

Revisión sistemática de literatura: Técnicas de aprendizaje automático (machine learning)

Anderson Damián Jiménez Alfaro¹ José Vicente Díaz Ospina²

Recibido: 00/00/0000. Aprobado: 00/00/0000.

Resumen. En la actualidad, existe una gran diversidad de modelos que permiten realizar predicciones, y para ello se cuenta con técnicas de Aprendizaje Automático (*Machine learning*) que pueden ayudar a las organizaciones a potenciar sus ventas a través de estos modelos predictivos. En este artículo se efectúa una búsqueda de literatura científica especializada que brinde claridad sobre cuáles son las técnicas más usadas y bajo qué criterios son efectivas. De acuerdo con las necesidades de la investigación, se han filtrado y seleccionado los artículos más relevantes para dilucidar cómo ejecutar un proyecto de machine learning para el pronóstico de ventas. A partir de la revisión realizada se puede afirmar que las diferentes técnicas de machine learning que se encuentran en la literatura son evoluciones de diferentes técnicas conocidas, lo cual es un componente importante para mantener la competitividad empresarial, y bien empleadas podrían convertirse en herramientas potenciadoras de las ventas en las organizaciones.

Palabras claves: *Machine learning*, pronóstico, inteligencia de negocios, márketing, *business management*.

Abstract. Currently, there are a great diversity of models that allow making predictions, and for this there are machine learning techniques that can help organizations to boost their sales through these predictive models. In this article a specialized search of scientific literature is carried out that provides clarity on which are the most used techniques and under what criteria are they effective. According to the research needs, the most relevant articles have been filtered and selected to elucidate how to execute a machine learning project for sales forecasting. From the review carried out, it can be affirmed that the different machine learning techniques found in the literature are evolutions of different known techniques, which is an important component to maintain business competitiveness, and if they are well used could become sales-enhancing tools in companies organizations.

Keywords: Machine learning, forecasting, business intelligence, marketing, business management.

1 Anderson Jiménez. Magíster en Ingeniería Industrial. Universidad Católica Luis Amigó. Medellín, Colombia. Correo electrónico: anderson.jimenezas@amigo.edu.co ORCID: 0000-0001-6601-7911

2 José Vicente Díaz Ospina. Ingeniero Financiero y de Negocios. Universidad Católica Luis Amigó. Medellín, Colombia. Correo electrónico: Jose.diazos@amigo.edu.co ORCID: 0000-0003-0882-5595

Introducción

En la actualidad, se puede encontrar en la literatura un sinnúmero de textos sobre las diferentes técnicas de Aprendizaje Automático (*Machine Learning*), que pueden ser utilizadas para realizar pronósticos en las ventas, lo que en muchas ocasiones genera confusión a la hora de determinar qué técnica es la más apropiada bajo ciertas condiciones o tipos de variables.

El objetivo de este estudio es hacer una búsqueda sistemática de información que permita resolver las preguntas y que brinde claridad sobre el problema de investigación: las técnicas de aprendizaje automático para realizar pronósticos en ventas. La búsqueda de información se hizo en diferentes fuentes de bases especializadas, y de acuerdo con las necesidades de la investigación se filtraron y seleccionaron los artículos más relevantes que ayudaran a determinar dicha claridad.

Las diferentes técnicas de aprendizaje automático que se encuentran en la literatura son la evolución de diferentes técnicas conocidas, lo cual es un componente importante para mantener la competitividad empresarial y que, bien empleadas, podrían convertirse en herramientas potenciadoras de las ventas en las organizaciones.

Marco teórico

Aprendizaje automático (*Machine learning*): el aprendizaje automático, comúnmente abreviado como ML, es un tipo de inteligencia artificial (IA) que “aprende” o se adapta con el tiempo. En lugar de seguir reglas estáticas codificadas en un programa, esta tecnología identifica patrones de entrada y contiene algoritmos que evolucionan con el tiempo [1].

Pronóstico: es la estimación sobre lo que se espera que pueda suceder respecto a una variable con base en un análisis numérico [2].

Inteligencia de negocio: La inteligencia de negocios (BI, por sus siglas en inglés) combina análisis de negocios, minería de datos, visualización de datos, herramientas e infraestructura de datos, así como las prácticas recomendadas para ayudar a las organizaciones a tomar decisiones más basadas en los datos [3].

Márquetin: es el conjunto de actividades y procesos destinados a crear y comunicar valor de marca, identificando y satisfaciendo las necesidades y deseos de los consumidores.

Business management (Gestión empresarial): gestiona la coordinación y organización de las actividades empresariales. Por lo general, esto incluye la producción de materiales, dinero y máquinas, e implica tanto la innovación como el márquetin [1].

Antecedentes

A continuación, se presenta una revisión de los artículos publicados entre los años 2017 a 2021 afines con el uso de métodos de aprendizaje automático para hacer predicciones.

Aruna et al. [1] realizaron una investigación sobre el uso de técnicas de aprendizaje automático a través de aprendizaje no supervisado con un modelo de regresión de bosque aleatorio, para predecir las ventas de medicamentos falsificados. Los resultados del estudio se utilizan para poder contrarrestar esta problemática identificando qué referencias proyectan ventas fraudulentas, el autor aporta algunas referencias que se constituyen en bases para la realización de proyectos de aprendizaje automático que buscan llevar a cabo un pronóstico de ventas.

De acuerdo con los autores, es importante evaluar diferentes modelos de aprendizaje automático, y de los resultados se debe elegir el que mejor se ajuste para la solución del problema, en este caso el algoritmo de regresión de bosque aleatorio.

Los autores concluyen que, con la ayuda de técnicas de aprendizaje automático, es posible predecir las ventas de medicamentos falsificados y, por lo tanto, contrarrestar estas actividades ilegales e inseguras en los países en desarrollo, porque se han convertido en una gran amenaza para las industrias farmacéuticas. El aprendizaje automático es el futuro y la toma de decisiones es más simple, ya que el conocimiento de la máquina es robusto y se puede confiar en los resultados dados por ella con base en el conocimiento previamente adquirido mediante el análisis de los datos.

Mosquera et al. [2] presentan los resultados de una investigación cuyos objetivos fueron:

- Predecir el nivel de riesgo psicosocial en docentes de colegios públicos colombianos utilizando técnicas de inteligencia artificial.
- Definir qué método de clasificación se ajustaba mejor al problema planteado.

En el estudio se tomó una muestra de 5 340 registros epidemiológicos, correspondientes a evaluaciones psicosociales de los docentes. En el tratamiento de los datos se realizaron varias pruebas estadísticas: un análisis de componentes principales con la finalidad de reducir la dimensionalidad de los conjuntos de datos sin perder mucha información de la base y un análisis de varianza con el fin de establecer si existen diferencias significativas en los resultados.

También se realizó un estudio de algoritmos evolutivos, los cuales son técnicas de optimización metaheurística inspiradas en la evolución biológica. Se mantiene una población de soluciones candidatas en cada generación, y cada solución candidata se codifica en un espacio apropiado para aplicar operadores como selección, reproducción y mutación. Se define una función de *fitness* para medir la calidad de los individuos [2].

Luego de la preparación de datos, los autores evaluaron tres algoritmos de aprendizaje automático con los que buscaron llevar a cabo su predicción: Redes neuronales artificiales, *Naive Bayes* y árboles de decisión, con el fin de obtener los resultados de precisión y tasa de error de cada uno. En este caso, los resultados mostraron que el porcentaje de clasificación y tasa de error del algoritmo de redes neuronales artificiales se ajusta más al objetivo del estudio.

El orden de ejecución de esta investigación aporta bases fundamentales para el desarrollo de proyectos de aprendizaje automático que pretendan predecir alguna variable con datos establecidos. Una vez más se identifica la importancia de validar diferentes algoritmos con el fin de evaluar cuál se ajusta más a la necesidad planteada por el investigador, es claro que con datos bien preparados las técnicas de aprendizaje automático son un aliado importante para la toma de decisiones de acuerdo con las predicciones.

Por su parte, Aggarwal et al. [3] evalúan en su investigación seis algoritmos de aprendizaje automático con el fin de identificar qué modelo es más apropiado para predecir la clasificación de estudiantes en función de su rendimiento.

Según los autores, se pueden utilizar varios algoritmos de clasificación para la predicción: *Naive Bayes*, Regresión logística, Máquina de vectores de soporte, Perceptrón multicapa, Árbol de decisión y bosque aleatorio. La evaluación se hace de acuerdo con el conjunto de datos recopilados después de su preprocesamiento. Los autores concluyen que el Perceptrón multicapa y el Bosque aleatorio demuestran ser los clasificadores más apropiados para predecir el desempeño del estudiante.

Lopes et al. [4] plantean en su investigación el desarrollo de un proyecto de aprendizaje automático, usando la metodología CRISP-DM que consta de los siguientes pasos:

- Comprensión de los datos
- Preparación de datos
- Modelado (diferentes modelos)
- Entrenamiento de datos
- Validación de los modelos
- Evaluación de modelos

Los autores señalan que, para cada conjunto de datos obtenido, se construyeron cuatro modelos predictivos utilizando la plataforma H2O integrada en R, se utilizaron los algoritmos de modelos lineales generalizados, Método de refuerzo de gradiente, bosque aleatorio y *deep learning*. Para realizar la predicción, se eligieron los tres primeros, ya que presentaron menos errores en la evaluación [4].

Los autores concluyen que los modelos de aprendizaje automático son herramientas que potencian los procesos bancarios, asistiendo eficazmente a sus clientes con cada producto o servicio.

Henrique et al. [5] proponen en su investigación que la predicción del precio de las acciones es fundamental para la formación de estrategias de inversión y el desarrollo de modelos de gestión del riesgo. Los autores utilizan la siguiente

metodología para el desarrollo de su modelo de aprendizaje automático:

- Extracción de datos
- Preparación de datos
- Modelado por vectores de soporte
- Entrenamiento de datos.
- Variación de entrenamiento de datos
- Validación de los modelos.
- Análisis de resultados

Los autores desarrollan métodos computacionalmente intensivos, utilizando precios pasados para facilitar una mejor gestión del riesgo de mercado para inversores y especuladores. Concluyen que desarrollar modelos de precios predictivos para el mercado de valores es un desafío, pero es una tarea importante a la hora de construir estrategias de transacciones rentables en los mercados financieros.

Justificación

En la actualidad existe una gran diversidad de modelos que permiten realizar algún tipo de pronóstico, por lo que es necesario efectuar una búsqueda de literatura científica especializada que brinde claridad sobre cuáles son las técnicas más usadas y bajo qué criterios son efectivas; las ventas son un tipo de variable-respuesta que se ha estudiado con el fin de determinar el comportamiento del mercado y así ajustar los recursos necesarios para cumplir con la demanda. Por esta razón, el conocimiento de las técnicas de aprendizaje automático para predecir debe ser amplio, con el fin de convertirlas en una herramienta potencial para la organización.

Metodología

Formulación de las preguntas de investigación

- a. ¿Qué técnicas de aprendizaje automático existen en la actualidad y cuáles son sus características?
- b. ¿Cómo se han aplicado las técnicas de aprendizaje automático para hacer pronóstico?
- c. ¿Qué otras técnicas se han aplicado en la predicción de bienes tangibles o intangibles?

Descripción del protocolo de investigación

- d. Empleando la ayuda de Tesouro de la Unesco, se buscan las palabras claves y relevantes que permitan llevar a cabo una revisión de literatura más próxima a la investigación.
- e. Se ingresa a Scopus para hacer una búsqueda de los autores y revistas especializados en el tema propuesto.
- f. El filtrado se realiza de acuerdo con las necesidades de la investigación:
 - Correspondan al periodo de los últimos 5 años
 - Artículos finales
 - Metodología clara
- g. Selección de los artículos con aportes más relevantes a la investigación.
- g. Discusión de los artículos.
- i. Conclusiones.

Proceso de búsqueda de documentos

- j. Tormenta de ideas
 - Pronósticos de venta basados en técnicas de IA.
 - Modelos de toma de decisiones basados en Inteligencia de negocios.
 - Comparación de los diferentes modelos de tomas de decisiones aplicados al sector real.
 - Diseño de una visualización de datos para la toma de decisiones con base en inteligencia de negocios: Caso de estudio.
- k. Palabras claves:
 - Machine learning
 - Predicción
 - Ventas
 - Inteligencia de negocios
- l. Búsqueda de las palabras claves en las plataformas especializadas de investigación:
 - Codificación
 - Método de aprendizaje
 - Presentación de estadísticas
 - Recopilación de datos
 - Visualización de datos
 - Procesamiento de datos
- m. Ecuaciones de búsqueda.
 - (TITLE-ABS-KEY (machine AND learning) ANDTITLE-ABS-KEY (forecast) AND TITLE-ABS-KEY (marketing)) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR,

- 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016)) AND (LIMIT-TO (DOCTYPE, "ar"))
- TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
 - TITLE-ABS-KEY ("Machine Learning" AND "Supervised Learning" AND "Unsupervised Learning") AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017)) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SRCTYPE , "j"))
- Proceso de selección de documentos**
- Si las respuestas a las siguientes preguntas son afirmativas, se establece que el artículo académico es relevante para la investigación:
- n. El título del artículo académico se relaciona con alguna de las temáticas de la investigación.
 - o. El abstract del artículo académico se relaciona con alguna de las temáticas de la investigación.
 - p. Las conclusiones del artículo académico se relacionan con alguna de las temáticas de la investigación.
 - q. Algunos de los subtítulos del artículo se relacionan con alguna de las temáticas de la investigación.
 - r. Algunas referencias bibliográficas del artículo se relacionan con alguna de las temáticas de la investigación.
- Proceso de evaluación de la calidad**
- Los criterios de calidad empleados para la selección de documentos fueron:
- Tener un número de citas de mínimo 35 artículos.
- s. Metodología de aplicación clara.
 - t. Que tenga cohesión.
 - u. Estadística y metodología detalladas.
 - v. Que no existan errores en la argumentación del autor.

w. Algunos de los argumentos no se relacionan con las posturas o propuestas del autor.

Proceso de extracción de datos

Se realiza la lectura completa de cada uno de los documentos seleccionados y se efectúan las preguntas de investigación planteadas con anterioridad para determinar cuál o cuáles interrogantes son contestados por el texto, después se extrae la información que da respuesta clara y precisa a las preguntas para un posterior análisis.

Resultados Proceso de búsqueda

En la Tabla 1 se muestran los resultados obtenidos de la búsqueda. corregir tabla

Tabla 1. Resultados del proceso de búsqueda.

Número de Ecuación	Número de elementos en la búsqueda
1	23
2	50
3	35
4	20
5	89
6	40

Fuente: elaboración propia (2021)

Discusión

En la actualidad existe una gran cantidad de técnicas de aprendizaje automático, entre las más conocidas encontramos el Perceptrón Multicapa (MLP), una técnica con estructura flexible que permite relacionar variables no lineales; por otro lado, encontramos diferentes tipos de redes neuronales artificiales como los mapas autoorganizados de Kohonen (SOM), el codificador automático (AuNN) o la autoorganizada basada en errores (SOEDNN), entre otras. Estos son sistemas de redes que ajustan sus parámetros de pesos y umbrales con el fin de producir las salidas deseadas a partir de las entradas establecidas, también son sistemas de aprendizaje sin supervisión que facilitan la detección de datos asociados que no tienen un

comportamiento lineal en series temporales, lo que les permite ser de utilidad en la predicción de datos económicos y financieros. También se encuentran los árboles de decisiones (DT), los cuales tienen un enfoque de clasificación no paramétrico y han sido de uso común en el campo de la calificación crediticia, los resultados predictivos de DT se pueden representar como un gráfico acíclico recitado en forma de árbol, por lo que los resultados del modelo pueden ser fácilmente entendidos por los humanos. Los árboles de regresión son capaces de estimar apareamientos de las relaciones entre variables, depende de modelar la relación funcional entre un dependiente variable (objetivo) y una o más variables independientes (predictores), la regresión constituye una poderosa estadística herramienta de predicción y previsión de un valor continuo objetivo, dado un conjunto de predictores. K-means (KNN) se basa en el concepto de centroides y es más utilizado como algoritmo de agrupamiento. El atractivo de K-means es que incluye simplicidad, escalabilidad y versatilidad; todos los métodos estadísticos son codificados de manera que procesan el mismo conjunto de datos para construcción de modelos y validación de conjuntos de prueba [4]-[10].

Resultados y discusión

Para realizar pronósticos y predicciones en diferentes ámbitos se están utilizando diferentes métodos de aprendizaje automático, por ejemplo, el modelo de Bosque aleatorio sirvió para predecir las ventas de medicamentos falsificados con el fin de contrarrestar esta problemática, entregando resultados que sirven para ejecutar acciones contra este comercio [1], este modelo no solo ha sido utilizado para predecir ventas, también se encontraron trabajos dedicados a la predicción de calificaciones de estudiantes basados en su rendimiento, donde la técnica de bosque aleatorio se ajustaba más y mostraba mejores resultados que otros métodos analizados como, Bayes, Regresión aleatoria, Máquina de Vectores de soporte, Árboles de decisión y Perceptrón multicapa, siendo este último un modelo que también sobresale para este tipo de pronósticos [3]. Otros autores han utilizado árboles de decisión para predecir niveles de deserción de empleados en la organización [13] y pronosticar los riesgos psicosociales en los docentes de colegios públicos colombianos [2].

La aplicación de técnicas predictivas de aprendizaje automático en el campo médico ha sido muy importante, trabajos como la predicción de pacientes con enfermedad del corazón a través de una Red Neuronal han demostrado que se puede ser más eficiente con los diagnósticos, procesando adecuadamente los datos de los pacientes a través de este algoritmo con gran cantidad de variables [14], también se han ejecutado métodos de aprendizaje automático para la predicción de artritis reumatoide en edad temprana procesando datos de pacientes a través de K-Means [15], y aunque en la mayoría de modelos del sector médico se utilizan para predecir enfermedades, también podemos encontrar trabajos donde estas técnicas se aplican para identificar las tendencias en salud en las redes sociales, un ejemplo es el modelo desarrollado en [16], donde a través de Naive Bayes procesaron los datos tomados de opiniones y actitudes de los usuarios en redes sociales para predecir estas tendencias y tomar decisiones sobre los servicios que se debían ofrecer.

Los modelos de aprendizaje automático también se han utilizado para predecir el valor de las acciones por minuto en Brasil, donde se tomaron los principales indicadores financieros de esta localidad y dos años de datos bursátiles para lograr esta predicción [5], con excelentes resultados. El trabajo en [4] propone la construcción de clasificadores basados en la aplicación de modelos de aprendizaje automático para predecir futuros incumplimientos de pago y clientes con potencial recuperación crediticia. En [17] se encuentra un ejemplo de aplicación de técnicas supervisadas de aprendizaje automático en el sector agrícola, donde se elaboró un modelo basado en Kernel para predecir el rendimiento de la caña de azúcar en la India.

Todos los modelos mencionados anteriormente tienen una metodología en común la cual se relaciona a continuación:

- y.** Procesamiento o preparación de datos
- z.** Elección de Modelo o Algoritmo ML
- aa.** Aplicación de algoritmos para reducción de la dimensionalidad
- bb.** Entrenamiento y evaluación con las técnicas de aprendizaje elegidas

cc. Análisis de estabilidad

Esto demuestra que los métodos de predicción de aprendizaje automático se han aplicado en diferentes ámbitos y que sus resultados suelen ser buenos cuando se sigue un proceso de implementación adecuado, empezando por la comprensión de los datos y su procesamiento y eligiendo el algoritmo adecuado de acuerdo con el requerimiento, para garantizar un buen entrenamiento de la herramienta y los ajustes necesarios al modelo final [1], [18].

Se sigue evolucionando permanentemente, en esencia, la base de muchos nuevos modelos sigue siendo los aquí mencionados, que se ajustan o reforman para que se adapten a un contexto específico. Entre estos encontramos FAIR que está diseñado para tratar la gran dimensionalidad de las variables explicativas y la confusión inducida por la regularización, al tiempo que incorpora efectos de series temporales heterogéneas y dinámicas. Las previsiones de FAIR pueden descomponerse en lo que se espera sobre la base de patrones históricos, el impacto de la comercialización propia y de categorías cruzadas, y el impacto de las perturbaciones aleatorias en las ventas de base [12], [19]-[26].

Conclusiones

Con base en los artículos consultados y su discusión se concluye lo siguiente:

Existen tres tipos de aprendizaje automático: aprendizaje supervisado, aprendizaje sin supervisión y aprendizaje de refuerzo, de estos se desprenden diversos algoritmos de aprendizaje automático que podemos utilizar de acuerdo con nuestro requerimiento y nivel de ajuste del modelo. Entre los más comunes se destacan: algoritmos de regresión, algoritmos bayesianos, algoritmos de agrupación, algoritmos de árboles de decisión, algoritmos de redes neuronales, algoritmos de reducción de dimensión y algoritmos de aprendizaje profundo.

Las predicciones y pronósticos a través de técnicas de aprendizaje automático se han aplicado en diferentes ámbitos, y han beneficiado sectores como la salud, la educación, las finanzas, la

producción y la agricultura, ya que sus resultados ayudan a tomar decisiones para mejorar de forma continua resultados, ventas, servicio al cliente y percepción de valor.

Las técnicas mantienen una dinámica evolutiva, lo que permite encontrar una gran variedad de técnicas que se pueden aplicar o transformar dependiendo de las variables de entrada y salida, por otra parte, se debe buscar la minimización de las perturbaciones que se pueden generar al interior del modelo.

Trabajo futuro

Se propone llevar a cabo una investigación sobre cómo predecir las ventas por referencia de un negocio cárnico utilizando modelos de aprendizaje automático, mediante implementaciones de pronóstico de la demanda y pedidos sugeridos para la ejecución del proyecto.

Referencias

- [1] R. Aruna Florence, S. Bethu, V. Sowmya, K. Anusha, y B. Sankara Babu, "Importance of supervised learning in prediction analysis", *Period. Eng. Nat. Sci.*, vol. 6, n.º 1, pp. 201-214, 2018.
- [2] R. Mosquera, O. D. Castrillón, y L. Parra, "Predicción de riesgos psicosociales en docentes de colegios públicos colombianos utilizando técnicas de inteligencia artificial", *Inf. Tecnológica*, vol. 29, n.º 4, pp. 267-280, 2018.
- [3] D. Aggarwal, S. Mittal, y V. Bali, "Prediction model for classifying students based on performance using machine learning techniques", *Int. J. Recent Technol. Eng.*, vol. 8, n.º 2, Special Issue 7, pp. 496-503, 2019.
- [4] R. G. Lopes, M. Ladeira, y R. N. Carvalho, "Use of machine learning techniques in the prediction of credit recovery", *Adv. Sci. Technol. Eng. Syst.*, vol. 2, n.º 3, pp. 1432-1442, 2017.
- [5] B. M. Henrique, V. A. Sobreiro, y H. Kimura, "Stock price prediction using support vector regression on daily and up to the minute prices", *J. Financ. Data Sci.*, vol. 4, n.º 3, pp. 183-201, 2018.
- [6] G. I. Ahmad y J. Singla, "Machine learning techniques for sentiment analysis of indian languages", *Int. J. Recent Technol. Eng.*, vol. 8, n.º 2, Special Issue 11, pp. 3630-3636, 2019.
- [7] S. Athey y G. W. Imbens, "Machine Learning Methods Economists Should Know About", Stanford Graduate School of Business", *Work. Pap.*, n.º 3776, March, 2019.
- [8] Z. Ge, Z. Song, S. X. Ding, y B. Huang, "Data Mining and Analytics in the Process Industry: The Role of Machine Learning", *IEEE Access*, vol. 5, pp. 20590-20616, 2017.
- [9] Z. Gong, P. Zhong, y W. Hu, "Diversity in Machine Learning", *IEEE Access*, vol. 7, pp. 64323-64350, 2019.
- [10] O. Simeone, "A Very Brief Introduction to Machine Learning with Applications to Communication Systems", *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, n.º 4, pp. 648-664, 2018.
- [11] M. Usama et al., "Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges", *IEEE Access*, vol. 7, pp. 65579-65615, 2019.
- [12] J. Wang, C. Jiang, H. Zhang, Y. Ren, K. C. Chen, y L. Hanzo, "Thirty Years of Machine Learning: The Road to Pareto-Optimal Wireless Networks", *IEEE Commun. Surv. Tutorials*, vol. 22, n.º 3, pp. 1472-1514, 2020.
- [13] J. Vasa y K. Masrani, "Foreseeing employee attritions using diverse data mining strategies", *Int. J. Recent Technol. Eng.*, vol. 8, n.º 3, pp. 620-626, 2019.
- [14] O. Terrada, S. Hamida, B. Cherradi, A. Raihani, y O. Bouattane, "Supervised

- machine learning based medical diagnosis support system for prediction of patients with heart disease", *Adv. Sci. Technol. Eng. Syst.*, vol. 5, n.º 5, pp. 269-277, 2020.
- [15] B. Jayanthi y C. Senthamarai, "Feature selection using K-means genetic clustering to predict rheumatoid arthritis disease", *Int. J. Recent Technol. Eng.*, vol. 8, n.º 3, pp. 7020-7023, 2019.
- [16] S. Saini, S. P. Singh, y R. Agarwal, "Augmented machine learning ensemble extension model for social media health trends predictions", *Int. J. Recent Technol. Eng.*, vol. 8, n.º 2 Special Issue 7, pp. 482-486, 2019.
- [17] R. Medar y V. S. Rajpurohit, "Supervised machine learning techniques for predicting sugarcane yield", *Int. J. Recent Technol. Eng.*, vol. 8, n.º 2, pp. 5662-5668, 2019.
- [18] R. Jafari-Marandi, "Supervised or unsupervised learning? Investigating the role of pattern recognition assumptions in the success of binary predictive prescriptions", *Neurocomputing*, vol. 434, pp. 165-193, 2021.
- [19] W. Bao, N. Lianju, y K. Yue, "Integration of unsupervised and supervised machine learning algorithms for credit risk assessment", *Expert Syst. Appl.*, vol. 128, pp. 301-315, 2019.
- [20] A. Martínez, C. Schmuck, S. Pereverzyev, C. Pirker, y M. Haltmeier, "A machine learning framework for customer purchase prediction in the non-contractual setting", *Eur. J. Oper. Res.*, vol. 281, n.º 3, pp. 588-596, 2020.
- [21] S. Mohanapriya y S. Mohana Saranya, "Sales prediction using machine learning algorithm", *Int. J. Adv. Sci. Technol.*, vol. 29, n.º 3 Special Issue, pp. 1049-1055, 2020.
- [22] D. Raditya, E. P. Nicholas, A. S. Ferarida, y N. Hanafiah, "Predicting Sneaker Resale Prices using Machine Learning", *Procedia Comput. Sci.*, vol. 179, pp. 533-540, 2021.
- [23] J. Salminen, V. Yoganathan, J. Corporan, B. J. Jansen, y S. G. Jung, "Machine learning approach to auto-tagging online content for content marketing efficiency: A comparative analysis between methods and content type", *J. Bus. Res.*, vol. 101, September 2018, pp. 203-217, 2019.
- [24] Z. Shahbazi, D. Hazra, S. Park, y Y. C. Byun, "Toward improving the prediction accuracy of product recommendation system using extreme gradient boosting and encoding approaches", *Symmetry (Basel)*, vol. 12, n.º 9, 2020.
- [25] A. Sreekant, P. Senthilnathan, G. Gopichand, M. Rajapandy, y N. Kannan, "Necessity of machine learning and data visualization principles in marketing investment management", *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, n.º 6, Special Issue 4, pp. 527-530, 2019.
- [26] T. Van Nguyen, L. Zhou, A. Y. L. Chong, B. Li, y X. Pu, "Predicting customer demand for remanufactured products: A data-mining approach", *Eur. J. Oper. Res.*, vol. 281, n.º 3, pp.543-558, 2020.