 8.1, Springer, una editorial reconocida que aporta y publica artículos científicos dirigidos a sectores corporativos y académicos, su índice es de valor 6.9 y solo 3 de los artículos le pertenecen. IEEE Xplore, editorial de ocho artículos y con un índice FP de 4.5 publica artículos de análisis de investigación, y una de sus áreas se enfoca a las ciencias de la

computación. Y por último Multidisciplinary Digital Publishing Institute (MDPI) le pertenecen dos artículos, tiene índice 3.5, es una editorial de revistas científicas.

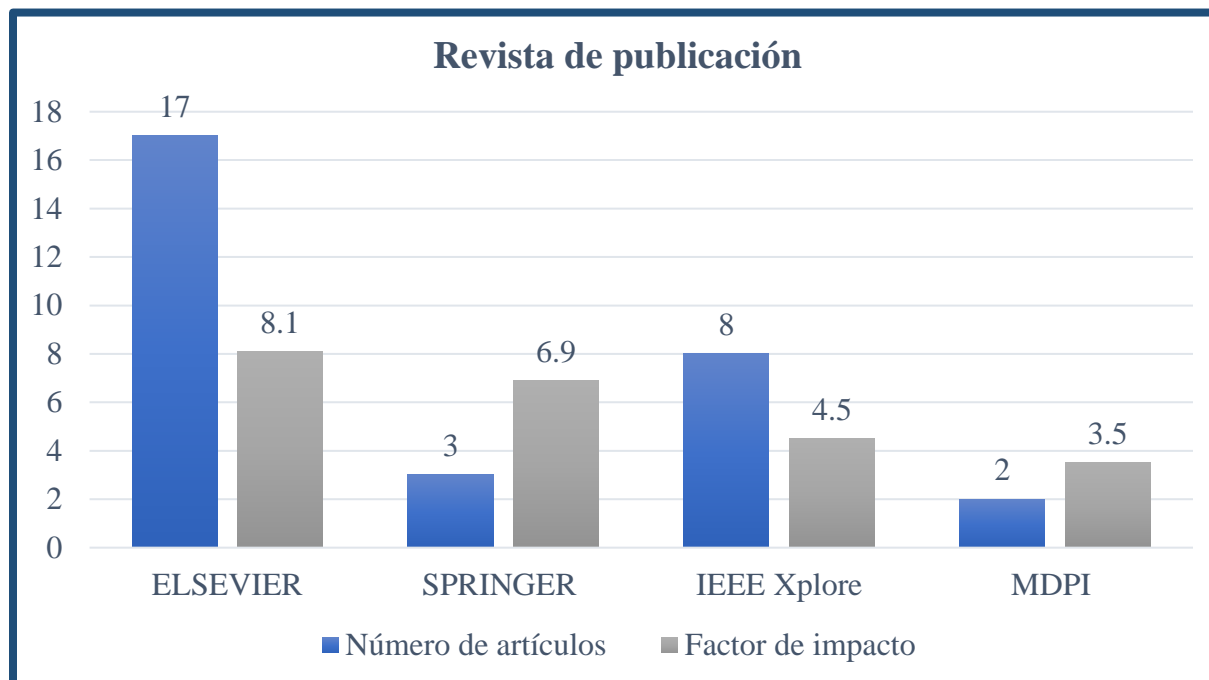


Figura 29. Número de artículos por revista y su factor de impacto. (Elaboración propia).

En la figura 30. se aprecian 16 países a los que pertenecen las publicaciones de los artículos, los que más destacan son Alemania, Estados Unidos e India con 4 artículos, les sigue China con tres artículos, Italia y Turquía con dos y los demás pertenecientes a diferentes países cada uno. En su totalidad suman 30 artículos relacionados con todo el entorno de implementación Machine Learning en dirección a la industria.

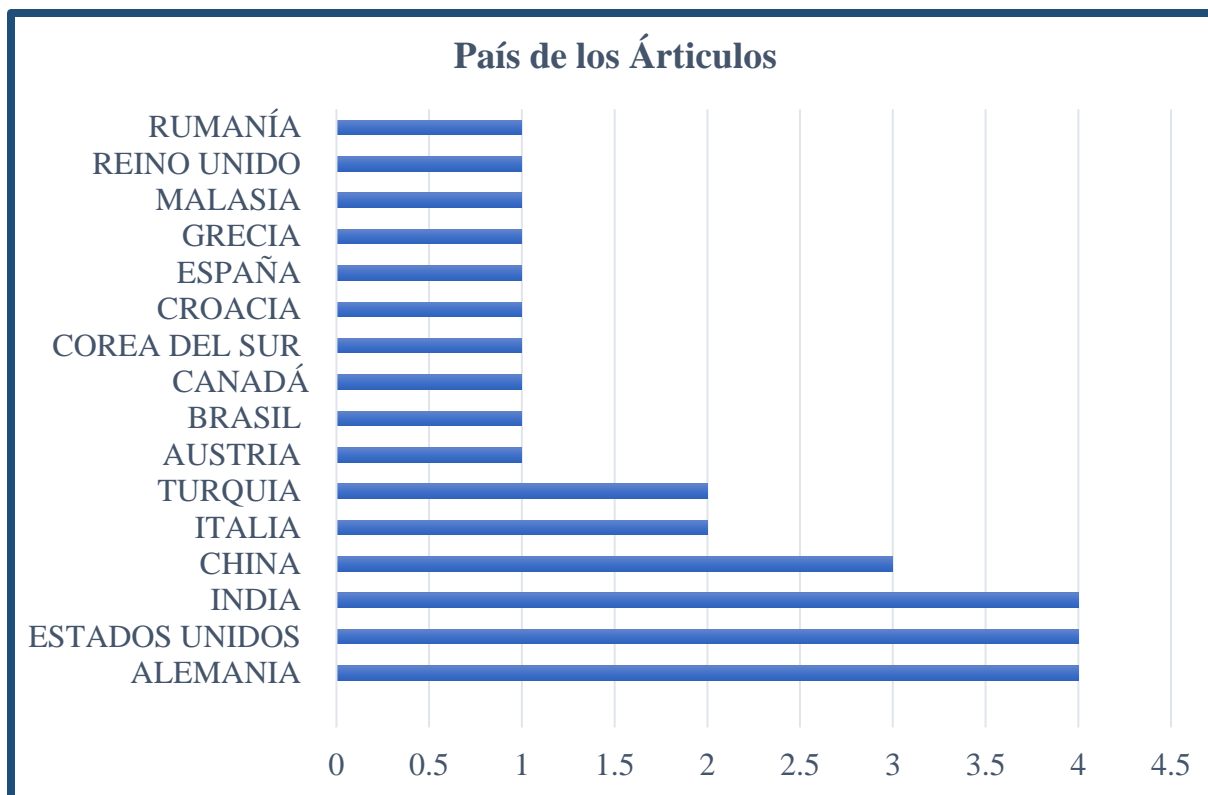


Figura 30. País de procedencia (Elaboración propia)

CAPÍTULO IV.
METODOLOGÍA Y
DESARROLLO

4.1 Descripción de las actividades

Continuando con la ejecución de los objetivos del presente trabajo, en la figura 31. se presenta un diagrama que contiene los pasos que se llevarán a cabo en la realización de este proyecto con el fin de cumplir los objetivos planeados, reduciendo la posibilidad de omitir puntos de importancia para el buen funcionamiento.

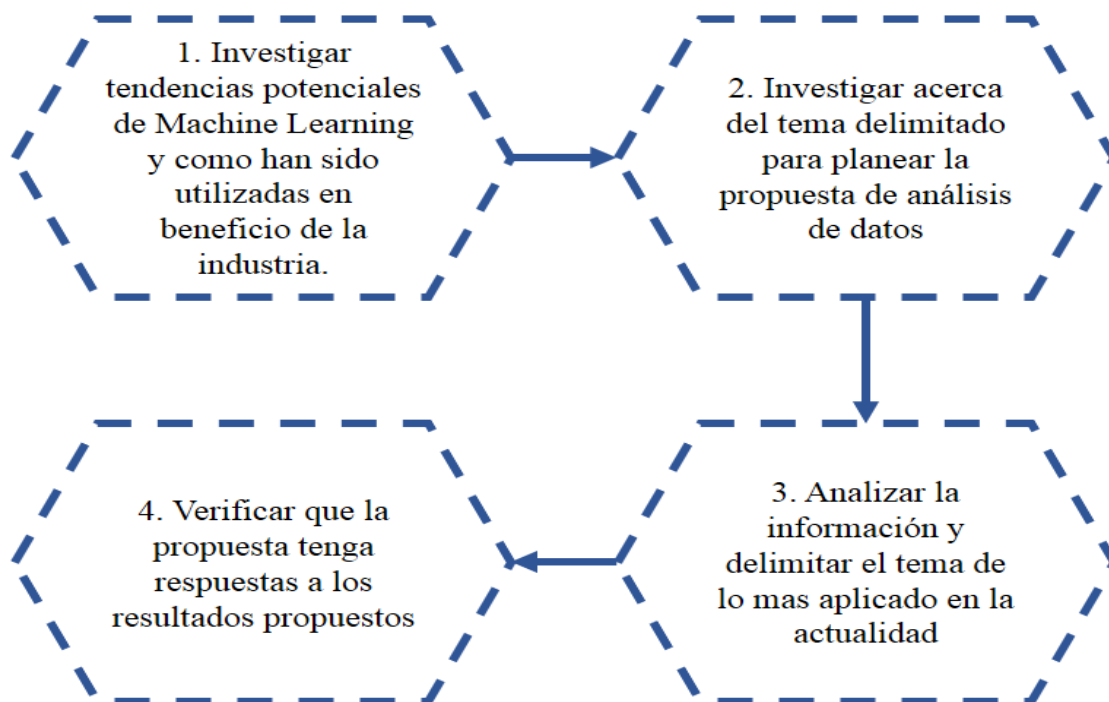


Figura 31. Bosquejo del método (Elaboración propia)

Para cada uno de los pasos, se explica en que consiste cada uno de ellos

1. Investigar las tendencias potenciales de Machine Learning y como han sido utilizadas en beneficio de la industria.

De acuerdo con el objetivo, existen ciertos parámetros que se toman en cuenta para que la investigación bibliográfica sea ejecutada de la mejor manera, como bien se ha documentado, Machine Learning, es una tecnología que, aunque ha tenido éxito en gran parte de sus

implementaciones, no existen aún, grandes espacios de aplicación, es decir, no ha sido implementada en rangos importantes.

El principal parámetro y requisito de la investigación se torna en la búsqueda de artículos en el buscador académico “Google Academy”, una plataforma que permite indagar sobre documentos de carácter académico (Figura 32). Se sustenta con información perteneciente a una gran variedad de fuentes de organizaciones académicas.

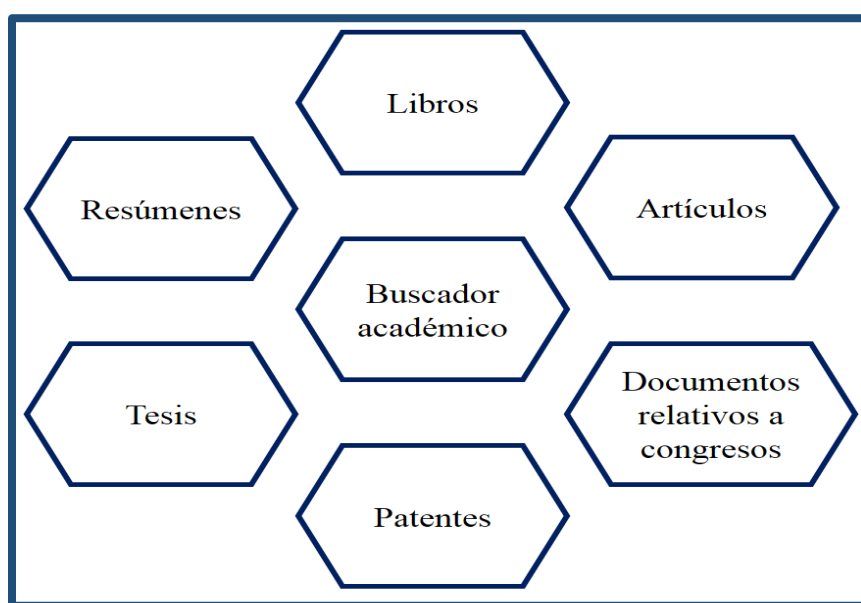


Figura 32. Documentos para indagar en el buscador académico (elaboración propia)

Dicho buscador, contiene herramientas que facilitan su funcionamiento el idioma, como requisito principal de esta búsqueda, ya que solo se permite que la información sea de artículos publicados con el idioma inglés.

2. Investigar acerca del tema delimitado para planear la propuesta de análisis de datos.

Si bien, Machine Learning es un campo que actualmente se aplica en varios ámbitos, este proyecto únicamente se direcciona al cumplimiento del beneficio en control de procesos industriales, por lo cual, es importante delimitar la búsqueda a implementaciones en áreas industriales para la optimización del flujo. Para ello se seleccionaron de manera minuciosa palabras clave (Figura 33) que concuerden correctamente en el uso de este sector.

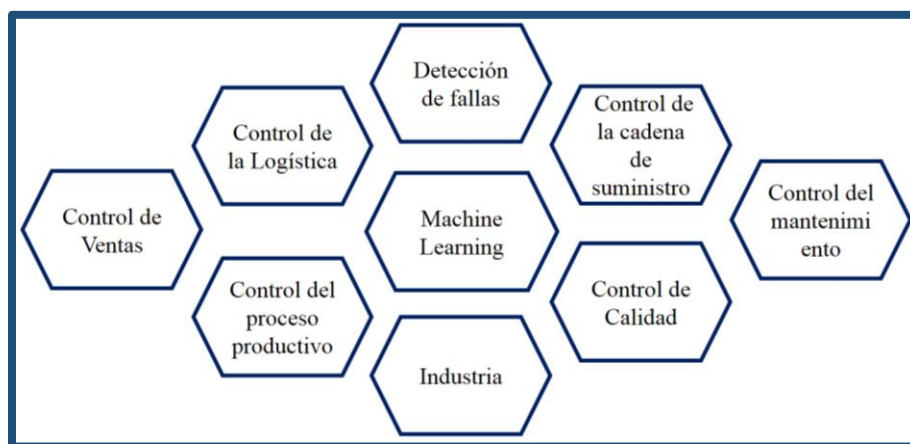


Figura 33. Palabras clave (Elaboración propia)

De esta manera, es posible omitir encontrar publicaciones de áreas como medicina, que, por el contrario, se establece como un ámbito en el que Machine Learning ha sido muy desarrollado, por sus características de emitir predicciones que ayuden con la detección y resolución de problemas de carácter medicinal.

3. Analizar la información y delimitar el tema de los más aplicado en la actualidad

Debido a que durante el desarrollo del inicio de esta investigación se encontró que existen publicaciones desde años atrás como 1900, 2000, etc., fue necesario establecer un rango el cual,

se puso en funcionamiento con la herramienta propia del buscador como la selección de los límites de años, con la cual se cubre el requisito de los años de publicación, para este proyecto se estandarizó la búsqueda de no mayor a 7 años antecesores a partir del año del desarrollo del proyecto, por lo cual el rango queda del 2015-2022. Una vez filtrada la búsqueda, los artículos pasan a ser analizados, este proceso consta de la revisión de la literatura de manera exhaustiva con el fin de clasificar la información.

Se genera un apartado dentro del proyecto mismo en el que se registran los puntos más importantes con los que se trabaja la propuesta (Tabla 5.) donde se estable las características de cada artículo.

Tabla 5. Desarrollo de información (Elaboración propia)

Aspecto	Descripción
Nombre del artículo	Considera como una breve introducción de lo que se hizo durante el desarrollo de la implementación
Herramientas utilizadas	Proporciona al lector una descripción clara y concisa de las herramientas que se utilizan, mencionando los softwares y lenguajes de programación, así como también en algunos casos las características del hardware para su utilización como referencia.
Resultados obtenidos	Se muestra un resumen de las conclusiones y recomendaciones que se dan para que se haga uso dentro de la propuesta en desarrollo en diversas áreas.

País	Conocer el país del que proviene la publicación, es importante, ya que podemos identificar las diferencias entre las ejecuciones de muchas actividades con respecto al país propio conocido.
Autor	Se menciona con el fin de evitar dificultades en cuestiones de plagio, así como también, para identificar la descripción profesional de quien ha desarrollado dicha ejecución.

Hacer un plan que determine cuál es el área en la que la herramienta tiene consecuencias con mayor éxito para la optimización efectiva entre las áreas con ML en creciente implementación para un futuro análisis específico de cada uno entre las que se encuentran; ventas, cadena de suministro, calidad, almacén, logística, producción y administración en general.

4. Verificar que la propuesta tenga respuestas a los resultados propuestos

De acuerdo con la propuesta generada, se debe verificar que se cumpla con una explicación clara de cada punto que se debe considerar para poder desarrollar una implementación real en cualquier área dentro de la industria. Comprobar y controlar a través del análisis exhaustivo de los resultados que la propuesta sea competente en su clasificación para evitar un estudio deficiente y desarrollar las posibilidades de la experimentación real de la implementación.

Designar el área en la que la implementación provoco cambios óptimos y genere resultados en la mejora continua del proceso, detallando la herramienta, método o software utilizado. Revisar que los resultados sean posibles candidatos a ser estudiados de manera aislada o exclusiva a otras áreas.

CAPÍTULO V.

RESULTADOS

5.1 Resultados y discusión

Una vez, clasificada la información, es preciso dar pauta a la documentación de la propuesta, con el fin de cumplir los objetivos de este documento, se presenta la sucesión de pasos propuestos de acuerdo con la organización Deming, esto debido a que se considera que dicha propuesta funge como mejora continua dentro de cualquier proceso.

Para iniciar, recordemos que el ciclo Deming es un sistema que se utiliza para ejecutar un plan de sistema de mejora en diferentes ámbitos, es conocido como Edwards Deming, o PDCA, cuenta con cuatro etapas (plan, do, check, act.) y es considerado como ciclo precisamente porque cuando se llega a la etapa final, se debe volver a iniciar, y así se une como estrategia perfecta de esta propuesta de implementación Machine Learning en cuanto al seguimiento de las evaluaciones y mejoras al rendimiento del modelo.

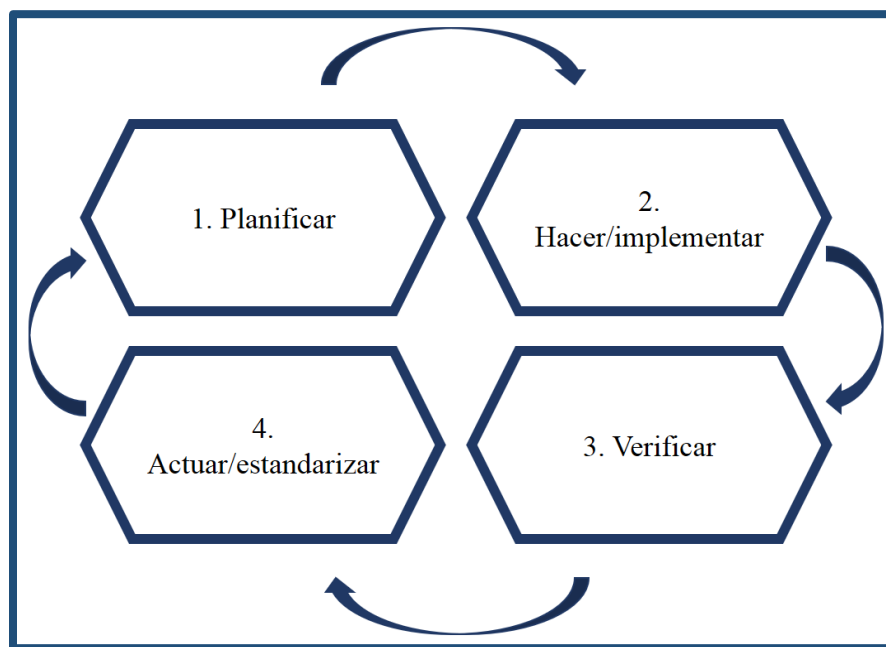


Figura 34. Ciclo Deming (Elaboración propia con información de Isniah S., et al., 2020)

Es así como las organizaciones ejecutan cualquier cambio dentro de sus procesos, con el fin de mantener una organización competitiva y mejorando la eficiencia de las mejoras.

5.1.1 Planificar

Si bien, dentro de una empresa es altamente probable encontrar problemas o simplemente áreas de mejora continua, es indispensable realizar la observación y análisis correcto de esto, pues el principal objetivo de esta propuesta es hacer que se tome como iniciativa desde el punto de origen del problema al que se tiene que dar solución.

Es recomendable utilizar herramientas que ayuden a entender y clasificar sobre los factores con mayor importancia del problema, omitiendo errores y la falta de información que se considere indispensable. Esto permite decidir desde que área es recomendable iniciar el proceso de solución. Las herramientas básicas y nuevas de la calidad (Tabla 6.) proporcionaran la capacidad de identificar de manera clara y precisa la raíz del problema y los aspectos más importantes a considerar.

Tabla 6. Herramientas para la medición de la calidad (Elaboración propia con información de Isniah S., et al., 2020)

Herramientas básicas	Herramientas nuevas
Histograma	Diagrama de relación
Gráfico de Pareto	Diagrama de matriz
Gráfico de causa y efecto	Diagrama de árbol
Gráfico de dispersión	Diagrama de afinidad
Hoja de verificación	Matrices de priorización
Gráfico de control	Diagrama de procedimiento

Herramientas graficas	Diagrama de programa de decisión de proceso
-----------------------	---

La elección de cualquiera de las herramientas de la calidad, depende de la necesidad del análisis que se esté ejecutando, y debe tomarse en cuenta que toda solución a un problema debe ser medible.

Una vez encontrada el área en la que se implementara la planificación, es necesario la selección del equipo de trabajo. Dentro de la propuesta se consideran cuatro dependencias más importantes (Figura 35), sin embargo, es importante mencionar que la alta dirección debe tener participación en todos los procesos, esto debido a que serán quienes tomen las decisiones del seguimiento del proyecto, así como también, mantener el abasto de la economía a invertir por mínima que sea. La dirección del área específica, es la principal fuente de información, pues será quien nos comparta todos los datos históricos, y el flujo de los procesos, resulta eficiente la participación de la totalidad de las áreas para que exista comunicación en el ambiente y los cambios no tengan consecuencias negativas dentro del proceso general de la producción. La parte fundamental se dirige al jefe del proyecto pues se encargará de darle estructura a la planificación sin que se omitan procesos indispensables y, por último, la creación de un área de programación, trabajar con especialistas en programación mejora el proceso y lo optimiza.

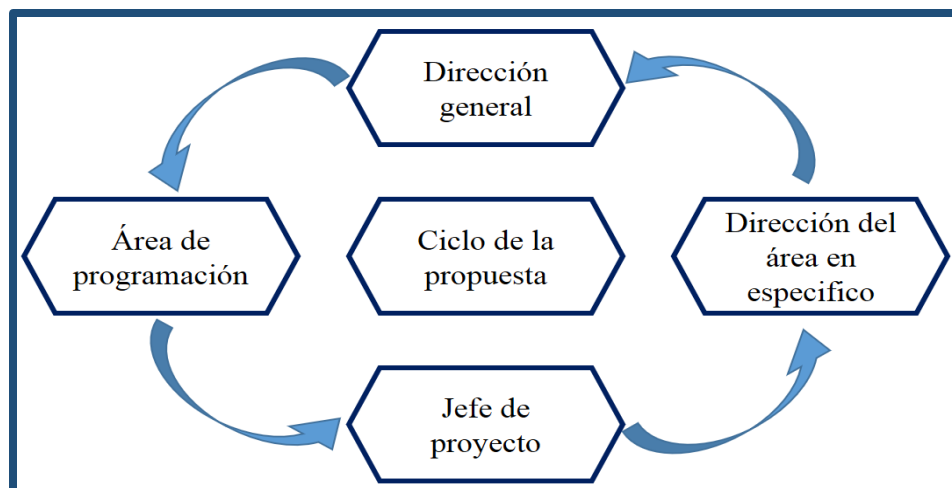


Figura 35. Dependencias (Elaboración propia)

De acuerdo con la revisión de casos de éxito en las implementaciones en el sector organizacional industrial de Machine Learning, se observa una elevada tendencia de su ejecución en las áreas de producción y calidad, sin embargo, es indispensable mencionar que solo se estudiaron artículos que se enfocaron en siete áreas. En la figura 36 se muestra la distribución de artículos según su variable. Planificación y control de la producción (PCP) contiene 11 artículos, gestión de la calidad (GC) tiene 7 artículos en total, hay 4 artículos de gestión del mantenimiento (GM), 3 artículos pertenecen la gestión de la cadena de suministro, y los últimos 5 se dirigen a la gestión de ventas (GV) gestión de fallas (GF) y gestión de la logística (GL) respectivamente.

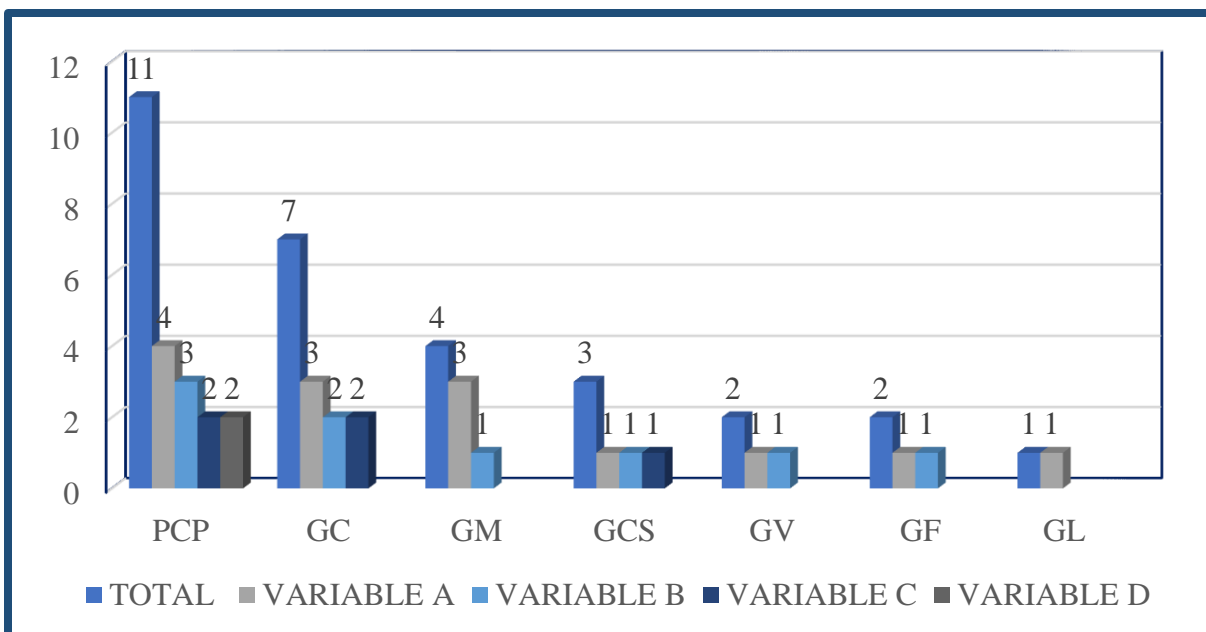


Figura 36. Artículos por dominio de implementación (Elaboración propia)

Cada una, relacionada a cada sector o área de la industria, pero contienen elementos diferentes, es decir, los campos de aplicación se dirigen a aspectos que contiene cada área. En la tabla 7 se muestran las especificaciones de cada una de ellas de acuerdo con el acrónimo asignado, se debe tener en cuenta que solo el área de PCP se divide en las 4 variables y las demás tienen una dispersión menor de acuerdo con el número total de artículos y sus variables.

Tabla 7. Designación e variables (Elaboración propia)

	TOTAL	VARIABLE A	VARIABLE B	VARIABLE C	VARIABLE D
PCP	Planificación y control de la producción	Monitoreo	Predicción para la construcción	Optimización	Planificación
	11	4	3	2	2
GC	Gestión de la calidad	Maquinaria y entorno	Producto	Proceso	
	7	3	2	2	
GM	Gestión del mantenimiento	Proceso	Maquinaria		
	4	3	1		
GCS	Gestión de la cadena de suministro	Pedidos pendientes	Insumos	Proveedores	
	3	1	1	1	
GV	Gestión de ventas	Predicciones	Previsiones en tiendas minoristas		
	2	1	1		
GF	Gestión de fallas	Maquinaria	Equipos		
	2	1	1		
GL	Gestión de logística	Tiempos de entrega			
	1	1			

Si bien, existen varios tipos de algoritmos en los que se clasifica Machine Learning, el conjunto de artículos, desarrollaron en su mayoría únicamente aprendizaje de tipo supervisado y semi-

supervisado, esto derivado de que se trata de implementaciones en industrias que contienen ya, información histórica y datos etiquetados para su estudio en la figura 36, se presenta la distribución del desarrollo y mención de cada técnica utilizada, RN que representa la cantidad de artículos enfocados al desarrollo de redes neuronales tanto convencionales como artificiales, 5 de ellos, desarrollaron técnicas de máquinas de vector soporte en la clasificación de los datos, 12 artículos, distribuyen y administran el resultado de acuerdo con la clasificación por árboles de decisión, utilizando técnicas como Random Forest, Gradient Boosting, arboles de clasificación y regresión, etc. Dos utilizan las técnicas K-medias y K- vecinos más cercanos y por último solo cuatro utilizan técnicas de regresión.

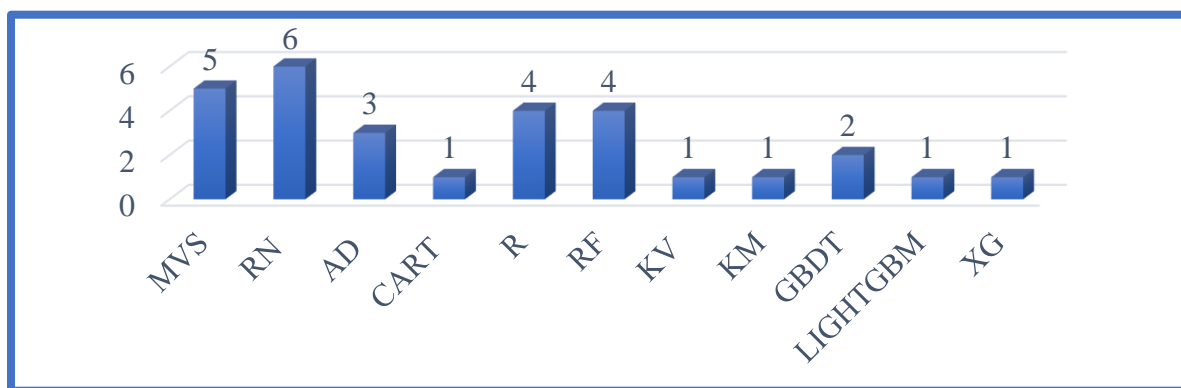


Figura 37. Algoritmos en producción (Elaboración propia)

La necesidad de implementar las tecnologías Machine Learning en el control de procesos específicamente industriales, conlleva a una revisión del éxito que se ha tenido a partir de las estrategias generadas en las industrias en funcionamiento de esta derivación. De acuerdo con el resultado de los análisis se obtiene información que conlleva una propuesta de posibles implementaciones en cualquier área organizacional figura 37.

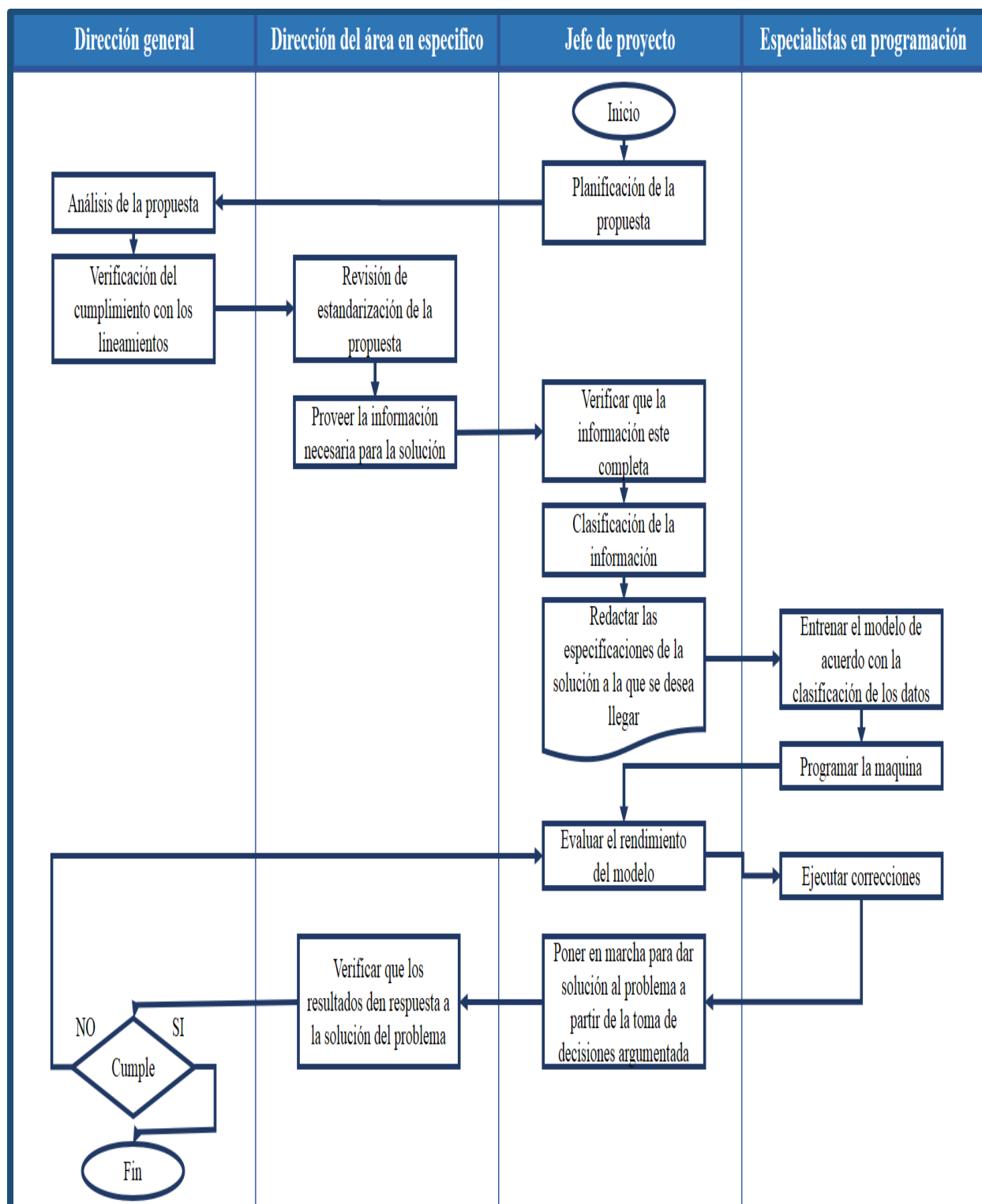


Figura 38. Actividades generales de cada responsable para la implementación (Elaboración propia).

1. Planificación de la propuesta

El primer paso conforma la elaboración de un plan que sea fácil de entender y que muestre los resultados específicos a los que se desea llegar, este paso es exclusivo para el encargado de desarrollar la implementación, se trata de estandarizar un proceso de la propuesta para que Machine Learning ayude con la clasificación de la información y sea posible encontrar patrones que demuestren el flujo óptimo de los procesos. Esta propuesta debe formularse de acuerdo al entorno en el que se encuentre el área y debe contener como requerimientos indispensables los procedimientos que se llevan a cabo y la totalidad de sus variables.

2. Análisis de la propuesta

Debido a que las tecnologías son un desarrollo nuevo dentro de la industria, es de vital importancia que los directivos se encarguen de analizar las nuevas propuestas que se tienen aun sabiendo que podría generar avances de eficiencia en cualquier proceso. Es así que se vuelve indispensable realizar el análisis exhaustivo de las ventajas y desventajas que el desarrollo de la implementación sería capaz de desencadenar, así como también de dar revisión a los costos que serían provocados de acuerdo a los requerimientos que se tengan. Después de revisar toda la información proporcionada por el jefe de proyecto, es necesario dar pauta a que se ejecuten nuevos acontecimientos que favorezcan el ambiente específico.

3. Verificación del cumplimiento con los lineamientos

es necesario también, que la mesa directiva, desarrolle un plan para que todos los procedimientos necesarios para la ejecución de la implementación no infrinjan con los lineamientos de ética o con el plan del proceso productivo, es decir, aunque las nuevas tecnologías representan un cambio.

4. Revisión de la estandarización de la propuesta

De acuerdo con la propuesta, la dirección del área en específico (producción, calidad, etc.) debe dar revisión a la planificación para dar respuesta a todas las necesidades que la ejecución contenga, es decir, debe proponer

5. Proveer la información necesaria para la solución

De acuerdo con la función principal de Machine Learning, es indispensable que los datos sean correctos y que demuestren importancia dentro del problema o simplemente el proceso, es así que la dirección del área, tiene la tarea de compartí toda la información necesaria.

6. Verificar que la información este completa

Una vez que se abastece la información pertinente (datos históricos) se genera una serie de datos a los que se desea llegar, así entonces, es posible especificar todos los puntos de importancia e iniciar el proceso principal

7. Clasificación de la información

En la figura 38. Se muestra una representación general del flujo propuesto para el desarrollo de las actividades que ejecuta el jefe de proyecto de acuerdo con los requerimientos de la solución. Iniciando por la clasificación de la información y los sistemas de recolección que este utilizara.

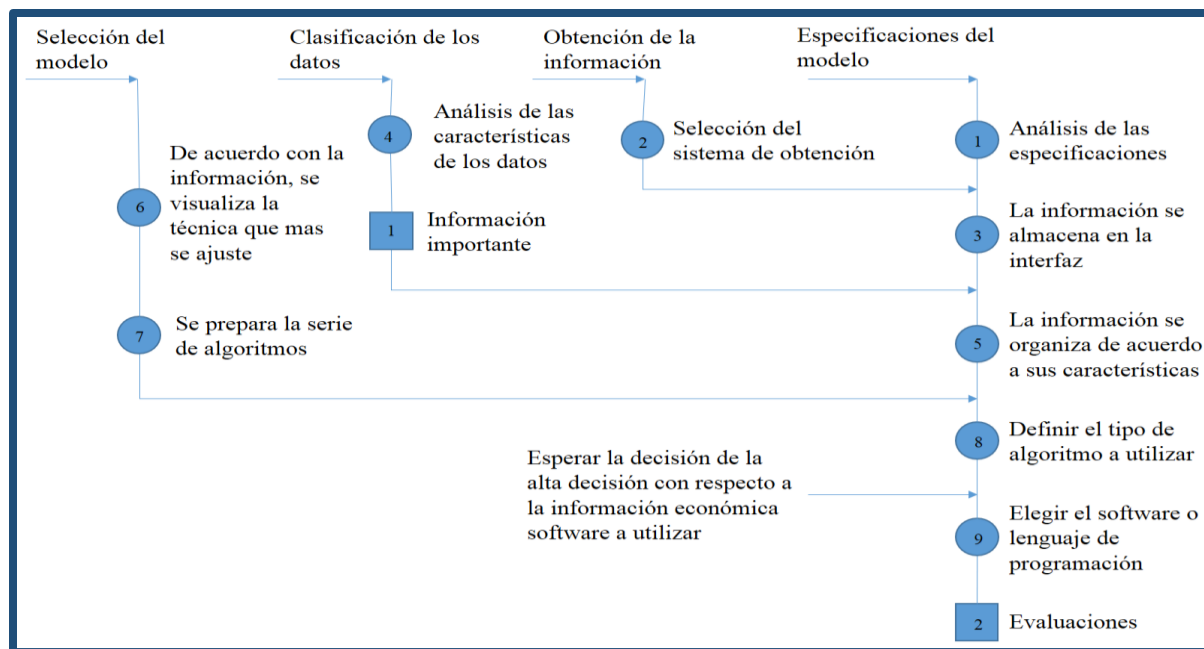


Figura 39. Proceso general de las actividades del jefe de proyecto (Elaboración propia).

Iniciando con el flujo de la información, como primera operación, el jefe de proyecto, deberá ejecutar un análisis de todos los requerimientos que se necesitan en cuanto a las soluciones. Y definirá internamente las características de los resultados que el modelo deberá tener.

La operación 2 se divide en dos atributos, para la obtención de los datos de manera manual, que es una técnica mayormente utilizada para la obtención de datos de calidad, logística, cadena de suministro en la que normalmente no existe un sistema implementado de la administración de la información más reciente. La información es trasladada por el usuario a la nube o cualquier interfaz que se considere capaz de almacenar grandes cantidades de datos para guardar en ella la mayor cantidad de datos. La segunda es la obtención de los datos a través de la información que los sensores emiten, es decir, existen sistemas de recolección de datos que sin necesidad de intervención humana son capaces de almacenar toda la información en tiempo real de manera autónoma, en las que se encuentran recomendados dos sistemas SCADA (como recolector de datos) y MES (como controlador de la gestión y ejecución de los datos de la producción. Los

datos se obtienen a partir de sensores o PLC's. o simplemente se pueden tomar en cuenta el uso de softwares de tecnologías de la información con DIGITIZELT o interfaces web.

Una vez generada la recolección de la información, se debe verificar que esta sea de vital importancia y que no se omita ningún tipo de dato jerarquizando su valor dentro de las especificaciones que los resultados emitan. De esta manera, se llega a la operación 3, se busca que la información sea almacenada en alguna interfaz capaz de guardar grandes cantidades de datos, o simplemente se pone en espera para ejecutarse al finalizar este pequeño proceso. Para la operación 4, la información se debe clasificar en etiquetada y no etiquetada de acuerdo a sus características. Sin importar la existencia o no de estos. Y se requiere la inspección exhaustiva para resolver si la información es realmente importante para ser utilizada en este proceso, la operación 5 continua con una organización simple de la organización y el análisis de esta para pasar a la operación 6 se recomienda tomar en cuenta cuatro alternativas principales haciendo hincapié en que existen una gran tipología de algoritmos y existen técnicas que se apegan cada vez más a una diversidad enorme de problemas a los que se desee dar solución. Los cuatro algoritmos principales para generar modelos de programas Machine Learning, independientemente de las especificaciones de las áreas, los datos pueden ser utilizados para entrenar el modelo. La elección de estos algoritmos depende totalmente de las características que la información contenga.

El algoritmo que más se adecua a la inexistencia de datos es el aprendizaje por refuerzo (Figura 40.) ya que es un sistema construido principalmente para que la maquina aprenda y ejecute acciones en funcionamiento a un entorno o ambiente industrial en el que se encuentra, dicho entorno, brindara información necesaria para poder ejecutar acciones con respecto al

estado en el que se encuentra y las recompensas que se emiten a través de experiencias prueba y error.

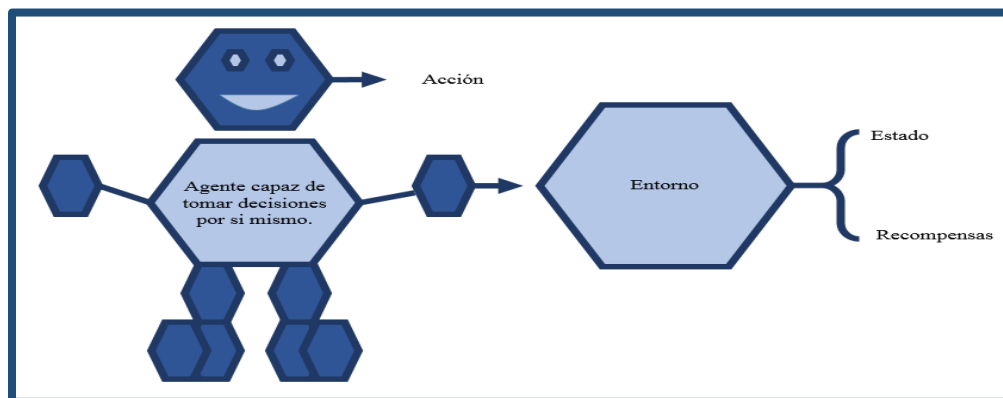


Figura 40. Proceso general del aprendizaje por refuerzo (Elaboración propia)

El aprendizaje no supervisado (Figura 41.) en solución a la inexistencia de datos etiquetados y grandes pérdidas de tiempo y dinero que podría costar etiquetarlos, este método, obtiene conocimientos a partir de las características de semejanza encontradas en los datos, es decir, clasifica la información, para poder aprender, y darle una característica con la que se logre etiquetar la información, sin embargo, en su aplicación, este modelo fue ajustado debido a que la clasificación emitida, suele ser información bruta, o datos que no especifican una respuesta.

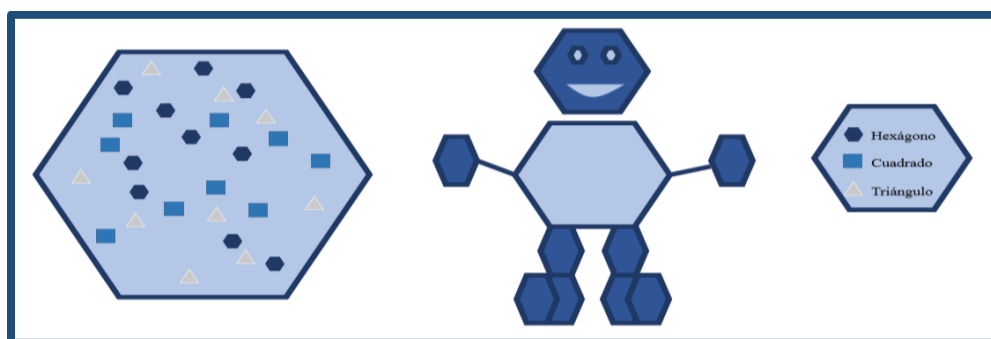


Figura 41. Aprendizaje no supervisado (Elaboración propia)

Considerando la posibilidad de que exista una mezcla entre información o datos etiquetados y no etiquetados, es posible hacer uso del algoritmo de aprendizaje semi-supervisado. Los modelos semi-supervisados se encargan de generar un proceso con el fin de evaluarse como aprendizaje supervisado. (Figura 42.)

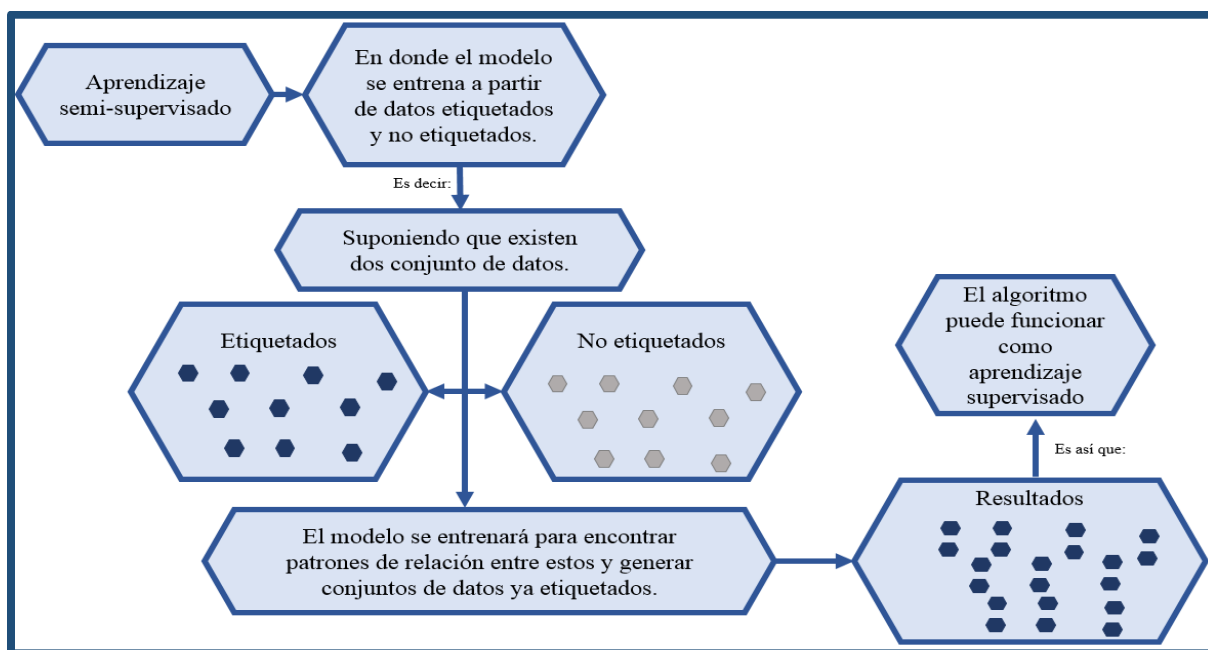


Figura 42. Proceso de datos etiquetados y no etiquetados para obtener elementos que permitan la implementación del aprendizaje supervisado. (Elaboración propia)

Si existen datos etiquetados, es decir, la información tiene estructura, es posible implementar aprendizaje supervisado con cualquiera de sus herramientas de clasificación (Figura 43.). De acuerdo con la revisión de los artículos, existe resiliencia superior en las ejecuciones del algoritmo de aprendizaje supervisado, por lo que se recomienda darle continuidad en este modelo, y ejecutar entrenamientos de un nuevo modelo con datos adquiridos a partir de la clasificación.

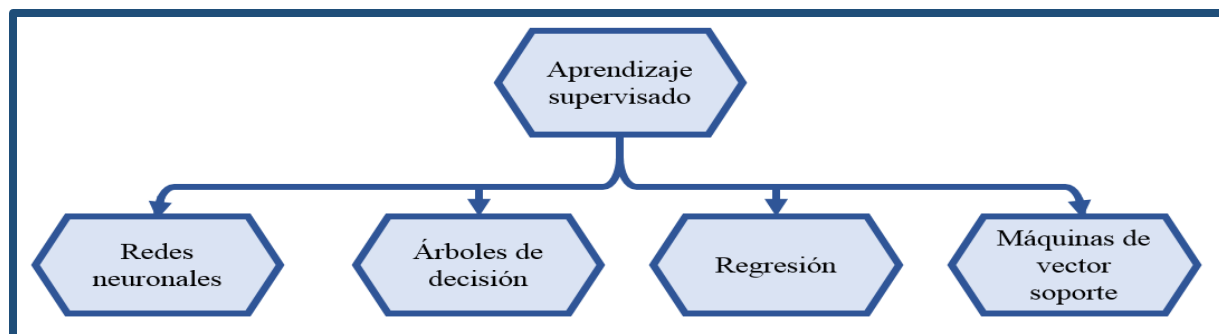


Figura 43. Técnicas y algoritmos de aprendizaje supervisado mayormente utilizados en las implementaciones (Elaboración propia)

Resulta eficiente, emplear modelos para poder clasificar y etiquetar la información de entrada, con ayuda de métodos como: árboles de decisión, con el cual, el modelo será capaz de analizar información a partir de observaciones y encontrar patrones que permitan generar predicciones dirigidas a la respuesta de múltiples preguntas o soluciones de cualquier problema, estos pueden ser agrupados, para manejar métodos Random Forest, que se encarga de encontrar la respuesta a partir de las predicciones que se generan en cada uno de los árboles que lo conforman o Gradient Boosting, encargado de mejorar las predicciones de cada árbol, con la elaboración de uno nuevo, las técnicas de modelado a partir de redes neuronales artificiales también suelen ser muy utilizadas, pues a partir de estas, el modelo es capaz de identificar similitudes en los datos y generar nuevos conocimientos y regresión lineal en el que la salida suele ser un valor numérico, así como también, las máquinas de vector soporte con las que se encuentra una forma óptima de clasificar la información, encontrando sus límites.

Se recomienda manejar los datos ya etiquetados para elaborar un sistema en el que la información aprenda a partir de datos, sin descartar la posibilidad de la creación de un sistema que pueda aprender por sí solo.

5.1.2 Hacer/implementar

Debido a que ejecutar una implementación de las nuevas tecnologías dentro de una empresa puede resultar difícil, el trabajo de un ingeniero industrial es relacionarse con la totalidad de los procesos que en ellos se presentan, es decir, dentro de la estandarización, el jefe de proyecto se encarga de buscar las mejores opciones evitando tomar un exceso de pruebas y erro con el fin de optimizar el proceso del proyecto.

Ahora bien, la programación del software, corresponde totalmente al especialista que se encargara de ello, y debe tomarse en cuenta que este debe tener conocimiento del funcionamiento general de cualquiera de los lenguajes de programación disponibles. Los softwares de modelado que más se utilizan son; Python, y su librería Scikit-learn, R y R STUDIO con paquetes de reglas y Matlab con la librería Toolbox (Tabla 8)

Tabla 8. Costos por lenguaje de programación (Elaboración propia).

Generalidades del estudio económico de los lenguajes de programación	
Python	Es un lenguaje que cuenta con múltiples plataformas de código abierto, por lo cual no tiene ningún costo, su principal ventaja es que permite desarrollar muchos softwares y programas de manera gratuita
MATLAB	Se trata de una plataforma de múltiples estancias en la cual es necesario un monto económico para poder ser ejecutado. El valor de la licencia oscila entre USD 149.00 Y cuenta con diferentes productos de la paquetería de toolbox que rondan entre los USD 45.00

R	Se trata de un lenguaje gratuito con herramientas estadísticas que permiten al usuario dirigir de la mejor manera una determinada cantidad de datos para que puedan ser analizados, se utiliza mayormente por el sector académico e industrial.
---	---

De acuerdo con la tabla, se dirige a tomar una decisión descartando las ventajas y desventajas que se tengan para el cumplimiento de los requisitos y acciones de acuerdo con el área. Dentro de las áreas de productividad, era más recurrente la utilización de las plataformas de PYTHON, debido a que los resultados que se obtuvieron, retomaban alto impacto dentro de los cambios en los procesos.

Es importante mencionar que cuando el rendimiento del modelo no es bueno, la planificación debe reestructurarse, ya que pueden surgir errores o deficiencias en el proceso. La mejor manera de medir el rendimiento del modelo es utilizar cálculos con ayuda de la metodología del error cuadrático medio, esto nos ayudara a analizar el impacto de las mediciones de cada error que se presente al implementar de manera real.

La tecnología, las redes de sensores, los controles informatizados, el software de gestión de la producción y, más en general, el Internet industrial de las cosas (IIoT) son requisitos básicos para que una empresa sea inteligente.

Dentro de este paso, la participación del jefe de proyecto es crucial, pues este, tendrá que monitorear como afecta el funcionamiento del modelo con respecto a las mejoras que se tienen, y se considera la posibilidad de hacer ciertas pruebas conocidas como pruebas piloto, así los

trabajadores podrían iniciar a conocer el nuevo entorno y los cambios para la posible optimización.

También, es recomendable utilizar herramientas como un diagrama de recorrido, este servirá para ejecutar la documentación sobre los registros de los cambios

5.1.3 Verificar

Esta etapa, resulta de la ejecución del registro de las evaluaciones realizadas al finalizar la puesta en marcha del modelo y los resultados arrojados. El error cuadrático medio o RMSE es una medida absoluta de ajuste que ayuda en la identificación de los errores. Es una herramienta interpretable como la desviación estándar que identifica los sucesos que sobresalen de los límites establecidos y que por lo tanto resultan ser poco exitosos o erróneos en el mayor de los casos. Los valores más bajos de RMSE indican un mejor ajuste. Su principal característica dentro de este proceso, es identificar si es posible que el modelo realizado emita las predicciones precisas dentro del proceso en específico o si sucede todo lo contrario.

El especialista en software hace pruebas del modelo en el lenguaje de programación o software de preferencia o elegido por la alta dirección, y este, está sujeto a los cambios que el jefe de proyecto le pida hacer. La dirección del área en específico debe realizar registros de las observaciones que la implantación de esta nueva tecnología provoca, ya sean positivas en el mejor de los casos o negativas.

Para este proceso, se recomienda hacer mediciones de las observaciones y estudios de comparación con respecto a los resultados de lo producido, haciendo un estudio con respecto a los indicadores de medición del desempeño que se tomaran en cuenta para hacer las evaluaciones, como se mencionó al inicio, es indispensable tomar en cuenta los porcentajes

medibles del problema y sus características para que se realice el análisis de los cambios nuevos ejecutados.

Durante la verificación, es importante seguir un recorrido con respecto a las fases que se obtienen generalmente Figura 44. En esta se observan los pasos que se registran para la solución.

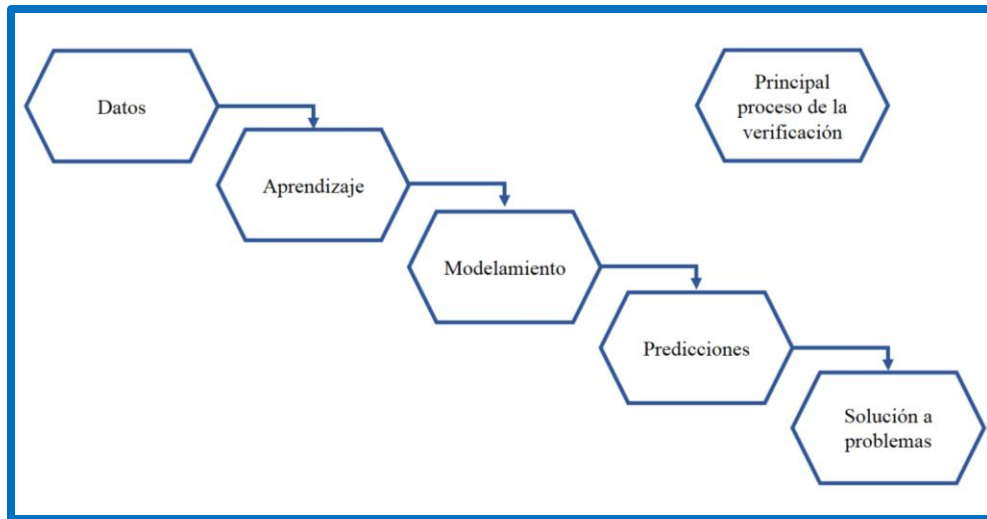


Figura 44. Pasos de verificación

El error cuadrático medio, ayudara como una técnica de evaluación para el jefe del proyecto a emitir la mejor respuesta de acuerdo con el procedimiento. Ya que, de este paso, existe la posibilidad de volver a iniciar incluso con la estandarización

5.1.4 Actuar/Estandarizar

Una vez que el rendimiento del modelo, tenga resultados óptimos a la resolución del problema inicial, este se analiza para saber si es posible o no, implementarlo en diferentes áreas o líneas de la misma área. Se recomienda utilizar diagramas de relación para que se encuentre la ruta correcta a la que le sigue el proceso. Sin embargo, desde un inicio, se debe evaluar que las fallas deben ser corregidas antes de iniciar la implementación, para poder medir la eficacia y eficiencia que el proceso nuevo tenga con respecto a las mediciones tomadas.

Si bien, este paso retoma un proceso difícil, pues el proyecto debe ajustarse a nuevas sucesiones en control a los procesos industriales, de acuerdo con las aplicaciones registradas, se considera que iniciar en áreas de producción puede resultar la mejor manera de incluir las tecnologías para predecir los cambios que existan y que mejoren las áreas. Esto debido a que dentro de las exámenes se concluye que todas las áreas podrían tener relación directa hacia un núcleo considerado la producción principal como se muestra en la figura 45.



Figura 45 Composición de las áreas (Elaboración propia)

Si bien, las áreas pueden sufrir cambios que mejoren el proceso en su totalidad, Machine Learning puede ajustarse a cualquier tipo de problema, los modelos deben ser programados con la mayor cantidad de información posible para evitar que se omita la información eficiente en respuesta a la toma de decisiones

5.2 Conclusiones

Derivado de la revisión de artículos obtenidos de un buscador académico, con el fin de que la información sea clara y coherente con el resultado que se especificó y se logró una clasificación específica de acuerdo con las áreas, algoritmos, y tipo de lenguaje de programación utilizado para poder resumir lo más importante de las implementaciones registradas. Se obtuvo una investigación de las tendencias potenciales e implementaciones en áreas industriales de Machine Learning,

Se realizó un estudio de las áreas en las que mayormente se implementa Machine Learning con éxito. Si bien en este proyecto concluyo que existe mayor rango de publicaciones específicamente en las área de planificación y control de la producción, es crucial considerar esta tecnología en una gran diversidad de áreas, como control de la calidad, mantenimiento, ventas, logística, cadena de suministro, y áreas dirigidas al capital humano, a partir de ellos es posible obtener previsiones y/o predicciones de los que podría suceder en el futuro y así aportar una toma de decisiones argumentada.

Con la revisión general de los principales tipos de algoritmos; aprendizaje supervisado, semi supervisado y no supervisado, se logro la comprensión para ejecutar un mejor modelode toma de toma de decisión. Se encontró que existe mayor tendencia en la aplicación de modelos de aprendizaje supervisado, pues resulta ser una técnica más flexible al momento de entrenar el programa, se considera que la información ya es más concisa y sin errores.

Se dan las características de la propuesta de los lenguajes mayormente recomendables según con los registros que se tienen en las implementaciones: Python, Matlab, R studio y Julia especificando las características de cada uno de ellos y así sea posible identificar el que mejor cumpla con las necesidades de los requerimientos de estandarización del proceso. De esta

manera mostramos que Machine Learning cuenta con una variedad de técnicas que hacen que su aplicación e implementación se de en formas más flexibles puesto que es posible adaptar el modelo según la necesidad de las soluciones al problema. Se determinó que las características del lenguaje de programación Python cumplen con el funcionamiento necesario para ejecutar la realización y activación de un modelo.

5.3 Recomendaciones

La presencia de un reto importante al abrirle las puertas a las nuevas tecnologías dentro del proceso, es cada vez mayor, sin embargo, debido al crecimiento de los avances tecnológicos, es crucial invertir en el desarrollo de fábricas inteligentes o por lo menos en la interferencia del aprendizaje automático en la gestión de las operaciones para el control de procesos industriales.

Existe una gran cantidad de implementaciones documentadas que representan el éxito a partir del uso de las tecnologías de la industria 4.0, y se demuestra con mayor importancia el incremento de este tipo de publicaciones y muy significativamente con mayor hincapié en los últimos años.

Para futuros proyectos de implementaciones de Machine Learning en la industria, se recomienda que se busque que la información o los datos sean ya etiquetados, esto debido a que existe una amplia variedad de alternativas en las que el modelo puede ser ejecutado, y los resultados suelen ser mayormente correctos.

Los lenguajes de programación, son la principal herramienta para la etapa de programación, por lo cual se recomienda hacer uso de la plataforma de libre acceso Python ya que permite que se ejecuten códigos sin la necesidad de aportar un monto económico, así es posible realizar varias experimentaciones y mejorar el rendimiento del modelo. Machine Learning es ahora una herramienta para el control de procesos que permite el paso a la automatización en una amplia variedad de áreas industriales.

Competencias desarrolladas

En la tabla 8 se muestra la descripción generalizada de las competencias adquiridas y aplicadas en este proyecto de investigación.

Tabla 9. Competencias adquiridas (Elaboración propia)

Materia	Competencia desarrollada
Taller de investigación I	Elaborar un protocolo de investigación en el que presenta soluciones científico - tecnológicas a problemáticas relacionadas con su campo profesional en diversos contextos.
Taller de investigación II	Consolidar el protocolo para ejecutar la investigación y obtener productos para su exposición, defensa y gestión de su trascendencia.
Industria 4.0	Conocer las industrias 4.0, características, áreas de aplicación, ventajas y desventajas y criterios de sustentabilidad, así como también, las diferencia entre industrias 3.0 y 4.0.
Simulación	Analizar problemas de líneas de espera, de inventarios, de producción de bienes o servicios, del medio ambiente, de instituciones gubernamentales, para determinar si existen cuellos de botella o sobredimensionamiento en los recursos asignados y mediante la simulación, obtener posibles soluciones, considerando también aspectos sociales, de sustentabilidad y costos.

Algoritmos y lenguaje de programación	Resolver problemas de programación mediante la aplicación de herramientas computacionales para el desarrollo de proyectos.
---------------------------------------	--

Referencias

- Amihai, I., Gitzel, R., Kotriwala, A. M., Pareschi, D., Subbiah, S., & Sosale, G. (2018). An Industrial Case Study Using Vibration Data and Machine Learning to Predict Asset Health. 2018 IEEE 20th Conference on Business Informatics (CBI), 178–185.
<https://doi.org/10.1109/CBI.2018.00028>
- Ayvaz, S., & Alpay, K. (2021). Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. *Expert Systems with Applications*, 173, 114598. <https://doi.org/10.1016/j.eswa.2021.114598>
- Barredo Arrieta, A., Díaz-Rodríguez, N., del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
<https://doi.org/10.1016/j.inffus.2019.12.012>
- Butte, S., Prashanth, A. R., & Patil, S. (2018). Machine Learning Based Predictive Maintenance Strategy: A Super Learning Approach with Deep Neural Networks. 2018 IEEE Workshop on Microelectronics and Electron Devices (WMED), 1–5.
<https://doi.org/10.1109/WMED.2018.8360836>
- Chen, X., & Voigt, T. (2020). Implementation of the Manufacturing Execution System in the food and beverage industry. *Journal of Food Engineering*, 278, 109932.
<https://doi.org/10.1016/j.jfoodeng.2020.109932>

- Chen, Y. (2019, April). Research on engineering quality management based on PDCA Cycle. In IOP Conference Series: Materials Science and Engineering (Vol. 490, No. 6, p. 062033). IOP Publishing.
- Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, 110, 12–22. <https://doi.org/10.1016/j.jclinepi.2019.02.004>
- Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., & de Felice, F. (2020). Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. *Sustainability*, 12(2), 492. <https://doi.org/10.3390/su12020492>
- Delli, U., & Chang, S. (2018). Automated Process Monitoring in 3D Printing Using Supervised Machine Learning. *Procedia Manufacturing*, 26, 865–870. <https://doi.org/10.1016/j.promfg.2018.07.111>
- Dong, G., Li, X., Zhao, J., Su, S., Misra, R. D. K., Xiao, R., & Shang, C. (2020). Machine learning guided methods in building chemical composition-hardenability model for wear-resistant steel. *Materials Today Communications*, 24, 101332. <https://doi.org/10.1016/j.mtcomm.2020.101332>
- Elangovan, M., Sakthivel, N. R., Saravanamurugan, S., Nair, Binoy. B., & Sugumaran, V. (2015). Machine Learning Approach to the Prediction of Surface Roughness Using Statistical Features of Vibration Signal Acquired in Turning. *Procedia Computer Science*, 50, 282–288. <https://doi.org/10.1016/j.procs.2015.04.047>

- Gao, K., Mei, G., Piccialli, F., Cuomo, S., Tu, J., & Huo, Z. (2020). Julia language in machine learning: Algorithms, applications, and open issues. *Computer Science Review*, 37, 100254. <https://doi.org/10.1016/j.cosrev.2020.100254>
- Garre, A., Ruiz, M. C., & Hontoria, E. (2020). Application of Machine Learning to support production planning of a food industry in the context of waste generation under uncertainty. *Operations Research Perspectives*, 7, 100147. <https://doi.org/10.1016/j.orp.2020.100147>
- Henrique da Silva Melo, B., Figueiredo Sales, R., da Silva Bastos Filho, L., Souza Povoas da Silva, J., Gabrielle Carolino de Almeida Sousa, A., Maria Camará Peixoto, D., & Pimentel, M. F. (2022). Handheld near infrared spectrometer and machine learning methods applied to the monitoring of multiple process stages in industrial sugar production. *Food Chemistry*, 369, 130919. <https://doi.org/10.1016/j.foodchem.2021.130919>
- Huo, Z. (2021). Sales Prediction based on Machine Learning. 2021 2nd International Conference on E-Commerce and Internet Technology (ECIT), 410–415. <https://doi.org/10.1109/ECIT52743.2021.00093>
- Islam, S., & Amin, S. H. (2020). Prediction of probable backorder scenarios in the supply chain using Distributed Random Forest and Gradient Boosting Machine learning techniques. *Journal of Big Data*, 7(1), 65. <https://doi.org/10.1186/s40537-020-00345-2>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Javaid, M., Haleem, A., Vaishya, R., Bahl, S., Suman, R., & Vaish, A. (2020). Industry 4.0 technologies and their applications in fighting COVID-19 pandemic. *Diabetes & Metabolic*

Syndrome: *Clinical Research & Reviews*, 14(4), 419–422.

<https://doi.org/10.1016/j.dsx.2020.04.032>

Jennings, C., Wu, D., & Terpenney, J. (2016). Forecasting Obsolescence Risk and Product Life Cycle With Machine Learning. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 6(9), 1428–1439.

<https://doi.org/10.1109/TCPMT.2016.2589206>

Jurkovic, Z., Cukor, G., Brezocnik, M., & Brajkovic, T. (2018). A comparison of machine learning methods for cutting parameters prediction in high speed turning process. *Journal of Intelligent Manufacturing*, 29(8), 1683–1693. <https://doi.org/10.1007/s10845-016-1206-1>

Kanawaday, A., & Sane, A. (2017). Machine learning for predictive maintenance of industrial machines using IoT sensor data. 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), 87–90.

<https://doi.org/10.1109/ICSESS.2017.8342870>

Kolokas, N., Vafeiadis, T., Ioannidis, D., & Tzovaras, D. (2018). Forecasting faults of industrial equipment using machine learning classifiers. 2018 Innovations in Intelligent Systems and Applications (INISTA), 1–6. <https://doi.org/10.1109/INISTA.2018.8466309>

Krauß, J., Frye, M., Beck, G. T. D., & Schmitt, R. H. (2019). Selection and Application of Machine Learning- Algorithms in Production Quality (pp. 46–57).

https://doi.org/10.1007/978-3-662-58485-9_6

Krishna, A., V, A., Aich, A., & Hegde, C. (2018). Sales-forecasting of Retail Stores using Machine Learning Techniques. 2018 3rd International Conference on Computational

Systems and Information Technology for Sustainable Solutions (CSITSS), 160–166.

<https://doi.org/10.1109/CSITSS.2018.8768765>

Lasisi, A., & Attoh-Okine, N. (2018). Principal components analysis and track quality index: A machine learning approach. *Transportation Research Part C: Emerging Technologies*, 91, 230–248. <https://doi.org/10.1016/j.trc.2018.04.001>

Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., & Monostori, L. (2018). Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. *Procedia CIRP*, 72, 1051–1056. <https://doi.org/10.1016/j.procir.2018.03.148>

Min, Q., Lu, Y., Liu, Z., Su, C., & Wang, B. (2019). Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry. *International Journal of Information Management*, 49, 502–519. <https://doi.org/10.1016/j.ijinfomgt.2019.05.020>

Moldovan, D., Cioara, T., Anghel, I., & Salomie, I. (2017). Machine learning for sensor-based manufacturing processes. 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), 147–154. <https://doi.org/10.1109/ICCP.2017.8116997>

Najah Ahmed, A., Binti Othman, F., Abdulmohsin Afan, H., Khaleel Ibrahim, R., Ming Fai, C., Shabbir Hossain, M., Ehteram, M., & Elshafie, A. (2019). Machine learning methods for better water quality prediction. *Journal of Hydrology*, 578, 124084. <https://doi.org/10.1016/j.jhydrol.2019.124084>

Niccolai, A., Caputo, D., Chieco, L., Grimaccia, F., & Mussetta, M. (2021). Machine Learning-Based Detection Technique for NDT in Industrial Manufacturing. *Mathematics*, 9(11), 1251. <https://doi.org/10.3390/math9111251>

- Odabaşı, Ç., & Yıldırım, R. (2020). Machine learning analysis on stability of perovskite solar cells. *Solar Energy Materials and Solar Cells*, 205, 110284.
<https://doi.org/10.1016/j.solmat.2019.110284>
- Okaro, I. A., Jayasinghe, S., Sutcliffe, C., Black, K., Paoletti, P., & Green, P. L. (2019). Automatic fault detection for laser powder-bed fusion using semi-supervised machine learning. *Additive Manufacturing*, 27, 42–53. <https://doi.org/10.1016/j.addma.2019.01.006>
- Quatrini, E., Costantino, F., di Gravio, G., & Patriarca, R. (2020). Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities. *Journal of Manufacturing Systems*, 56, 117–132. <https://doi.org/10.1016/j.jmsy.2020.05.013>
- Rajawat, A. S., Rawat, R., Barhanpurkar, K., Shaw, R. N., & Ghosh, A. (2021). Robotic process automation with increasing productivity and improving product quality using artificial intelligence and machine learning. In *Artificial Intelligence for Future Generation Robotics* (pp. 1–13). Elsevier. <https://doi.org/10.1016/B978-0-323-85498-6.00007-1>
- Robinson, C., Dilkina, B., Hubbs, J., Zhang, W., Guhathakurta, S., Brown, M. A., & Pendyala, R. M. (2017). Machine learning approaches for estimating commercial building energy consumption. *Applied Energy*, 208, 889–904.
<https://doi.org/10.1016/j.apenergy.2017.09.060>
- Sarkar, D., Bali, R., & Sharma, T. (2018). *Practical Machine Learning with Python*. Apress.
<https://doi.org/10.1007/978-1-4842-3207-1>
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>

- Schmitt, J., Bönig, J., Borggräfe, T., Beitinger, G., & Deuse, J. (2020). Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing. *Advanced Engineering Informatics*, 45, 101101. <https://doi.org/10.1016/j.aei.2020.101101>
- Shahbazi, Z., & Byun, Y.-C. (2020). A Procedure for Tracing Supply Chains for Perishable Food Based on Blockchain, Machine Learning and Fuzzy Logic. *Electronics*, 10(1), 41. <https://doi.org/10.3390/electronics10010041>
- Stearns, P. N. (2020). *The Industrial Revolution in World History* (P. N. Stearns, Ed.). Routledge. <https://doi.org/10.4324/9781003050186>
- Turetskyy, A., Wessel, J., Herrmann, C., & Thiede, S. (2021). Battery production design using multi-output machine learning models. *Energy Storage Materials*, 38, 93–112. <https://doi.org/10.1016/j.ensm.2021.03.002>
- Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1), 9. <https://doi.org/10.1186/s40537-016-0043-6>
- Woschank, M., Rauch, E., & Zsifkovits, H. (2020). A Review of Further Directions for Artificial Intelligence, Machine Learning, and Deep Learning in Smart Logistics. *Sustainability*, 12(9), 3760. <https://doi.org/10.3390/su12093760>
- Yadav, G., & Paul, K. (2021). Architecture and security of SCADA systems: A review. *International Journal of Critical Infrastructure Protection*, 34, 100433. <https://doi.org/10.1016/j.ijcip.2021.100433>