

**Diseño de un modelo predictivo basado en redes neuronales para el total de ventas de una
empresa del sector comercio**

Sahara Fernanda Benavides Pasmíño

Asesor

Rafael Roberto Ruiz Escorcía

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Ciencia de Datos y Analítica

2025

Nota de Aceptación

Nombre director de Trabajo de Grado

Jurado

Jurado

Agradecimientos

En primer lugar, agradezco a Jehová, Dios, por haberme brindado la oportunidad de emprender este camino y por abrirme las puertas hacia una vida mejor. Su guía y fortaleza han sido fundamentales para lograrlo.

A mi madre, Alicia, una mujer guerrera y fuerte, quien con su amor incondicional y su apoyo constante me ha dado las fuerzas necesarias para seguir adelante. Este logro es también suyo. Agradezco a mi familia, amigos y a mi asesor quienes han contribuido a que este sueño se haga realidad sus consejos y su apoyo han sido muy útiles durante este proceso.

Resumen

Este proyecto tiene como objetivo diseñar un modelo predictivo utilizando técnicas de aprendizaje profundo, orientado a una empresa del sector comercio tecnológico. Se busca obtener una buena precisión del modelo basándose principalmente en la métrica MAE, además realizar un análisis de las ventas, permitiendo tomar decisiones informadas y más asertivas por parte de la administración. Para su desarrollo se adoptará un enfoque en la metodología CRISP-DM, ajustando sus fases al proyecto. Esto proporciona una estructura que facilita la realización de ajustes eficientes durante el proceso. El modelo estará basado en redes neuronales LSTM (Long Short- Term Memory), reconocidas por su alta precisión en tareas de predicción con series de tiempo. La red se entrenará con datos históricos diarios de ventas de la empresa comprendidos de 2021 a 2024, y se analizarán características relevantes como días de la semana y temporada que se deben incluir. Se desarrolla este proyecto con el fin de proporcionar una herramienta eficaz para la gestión de inventarios y diseños de estrategias de venta, además de ofrecer un enfoque innovador que utiliza herramientas avanzadas de la inteligencia artificial para beneficio de la empresa o de otro sectores.

Palabras clave: Análisis predictivo, aprendizaje profundo, red neuronal, inteligencia artificial, LSTM.

Abstract

This project aims to design a predictive model using deep learning techniques, tailored for a company in the technology commerce sector. The goal is to achieve high model accuracy, primarily based on the MAE (Mean Absolute Error) metric, while also conducting a sales analysis to enable informed and more assertive decision-making by management team. For its development, the CRISP-DM methodology will be employed. This methodology provides a structured approach and facilitates efficient adjustments throughout the process. The model will be based on LSTM (Long Short-Term Memory) neural networks, which are renowned for their high accuracy in time series prediction tasks. The network will be trained on the company's historical daily sales data from 2021 to 2024, and relevant features such as days of the week and seasonality will be analyzed for inclusion. This project is developed to provide an effective tool for inventory management and sales strategy design. Additionally, it offers an innovative approach that leverages advanced artificial intelligence tools for the benefit of the company and others sectors.

Keywords: Predictive analysis, deep learning, neural network, artificial intelligence, LSTM

Tabla de Contenido

Introducción	11
Planteamiento del Problema	13
Justificación	14
Objetivos	15
Objetivo General	15
Objetivos Específicos.....	15
Marco de Referencia	16
Estado del Arte.....	16
Marco Contextual.....	19
Marco Teórico.....	20
Marco Conceptual.....	35
Metodología	37
Inicio del Proyecto	38
Recolección de Datos	40
Preparación de Datos.....	40
Modelamiento de Datos	50
Entrenamiento	52
Resultados	54
Primera Red Neuronal.....	55
Evaluación Red Neuronal General Uno	60
Predicciones Red General Uno.....	61
Segunda Red Neuronal.....	62

Evaluación Red Neuronal Dos	68
Tercer Red Neuronal	69
Evaluación Red Neuronal Tres.....	74
Conclusiones.....	76
Recomendaciones	78
Referencias Bibliográficas	80

Lista de Tablas

Tabla 1 <i>Comparación de Facturas Emitidas Desde 2021 a 2024</i>	44
Tabla 2 <i>Top 5 de Grupos de Productos del Año 2024</i>	45
Tabla 3 <i>Tabla de Parámetros de Redes Neuronales</i>	54
Tabla 4 <i>Resumen de Modelo de Red Neuronal General</i>	55
Tabla 5 <i>Métricas RMSE, MAE y R^2 de Red Neuronal Uno</i>	61
Tabla 6 <i>Resumen de Modelo de Red Neuronal Dos</i>	63
Tabla 7 <i>Métricas RMSE, MAE y R^2 de Red Neuronal Dos</i>	69
Tabla 8 <i>Resumen de Modelo de Red Neuronal Tres</i>	70
Tabla 9 <i>Métricas RMSE, MAE y R^2 de Red Neuronal Tres</i>	75

Lista de Figuras

Figura 1 <i>Línea del Tiempo de Datos Más Sobresalientes de la IA Hasta Nuestros Días</i>	20
Figura 2 <i>Estructura de Técnicas de Machine Learning</i>	22
Figura 3 <i>Disciplinas de la Inteligencia Artificial</i>	28
Figura 4 <i>Anatomía de Neurona Biológica</i>	29
Figura 5 <i>Modelo de Neurona Estándar (Perceptrón)</i>	30
Figura 6 <i>Funciones de Activación</i>	31
Figura 7 <i>Estructura Red Neuronal Monocapa</i>	32
Figura 8 <i>Estructura Red Neuronal Multicapa</i>	33
Figura 9 <i>Estructura de RNN</i>	33
Figura 10 <i>Proceso de Creación de Modelo de Metodología CRISP-DM</i>	38
Figura 11 <i>Top 5 de Ranking de Lenguajes de Programación Más Utilizados</i>	39
Figura 12 <i>Ventas Diarias Antes de Ponderar los Valores Extremos</i>	41
Figura 13 <i>Aporte de Ventas por Vendedor(%)</i>	42
Figura 14 <i>Comparación Mensual de Monto de Ventas Desde 2021 a 2024</i>	43
Figura 15 <i>Comparación Ventas Anual por Sede</i>	43
Figura 16 <i>Comparación de Datos Original y Suavizado con Kalman Segmentación Uno</i>	46
Figura 17 <i>Comparación de Datos Original y Suavizado con Kalman Segmentación Dos</i>	47
Figura 18 <i>Comparación de Datos Original y Suavizado con Kalman Segmentación Tres</i>	48
Figura 19 <i>Construcción del Modelo LSTM Ejecutado del Proyecto</i>	51
Figura 20 <i>Entrenamiento de Red Neuronal Global</i>	53
Figura 21 <i>Entrenamiento de Red Neuronal Dos</i>	53
Figura 22 <i>Entrenamiento de Red Neuronal Tres</i>	53

Figura 23 <i>Código de Realización de Predicciones</i>	54
Figura 24 <i>Gráfica de Aprendizaje Red Uno</i>	56
Figura 25 <i>Datos Reales Vs Datos Predichos de Entrenamiento</i>	57
Figura 26 <i>Datos Reales Vs Datos de Validación</i>	58
Figura 27 <i>Datos Reales Vs Datos de Prueba</i>	59
Figura 28 <i>Gráfica Completa Entrenamiento Real Más Predicciones de Validación y Test</i>	59
Figura 29 <i>Predicción de Fechas Futuras Red General</i>	61
Figura 30 <i>Predicción Futura</i>	62
Figura 31 <i>Gráfica de Aprendizaje Red Dos</i>	64
Figura 32 <i>Datos Reales Vs Datos Predichos de Entrenamiento</i>	65
Figura 33 <i>Comparación Conjunto de Validación Vs Valores Reales</i>	66
Figura 34 <i>Comparación Base de Datos Test Vs Valores Originales</i>	67
Figura 35 <i>Datos Continuos Entrenamiento Reales Más Predicción Validación y Test</i>	67
Figura 36 <i>Gráfica de Aprendizaje Red Neuronal Tres</i>	71
Figura 37 <i>Datos Reales y Datos Predichos de Entrenamiento</i>	72
Figura 38 <i>Comparación de Validación con Datos Reales</i>	73
Figura 39 <i>Comparación de Prueba con Datos Reales</i>	73
Figura 40 <i>Comparación de Validación con Datos Reales</i>	74

Introducción

El avance tecnológico y el auge de la inteligencia artificial en los últimos años, ha llevado que la mayoría de las empresas deban adaptarse a nuevas tecnologías para seguir siendo competitivas en el mercado. La revolución digital ha llevado a que las organizaciones cambien la forma de operar, volviendo a los datos un recurso fundamental para la toma de decisiones estratégicas. Por esto las empresas buscan conocer información clave para optimizar procesos, mejorar servicios y tomar mejores decisiones.

Los datos no solo son importantes, sino que cada optimización de técnicas o herramientas permiten a las empresas obtener una visión más detallada de sus operaciones. Y un desafío en la mayoría de las empresas es la gestión de inventarios y capacidad de satisfacer la demanda del mercado de manera correcta. Por este motivo es necesario implementar técnicas avanzadas de aprendizaje profundo como las redes neuronales LSTM o Long Short-Term-Memory, esta técnica nos ayudará a diseñar un modelo capaz de predecir la demanda con mayor precisión y escalabilidad aportando una mejora financiera para las empresas.

Por lo tanto, este proyecto tiene como objetivo diseñar un modelo predictivo basado en redes neuronales LSTM, para estimar las ventas de la empresa ubicada en el sector comercio tecnológico, Se realizará basándose en los datos históricos diarios de ventas de cuatro años, además de realizar el análisis correspondiente sobre que variables exógenas añadir. Llevar a cabo este proyecto es buscar la eficiencia operativa de las empresas, y reducir costos asociados a escasos o exceso de inventarios.

El proyecto contiene en primer lugar una revisión literaria sobre técnicas de predicción y conceptos claves, necesarios para comprender la construcción de una red neuronal. Posteriormente se describe la metodología utilizada y como esta nos beneficia realizar ajustes al

modelo, además se presentan los resultados gráficos y con métricas de los modelos, para finalmente terminar con las recomendaciones encontradas al realizar este proyecto.

Planteamiento del Problema

En el sector del comercio electrónico, la alta competencia y rápida rotación de productos exigen a las empresas tener desarrollado una gestión de inventarios eficiente y robusta. Sin embargo, la mayoría de las empresas realizan el proceso de verificación de inventario de manera manual y basan sus pronósticos de demanda en movimientos mensuales anteriores.

Este enfoque tan tradicional no es suficiente para lograr obtener los datos de manera precisa, lo que deriva a tener dos escenarios críticos, el exceso de inventario, que genera incrementos de costos de almacenamiento, endeudamiento sin retorno y alto riesgo de obsolescencia, y la falta de stock que conduce a clientes insatisfechos, pérdida de clientes y disminución de las ganancias. En ambos escenarios tiene un impacto negativo en la rentabilidad y la competitividad de las empresas, además de afectar su buen nombre y posicionamiento en el mercado.

Dado que los estudios demuestran que los métodos tradicionales de pronóstico no son suficientes, se propone diseñar un modelo predictivo basado en redes neuronales LSTM (Long Short- Term Memory), este tipo de red es capaz de capturar estacionalidad, complejidad y patrones no evidentes de los datos suministrados, todo esto con el fin de optimizar la precisión en las predicciones de la demanda y la gestión de inventarios.

Justificación

Para cubrir la necesidad de una empresa en el sector comercio se ve necesario desarrollar una estimación de ventas que nos permita prever la demanda de manera más precisa y confiable.

El diseño de un modelo predictivo de ventas basado en redes neuronales nos permitirá brindar a la empresa la capacidad de obtener estimaciones de ventas. Esto ayudará a optimizar la gestión de inventario, una utilización óptima de los espacios de almacenamiento y mejorará la rotación de productos. Además, este método contribuirá a una mayor satisfacción del cliente, ya que se garantizará la disponibilidad de productos.

Finalmente, la empresa se beneficiará al tener a disposición herramientas que le permitan tomar decisiones estratégicas informadas, lo que mejorará su rentabilidad, su eficiencia operativa y su sostenibilidad en el mercado.

Objetivos

Objetivo General

Crear un modelo de análisis predictivo utilizando técnicas de redes neuronales para mejorar la precisión en la estimación de ventas, permitiendo realizar una planificación estratégica y optimizando la toma de decisiones comerciales.

Objetivos Específicos

Elaborar un análisis exploratorio de los datos para identificar información clave que permita a la empresa elaborar objetivos y mejoras.

Diseñar y entrenar el modelo de predicción con la red neuronal LSTM usando datos históricos de ventas, número de grupo y cantidad vendida.

Evaluar el modelo mediante pruebas de validación como distancia media cuadrática mínima y error absoluto medio (RMSE, MAE), asegurando su eficiencia, precisión y capacidad de generalización para una futura implementación.

Marco de Referencia

Estado del Arte

La tecnología ha sido un pilar fundamental para el crecimiento económico global. En el sector empresarial, la implementación de herramientas basadas en la Inteligencia Artificial (IA) ha logrado optimizar procesos y obtener mejores resultados. Una de las claves del éxito para las empresas es la capacidad de poder anticipar y satisfacer la demanda y es posible con técnicas de Machine Learning o Aprendizaje supervisado.

La técnica de predicción de ventas ha ido evolucionando en los últimos años y hemos pasado por métodos tradicionales, como regresión lineal, modelos de series temporales como ARIMA, SARIMA, SARIMAX y hemos llegado a métodos más complejos y precisas basadas en aprendizaje profundo como las Redes Neuronales Recurrentes (RNN) y las redes Long Short Term Memory (LSTM).

Revisaremos algunos trabajos relacionados con el tema de para obtener una visión más clara y una buena fundamentación para nuestro problema.

La investigación titulada “Comparación de predicción basada en redes neuronales contra métodos estadísticos en el pronóstico de ventas” de Ruelas y Laguna (2014), nos muestra que los métodos estadísticos convencionales para pronósticos dependen de los datos históricos y asumen que el futuro seguirá el comportamiento pasado. Sin embargo, si alguna variable cambia, estos métodos se vuelven ineficaces ya que suponen que la población siempre tendrá una distribución normal. Por otra parte, las redes neuronales no requieren especificar distribución de probabilidad o patrón de comportamiento, ni que los datos sean analizados para verificar tendencias o estacionalidad, dado que aprenden a partir de los ejemplos proporcionados, en su artículo

menciona que las redes neuronales tuvieron mejor desempeño basándose en la métrica del error cuadrático medio (MSE).

Ensafi, Y. et, al. (2022) en la investigación titulada “Time-series forecasting of seasonal items sales using machine Learning – A comparative analysis” nos comenta lo crucial que es la previsión de ventas en la industria, especialmente las que manejan productos de temporada. Evaluó los diversos modelos de redes neuronales para prever ventas de artículos de temporada y los comparo con métodos tradicionales. Se destacó los modelos de redes neuronales principalmente el LSTM como el más eficaz. Cabe resaltar que el modelo de CNN, mostró un rendimiento notable. Además, el modelo Prophet, enfocado en generar pronósticos rápidos y precisos, resultó muy eficiente sobre todo cuando se incorpora el factor de vacaciones. Nos recomienda para futuros estudios podrían explorar LSTM y CNN más complejos, pronósticos multivariantes y el desarrollo de modelos híbridos que combinen enfoques clásicos y modernos para mejorar aún más los pronósticos.

En la investigación titulada “Series Temporales Avanzadas: Aplicación de redes Neuronales para el Pronóstico de Series de Tiempo” de Fernández Salguero (2021), nos cuenta que la prevención de series de tiempo financieras se ha basado en modelos económicos y estos presentan dificultades como los sobreajustes y la incapacidad para manejar grandes volúmenes de datos ruidosos. Con el aprendizaje profundo (Deep Learning), con las redes neuronales multicapa ha generado una solución efectiva, sin dejar a un lado los modelos tradicionales que siguen siendo útiles para problemas más simples.

En la investigación “Modelo basado en redes neuronales recurrentes LSTM para la predicción de la siguiente actividad en procesos de negocio” de Ramírez, U, et al. (2018), nos presenta una aplicación innovadora en el campo de la minería de procesos al utilizar redes

neuronales LSTM para procesar datos de registros de eventos en procesos de negocio, se demostró que el método es efectivo para predecir actividades futuras, obteniendo una tasa de exactitud aceptable en todos los casos de estudio, tanto sintéticos como reales.

En el artículo de Zhao Kui, et al. (2017), titulado “Sales Forecast in E-commerce using convolutional Neural network”, nos muestra un enfoque novedoso de obtener características a partir de datos estructurados usando redes neuronales convolucionales, disminuyendo tiempo y evitando complicaciones al hacerlo de manera manual. Prueba la metodología y obtiene un sólido rendimiento. Recomendando que si es posible identificar los indicadores más importantes para pronóstico de ventas. Y añade como reto proponer desarrollar un marco unificado que permita extraer de manera automática características de los datos.

Salazar. y Cabrera, M. (2007) en su investigación titulada “Pronóstico de demanda por medio de redes neuronales artificiales” donde se enfrenta al desafío de planear la demanda de la empresa de telecomunicaciones, donde para lograrlo requiere pronósticos precisos para determinar la capacidad necesaria o la expansión de la red. Aunque el análisis de series de tiempo ha sido dominado por métodos estadísticos lineales, las técnicas de redes neuronales (RNAs) son más efectivas y prometedoras.

Marco Contextual

Este proyecto se centra en el análisis de datos y la predicción de series temporales, con un enfoque en la técnica de la red neuronal LSTM (Long Short-Term Memory). Estas redes son reconocidas por su efectividad demostrada en varios proyectos, donde son capaces de capturar patrones a largo plazo y manejar problemas como desvanecimiento del gradiente y la estacionalidad, características esenciales para obtener pronósticos precisos en datos temporales.

Es importante conocer el entorno actual de las empresas, en este caso hablamos de una empresa mediana comercializadora tecnología como sonido profesional, repuestos de electrónica, iluminación y todo tipo de tecnología, está se enfrenta a las constantes fluctuaciones de la demanda de los productos, por lo que tener herramientas que ayuden a optimizar los procesos de distribución y logística, permitirá anticipar con precisión las ventas futuras, reflejándose en mejores tomas de decisiones en la gestión de inventarios, reduciendo costos y maximizando rentabilidad.

Este proceso de la creación de una red neuronal LSTM comienza en la recopilación de datos históricos de ventas diarios, teniendo en cuenta la temporada de aumento de ventas, se preprocesa los datos para garantizar la estabilidad del modelo, comprendiendo que la empresa maneja distintos grupos y sedes, se han diseñado varios modelos, segmentados según las líneas de mayor rotación y un modelo de manera global.

Este enfoque detallado busca prevenir tanto el desabastecimiento como la acumulación de inventario, optimizando costos operativos.

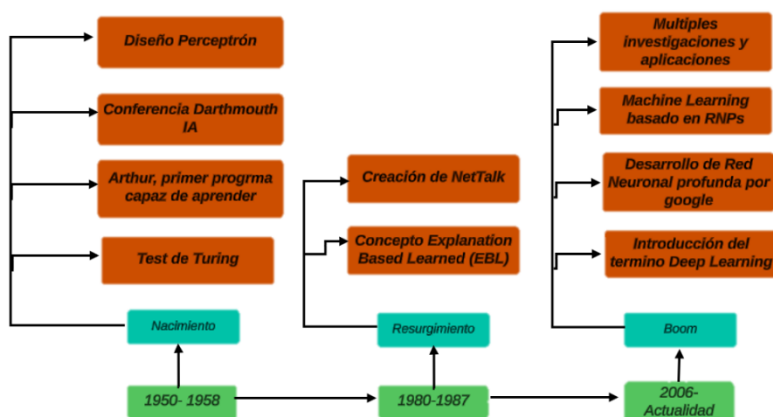
Marco Teórico

Inteligencia Artificial (IA) es una rama de la tecnología que busca simular la inteligencia humana y la capacidad de resolver problemas que incluyen el reconocimiento de patrones, la toma de decisiones y predicciones de eventos futuros basados en datos históricos (Russell & Norvig, 2010). Alan Turing, considerado el padre de la computación moderna, introdujo el concepto a través del “*Test de Turing*”, prueba diseñada para evaluar la capacidad de las máquinas en imitar el comportamiento humano, todas sus investigaciones sentaron las bases para el desarrollo de esta disciplina.

En 1956 John McCarthy acuñó el término Inteligencia Artificial y lo definió como “la ciencia y la ingeniería de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes”, y unos meses después, se diseñó la primera red neuronal por el psicólogo Frank Rosenblatt usando las investigaciones de McCulloch y Pitts (Rodríguez, 2018). Desde entonces ha evolucionado hasta convertirse en una herramienta esencial para resolver problemas complejos en diversos campos, incluyendo análisis de series temporales y predicción de ventas (Goodfellow, Bengio & Courville, 2016).

Figura 1

Línea del Tiempo de Datos Más Sobresalientes de la IA Hasta Nuestros Días



Aprendizaje Automático (Machine Learning): Ciencia de desarrollo de algoritmos y modelos estadísticos que utilizan el sistema de cómputo sin necesidad de dar instrucciones explícitas, hace parte de la IA, lo que significa que todo modelo de ML es una IA, pero no toda la IA es un modelo de ML.

Comprendiendo la historia ha sufrido dos largos periodos de invierno o quietud el primero desde 1974 a 1980, seguido por un boom en los años 80 donde lo más destacable es por Gerald Dejong que introduce el término “Explanation Based Learning” (EBL), donde la computadora analiza los datos de entrenamiento y genera reglas generales que permiten descartar los datos menos importantes. (González et al, 2019), Otro importante avance fue de Terry Sejnowski en 1985 que invento *NetTalk* capaz de pronunciar palabras, esto ayudo a sentar las bases del aprendizaje supervisado, pero no es sino hasta los años 90 donde el ordenador Deep Blue de IBM vence al campeón mundial de ajedrez, demostrando la capacidad de aprendizaje de la maquinas.

Hoy en día el Machine Learning es una disciplina necesaria para todo tipo de negocios, investigaciones, al combinarlo con modelos de redes neuronales su rendimiento es mucho más efectivo. Existen varios tipos de algoritmos, analizaremos los dos más grandes, técnicas supervisadas y no supervisadas.

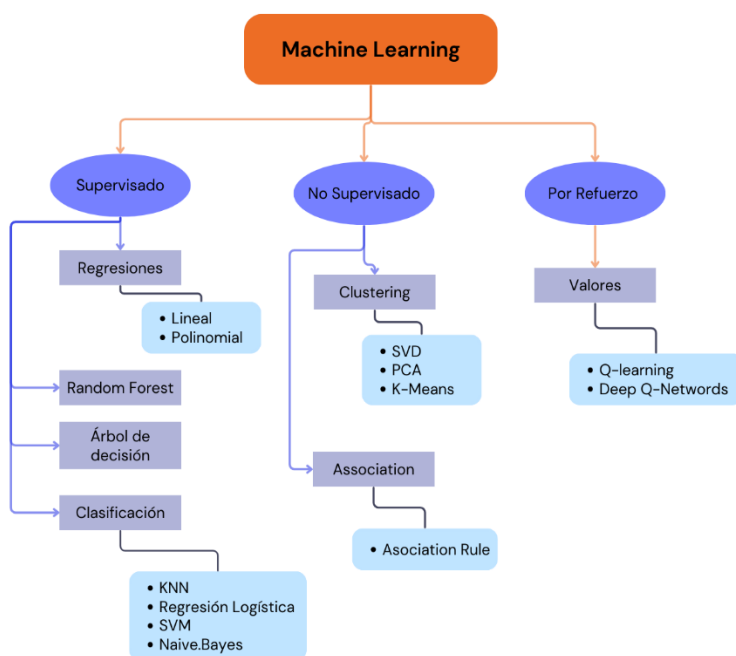
Aprendizaje Supervisado Datos etiquetados y definidos para evaluar las correlaciones, tiene como ventaja la simplicidad y facilidad de diseño, ayuda a predecir un posible conjunto limitado de resultados. Como principal dificultad es etiquetar millones de datos, sin embargo, es muy utilizado para detectar patrones, clasificar y predecir.

Aprendizaje No supervisado La máquina encuentra patrones de datos no etiquetados superando la dificultad del aprendizaje supervisado, entonces la máquina asigna etiquetas o

encuentra patrones, conexiones significativas, es muy utilizado para determinar patrones de compra de clientes, recomendaciones de anuncios, publicidad enfocada entre muchos otros.

Figura 2

Estructura de Técnicas de Machine Learning



Precisión (Accuracy): Porcentaje de ejemplos correctamente estimados del total de predicciones hechas.

Error cuadrático medio (MSE): Promedio de los errores al cuadrado, muy sensibles a valores extremos.

Desviación absoluta del promedio: Media de las desviaciones absolutas.

Sensibilidad: Tasa positiva real, son los valores positivos que se identifican correctamente.

Especificidad: Tasa negativa real, son los valores negativos que se identifican correctamente.

Pronóstico de ventas: Es una proyección de ingresos futuros conociendo datos de períodos anteriores, por lo tanto, se puede predecir la efectividad de distintos sectores a corto y largo plazo. (Douglas et al. 2022). Su principal función es anticipar el comportamiento de los datos, esto ayudara a tomar decisiones informadas, aprovechando cada oportunidad. Aplicando al desarrollo del proyecto para las empresas sus principales ventajas de aplicar pronósticos de ventas son:

- Asignación de recursos de manera eficiente.
- Adecuado nivel de inventarios.
- Planificación de crecimiento de las empresas.

Existen dos grandes tipos de pronósticos el cuantitativo y el cualitativo, en el primero se utiliza datos medibles para formulas hechos y encontrar patrones de comportamientos, se obtienen mediante estudios, sistemas contables, encuestas entre otros. Y en los cualitativos se enfocan en mencionar lo que se destacará sin incluir ingresos o datos medibles, se utiliza a menudo para comprender razones, opiniones o motivaciones. (Montserrat, et al. 2022).

Series de tiempo: Es una secuencia de N observaciones o datos ordenados y usualmente equiespaciados cronológicamente sobre una o varias características de una unidad observable en distintos momentos. (José et al. 2007)

Series univariantes

$$y_1, y_2, \dots, y_N; (y_t)_{t=1}^N; (y_t; t = 1, \dots, N),$$

Donde y_t es la observación

$$n^\circ t (1 \leq t \leq N)$$

En el análisis de series de tiempo se refiere al proceso de analizar los datos y poder descubrir tendencia o patrones en los datos. Además, que permite extraer y modelar las

relaciones entre los datos a lo largo del tiempo, extrapolando o interpolando el comportamiento de datos no observados. (Rojas, et al. 2022). Se comprende que las series de tiempo cuentan con componentes como la tendencia, estacionalidad, ciclicidad y aleatoriedad.

Tendencia: comportamientos repetitivos o patrones subyacentes a lo largo del tiempo.

Estacionalidad: Fluctuaciones en los datos influenciados por periodos fijos como días, semanas, mes.

Ciclicidad: Variación irregular en los datos que tienden a ocurrir en ciclos de 2 años o más.

Aleatoriedad: Partes inexplicables de los datos.

Ruido blanco: Cuando la serie tiene valores que son aleatorios, independientes entre sí, con una media de cero y varianza constante.

Modelo ARIMA

Acrónimo de Autoregressive Integrated Moving Average o Modelo Autorregresivo Integrado de Media Móviles. Desarrollado por Box y Jenkins en 1976, es una técnica básica para el pronóstico de series de tiempo su uso principal es en estadística y econometría, no es usado para modelar series de tiempo con comportamiento cíclico, este modelo utiliza variaciones y regresiones de datos estadísticos para encontrar patrones para realizar predicciones futuras. (Palma et al, 2022).

Utiliza promedios móviles rezagados para suavizar los datos. Mide la fortaleza de una variable dependiente en relación con las variables independientes. Además, que predice basándose en las diferencias entre los valores de la serie en lugar de los valores reales.

Componentes del modelo ARIMA

AR (p): Indica el número de observaciones pasadas de la serie de tiempo, similar a un modelo de regresión donde los retardos (lags) de los datos históricos sirven como predictores. Es decir, prueba los datos para determinar su estacionariedad.

$$y_t = C + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

Donde,

y_t : Son los datos que se aplican al modelo ARMA, donde ya tiene una transformación de potencia y la serie está diferenciada.

ϕ_1 : Coeficientes de AR.

I(d): Indica el número de las diferencias para lograr que la serie sea estacionaria. Esto se logra al aplicar la diferenciación, restando valores consecutivos para eliminar tendencias y estacionalidad, todo para lograr que la media y la varianza sean constantes en el tiempo, haciendo modelos más precisos y evitando el ruido.

MA(q): Analiza la relación entre las observaciones actuales y los errores residuales de observaciones pasadas. Nos ayuda a corregir predicciones y mejorar la precisión del modelo al detectar patrones no evidentes a la vista.

$$y_t = C + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Donde,

θ_1 : Son los coeficientes de MA.

Para llevar a cabo modelos de ARIMA es recomendable seguir el Método de Box-Jenkins, el cual nos dice:

Identificar y seleccionar el modelo, se determina si la serie temporal es estacionaria, si no lo es se debe calcular cuántas diferenciaciones son necesarias para lograrlo. Aquí se analiza los

gráficos de autocorrelación (ACF) y autocorrelación parcial (PACF) para definir los parámetros del modelo.

Estimación de parámetros autorregresivos (AR), se ajusta el modelo utilizando los datos, calculando los coeficientes de los parámetros.

Se comprueba el modelo, evaluando con datos disponibles para detectar sobreajustes, analizar residuales y detectar mejoras.

Se realiza el pronóstico utilizando el modelo ajustado para predecir los valores futuros. (Hayes et al, 2024), (Noble et al.2024)

Modelo SARIMA: Autoregressive Integrated Moving Average with Seasonalite. Este modelo generaliza todos los modelos de la familia ARIMA, nos permite modelar series estacionarias y no estacionarias en las series temporales, proporcionando una herramienta flexible para analizar y predecir datos que contengan patrones repetitivos. (Amézquita et al., 2021).

(p, d, q): Representan los parámetros el modelo ARIMA

(P, D, Q, s): Representan los parámetros estacionales:

P: Orden autorregresivo estacional. (SAR)

D: Diferenciación estacional.

Q: Orden de promedios móviles estacional. (SMA)

s: Frecuencia de la estacionalidad.

$$y_t = C + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^p \Phi_i y_{t-is} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^q \Theta_i \varepsilon_{t-is}$$

Donde:

y_t : Valor de la serie en el tiempo

C: Constante

ϕ_i : Coeficientes del componente autorregresivo.

Φ_i : Coeficientes del componente autorregresivo estacional.

ε_t : Término de error en el tiempo

θ_i : Coeficientes del componente de promedios móviles.

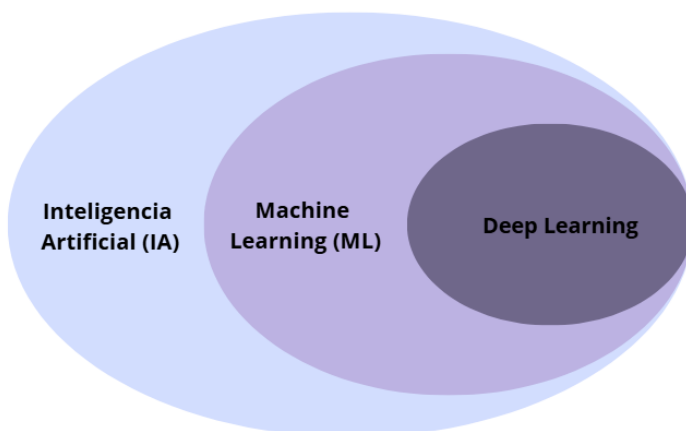
Θ_i : Coeficientes de componente de promedios móviles estacionales.

Modelo muy utilizado en contextos donde la serie temporal contiene estacionalidad, como ventas minoristas y mayoristas y contiene temporadas de navidad, de festividades o promociones, También en consumo energético que depende del invierno o el verano y muchas aplicaciones más.

Deep Learning: Aprendizaje profundo es un subconjunto de Machine Learning que utiliza algoritmos para modelar abstracciones de alto nivel mediante arquitecturas compuestas de transformaciones no lineales. Contiene principios específicos y arquitecturas especializadas para tareas más profundas de lo que abarca el Machine Learning tradicional. Algunas ventajas del aprendizaje profundo es que son modelos más robustos, son muy adaptables, procesan información en paralelo, capaces de trabajar en sistemas no lineales, tienen un alto nivel de tolerancia a fallas. Sin embargo, contienen limitaciones como no lograr identificar claramente el mecanismo que ha resuelto el problema, son modelos más complejos de entender. (Rojas, et al. 2022)

Figura 3

Disciplinas de la Inteligencia Artificial



Red Neuronal Biológica: Con los descubrimientos de *Santiago Ramón y Cajal* en el siglo XIX, sabemos que las neuronas son la base celular del cerebro y dijo “*el sistema nervioso está formado por células independientes: las neuronas, que contactan entre sí en lugares específicos.*” Entonces comprendemos que la neurona es el componente principal del sistema nervioso y su función es recibir, procesar y transmitir información a través de señales químicas y eléctricas.

Se estima que el sistema nervioso tiene cerca de cien mil millones de neuronas organizadas en una red compleja, y cada una puede estar conectada a miles de neuronas distintas.

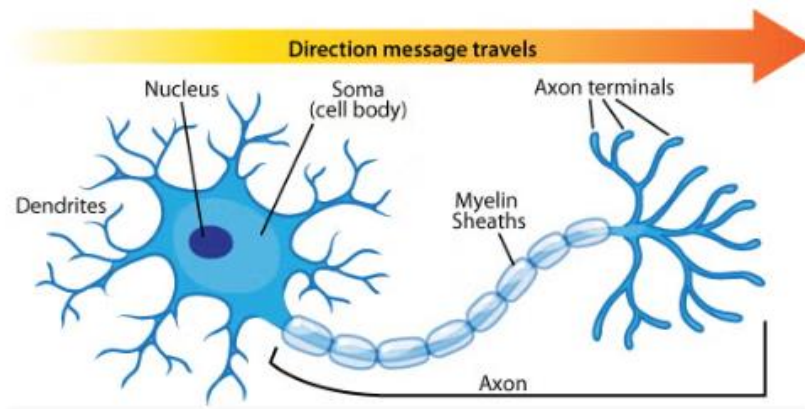
Donal O. Hebb contribuye a explicar el proceso psicológico o el aprendizaje y establece una conexión entre las funciones biológicas de la neurona y lo que conocemos como mente.

Su estructura dendritas canal de entrada de la información, el soma órgano de cómputo y el axón canal de salida, la mayoría de las neuronas contiene una estructura de árbol llamada dendritas y se unen a otras neuronas a través de uniones llamadas sinapsis. El tipo de sinapsis más común no existe contacto físico entre ellas y permanecen separadas por 0.2 micras. Y en

ellas existen las *presinápticas*, es decir, envían las señales, y las *postsinápticas*, las que reciben las señales, concluyendo que la información fluye en un único sentido. (Grupo US).

Figura 4

Anatomía de Neurona Biológica

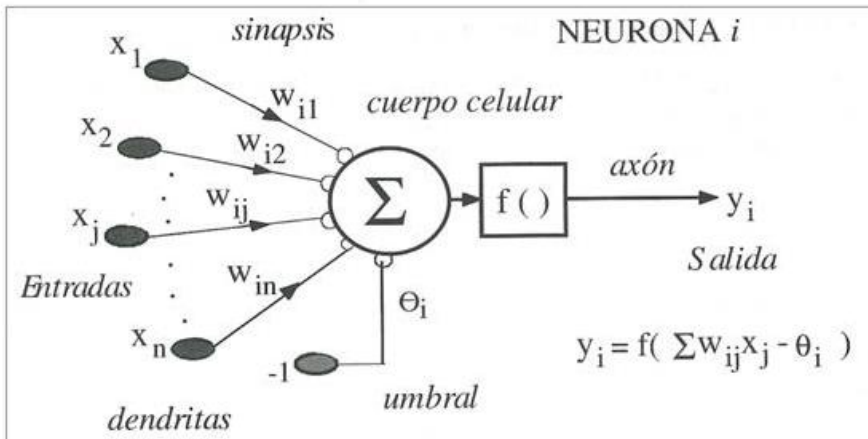


Nota. Tomado De ASU School Os Life Sciences.

Perceptrón: La neurona como parte individual de una red neuronal es un procesador fundamental que a partir de un elemento de entrada procedente de datos iniciales u otras neuronas proporcionara una única salida. (Grupo US). Se puede decir que el perceptrón es una función matemática.

Figura 5

Modelo de Neurona Estándar (Perceptrón)



Nota. Tomado de Grupo.us.es

Compuesta por:

Entrada: vector de entradas (x_i) pueden ser binarias o continuas.

Pesos sinápticos: valores que representan la intensidad de interacción entre cada neurona presináptica y postsináptica (W_{ij})

Regla de propagación: Se basa en una suma ponderada de las entradas con los pesos sinápticos, representados:

$$h_i(t) = \sum_j W_{ij}x_j = W_i^T x$$

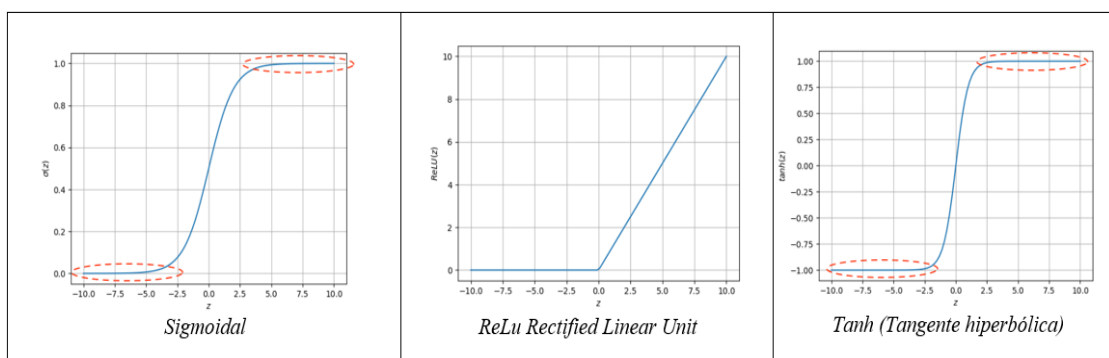
Función de activación: Proporciona el estado de activación en función del estado anterior y del calor postsináptico.

Sigmoide: Toma cualquier rango de valores a la entrada y los mapea al rango de 0 a 1 en la salida. Usada principalmente para clasificación binaria.

ReLU: Unidad lineal rectificada, rectifica los datos negativos y los convierte a cero a la salida y los valores positivos no sufren modificaciones, es el más recomendable dado que no tiene problemas de saturación, es fácil de implementar, eficiente y ayuda a que el gradiente descendente converja más rápido.

Figura 6

Funciones de Activación



Nota. Tomado de Codificandobits

Tanh: Tangente hiperbólica, similar al sigmoide pero los valores de salida oscilan de -1 a 1.

Función de salida: Emite la salida en función del estado de activación.

Red Neuronal Artificial: Inspirado en las redes biológicas, contiene un conjunto de unidades conectadas entre sí para transmitir señales. Funcionan al combinar ejemplos de datos de entrenamiento e identificando de manera automática las correlaciones entre las variables suministradas. Cuando ya está entrenado, el algoritmo puede usar las asociaciones realizadas para interpretar nuevos datos. (Rojas et al. 2022)

Como concepto de red neuronal es necesario hacer uso del concepto matemático de *grafo*, dirigido con las siguientes propiedades:

A cada nodo i se le asocia una variable de estado x_i

A cada conexión (i, j) de los nodos i y j se le asocia un peso $W_{ij} \in R$

A cada nodo i se le asocia un umbral θ_i

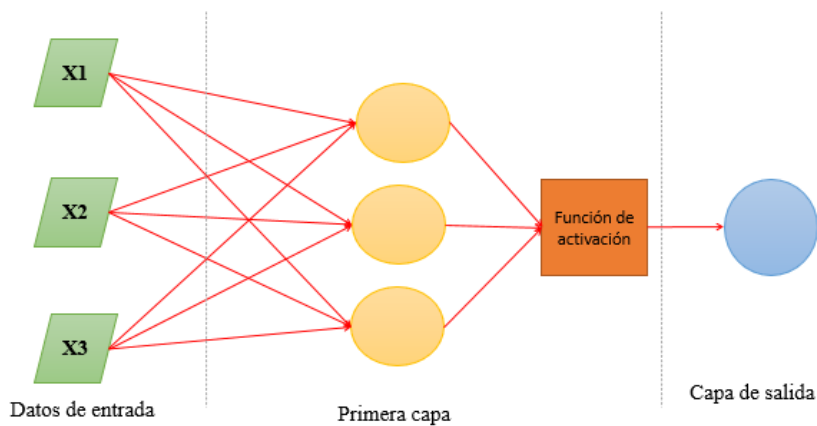
Para cada nodo i se define una función $f_i(x_i, W_{ij}, \theta_i)$, que depende de los pesos de sus conexiones, del umbral y de los estados de los nodos j a él conectados. Y esta función proporciona el nuevo estado del nodo. (Grupo US)

Red Neuronal Monocapa o perceptrón simple

Compuesto por una capa de neuronas que envían las entradas a una capa de neuronas de salida donde realiza los cálculos. (Calvo Diego et al. 2017)

Figura 7

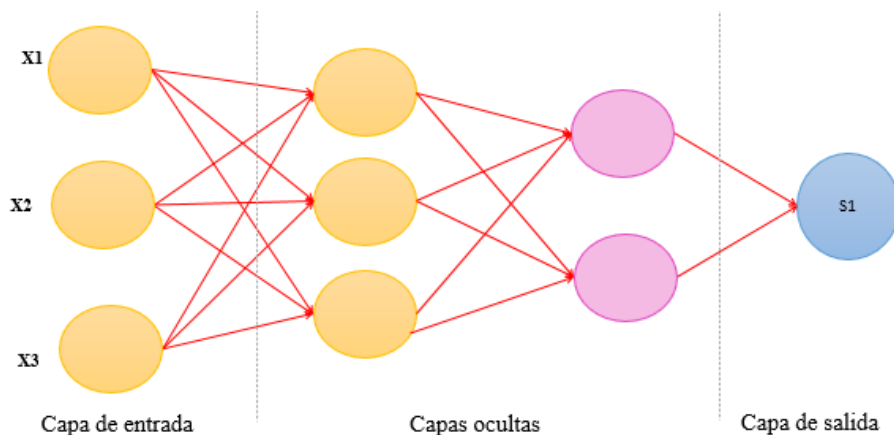
Estructura Red Neuronal Monocapa



Red Neuronal Multicapa: Contiene un conjunto de capas intermedias conocidas como capas ocultas, se ubican entre la capa de entrada y la de salida, pueden estar parcial o totalmente conectadas. (Calvo Diego et al. 2017).

Figura 8

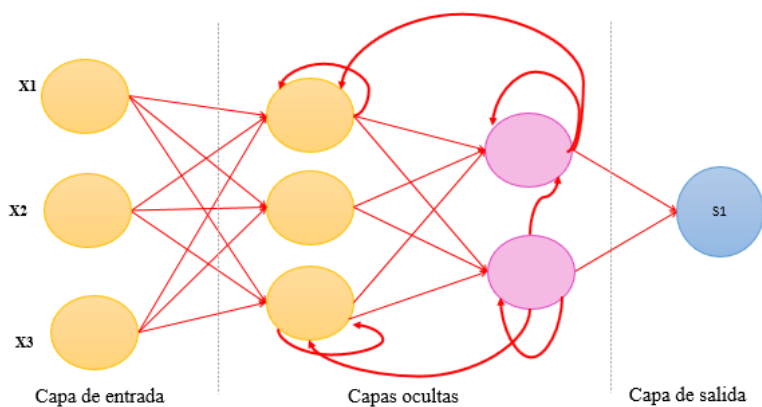
Estructura Red Neuronal Multicapa



Red Neuronal Recurrente: Almacenan de manera más eficiente la información, permite conexiones arbitrarias entre redes, con lo cual consigue crear la temporalidad y la red consiga memoria, contiene una memoria a corto plazo (Calvo Diego et al. 2017).

Figura 9

Estructura de RNN



Red Neuronal LSTM: Long Short Term Memory o Memoria corta a largo plazo, hace parte de las redes neuronales recurrentes (RNN), diseñada especialmente para ayudar con el problema de la desaparición del gradiente, que se produce cuando se retro propaga el error a través de las múltiples capas y la información se pierde.

Este tipo de redes utilizan una estructura de celdas con puertas que permite a la red controlar la cantidad de información que se almacén y se olvida en cada paso de tiempo.

Celda de estado: Elemento diferenciador de las redes RNN, donde se decide los datos que se pueden añadir o remover de la memoria de la red.

- Compuerta Forget: permite eliminar elementos de la memoria.
- Compuerta update: permite añadir nuevos datos a la memoria.
- Compuerta de salida: Crea un estado oculto actualizado.

Estas compuertas se encargan de detectar y clasificar la información necesaria y útil para la red.

Su uso es cada vez mayor, generalmente es usada en generación de texto, imágenes, análisis de sentimientos, predicciones financieras, composición musical.

Marco Conceptual

Épocas: Ciclo de corrección que se realiza para reducir la pérdida. Estas realizan una propagación hacia atrás y hacia adelante para corregir los errores.

Híper parámetros: Son variables de configuración externa que los científicos de datos utilizan para administrar el entrenamiento de modelo de machine Learning. Logran determinar características clave como la arquitectura del modelo, la tasa de aprendizaje y la complejidad del modelo. (Amazon web service)

Lotes: Número de muestras de entrenamiento que se envían a la red neuronal a la vez.

Optimizador Adam: Adaptativa de momentos es Algoritmo de optimización utilizado para el aprendizaje automático y profundo, mejora el rendimiento del entrenamiento de las redes neuronales. Principalmente ajusta automáticamente las tasas de aprendizaje en función de cada parámetro. Se adapta dinámicamente a los ritmos de aprendizaje, permitiendo obtener buenos resultados.

Parámetros: Elementos internos derivados de forma automática durante el proceso de aprendizaje.

Suavizado Kalman: Algoritmo basado en el filtro de Kalman, que combina modelos matemáticos con datos reales para producir estimaciones de estado óptimas. Procesa las mediciones en un paso hacia adelante seguido de un paso hacia atrás.

Error cuadrático medio (MSE): Mide la magnitud cuadrática promedio de los errores.

$$MSE = \frac{\sum_{i=1}^n e_i^2}{n}$$

Desviación típica de los errores (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n - 1}}$$

Estas dos métricas MSE y RMSE son usados para medir la dispersión de los errores, es decir que tanto se acercan o se alejan los valores del pronóstico de los valores reales. Si este valor es pequeño significa que el pronóstico es muy cercano a los valores reales y si es muy grande existen errores en el pronóstico.

Error absoluto medio (MAE)

Mide la magnitud promedio de los errores entre valores predichos y valores reales. Calculando la suma de las diferencias absolutas entre las predicciones y las observaciones, dividida por el número de muestras.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Error porcentual absoluto medio (MAPE): Nos ayuda a medir el rendimiento del pronóstico.

$$MAPE = \frac{\sum_{i=1}^n |PE_i|}{n}$$

Donde,

n= número de observaciones

Metodología

La metodología que se propone utilizar está basada en CRISP-DM (Cross Industry Standard Process for Data Mining) se encuentra dividida en seis etapas fundamentales el profesor Helder Prado nos dice “Las tres primeras etapas tienen como objetivo la recolección y organización de los datos para ser analizados.”

Entendimiento del negocio: Comprender el objetivo del proyecto y las necesidades del negocio o empresa.

Entendimiento de los datos: Recolección inicial de datos, familiarizarse con los datos, descubrir aspectos interesantes y resaltables.

Preparación de datos: Seleccionar, organizar y limpiar los datos, escoger variables y tipos de variables a utilizar. En este punto se debe revisar el tratamiento que se debe dar a los datos, si es necesario escalarlos, qué hacer con los valores nulos, reducir el ruido, suavizar los datos, entre otros.

Las siguientes tres etapas tienen como objetivo la creación del modelo y ejecución de este, es donde el modelo será evaluado e implementado. (Abid Alli, et al. 2024)

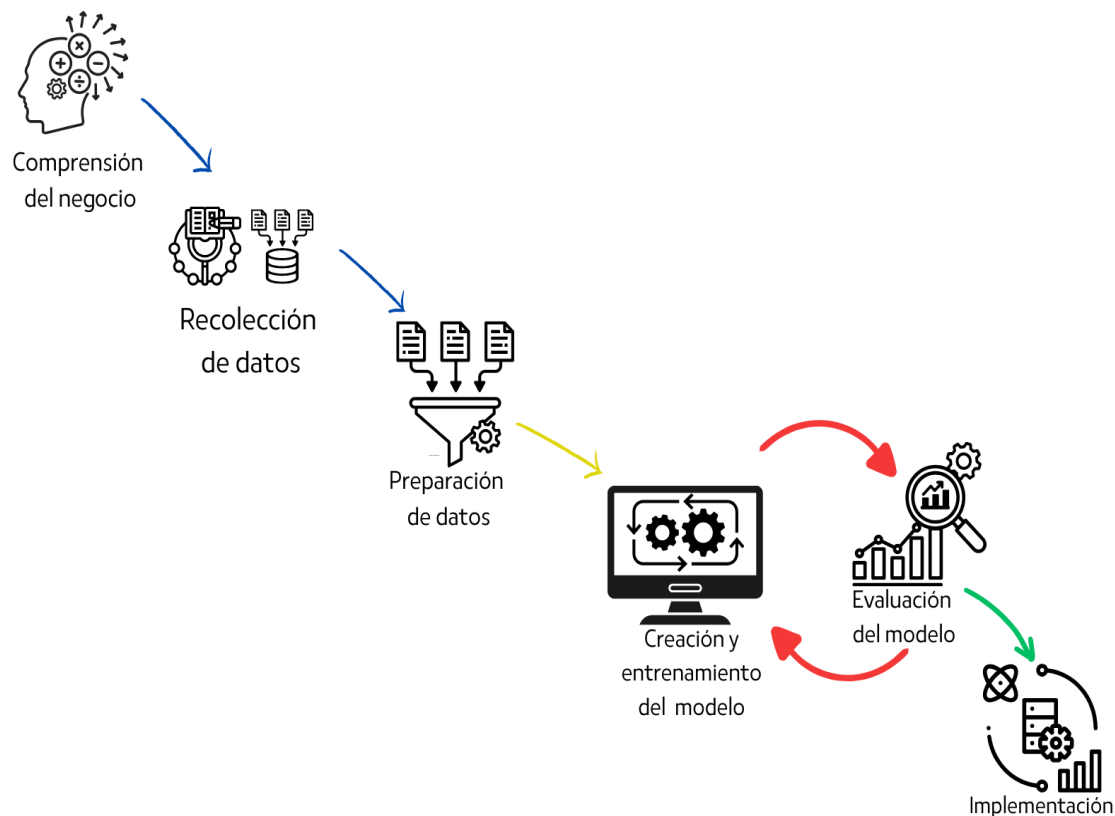
Modelado: Seleccionar y aplicar los modelos definidos. Se debe construir el modelo, definir las métricas a aplicar, entrenar y validar el conjunto de datos de entrenamiento y validación. Reajustar el modelo hasta conseguir los resultados esperados.

Evaluación: Con las métricas escogidas se revisa si los resultados corresponden a lo esperado, se toma decisiones sobre las posibles mejoras y si es necesario se corre nuevamente el modelo.

Implementación: Modelo debe ser llevado a producción, generalmente son almacenados en la nube o servidores locales, donde deben estar en constante supervisión y mantenimiento garantizando tener un buen rendimiento. (Caio et al. 2022)

Figura 10

Proceso de Creación de Modelo de Metodología CRISP-DM



Inicio del Proyecto

Siguiendo las bases de la metodología CRISP-DM en la construcción de un modelo basado en redes neuronales LSTM para la predicción de ventas de la empresa de comercio electrónico. Plasmaremos el proceso.

Lenguaje de programación: Se escoge como lenguaje a utilizar Python dado que ocupa la posición número uno en el índice TIOBE, además resalta por su versatilidad, legibilidad y

sencillez, considerado un lenguaje intuitivo y abarca muchas áreas de las tecnologías desde desarrollo web hasta Inteligencia Artificial y es el más común en la ciencia de datos y en aplicación de modelos de machine Learning. (Bilib et, al. 2024)

Figura 11

Top 5 de Ranking de Lenguajes de Programación Más Utilizados

Jan 2025	Jan 2024	Change	Programming Language	Ratings	Change
1	1		 Python	23.28%	+9.32%
2	3	▲	 C++	10.29%	+0.33%
3	4	▲	 Java	10.15%	+2.28%
4	2	▼	 C	8.86%	-2.59%
5	5		 C#	4.45%	-2.71%

Nota. Tomado de Tiobe.index

Librerías utilizadas:

Pandas: Su principal uso es para cargar, manipular, alinear los datos, engloba la transformación de datos.

1. Matplotlib: Especializada en creación de visualizaciones gráficos de 2D.
2. Numpy: Utilizada para cálculo numérico y análisis de datos
3. Sklearn: Utilizado para realizar algoritmos eficientes, incluye procesamiento de datos y evaluación de los modelos.
4. Keras: API de alto nivel que permite crear y entrenar modelos de machine Learning y Deep Learning, en este caso las redes neuronales LSTM.

Recolección de Datos

Los datos fueron recopilados directamente desde la base contable de la empresa, abarcando el período comprendido entre 2021 y 2024. Estos datos están organizados por grupo de venta, el cual corresponde a las diversas líneas de productos que maneja la empresa.

Preparación de Datos

El proceso de preparación de datos comienza con la modificación de los nombres de las columnas en los conjuntos de datos para asegurar su consistencia y facilitar su manejo posterior. Dado que los datos provienen están en distintos archivos se procede a unirlos en único conjunto de datos, que permita trabajar de manera más eficiente.

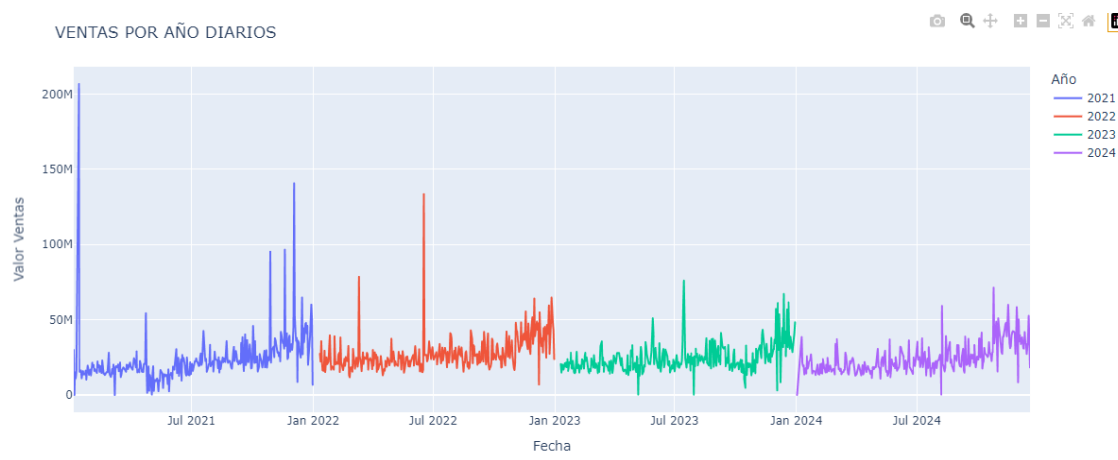
Se realizan conversiones de tipos de datos en varias características, por ejemplo, la columna fecha se convierte al tipo data time para facilitar la manipulación en el análisis temporal. Una vez realizados estos ajustes iniciales, se procede a examinar el contenido de los datos, se observa que el conjunto de datos tiene facturas de ventas y devoluciones, pero este análisis se centrará únicamente en las facturas de ventas por lo tanto se filtran las devoluciones y se las excluye del análisis.

También se revisa que todas las facturas de venta tengan un valor superior a cero, eliminado así cualquier error o anomalía que pueda existir en los datos.

Además, se detecta que en los datos existen valores extremos como se observa en la figura 12, se conoce que estos datos corresponden a los días sin IVA promocionado por años anteriores por el gobierno, por lo tanto, se procede a ponderar con la media estos valores.

Figura 12

Ventas Diarias Antes de Ponderar los Valores Extremos



En este análisis exploratorio identifica el aporte de cada vendedor al monto total ventas de la empresa del año 2024, y se lo refleja en porcentaje. Esta información resulta de gran utilidad para la gerencia de la empresa, ya que permite identificar a los mejores colaboradores en términos de rendimiento y ventas.

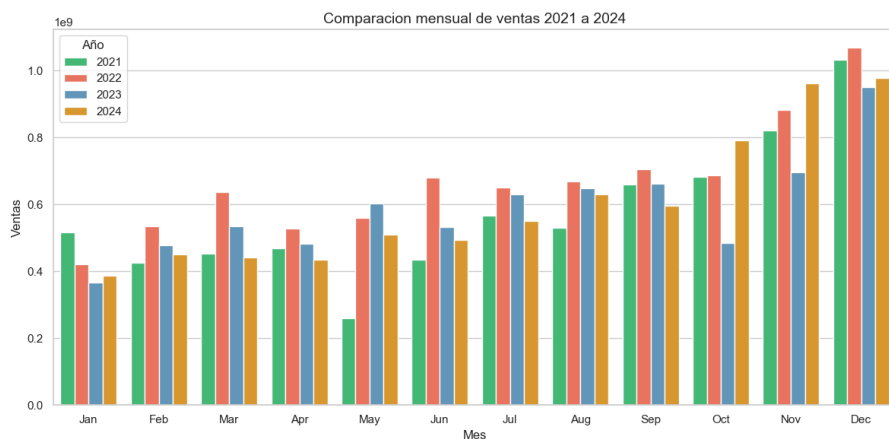
En esta gráfica se observa como el vendedor con código 1 aporta un 13.3% al total del monto de ventas a la empresa y luego el vendedor con código 26 aporta un 7.6%, se evidencia una diferencia significativa entre ambos. Con esto se puede tomar decisiones informadas sobre incentivos, estrategias de ventas y programas de capacitación, permitiendo reducir la brecha de desempeño y optimizar el rendimiento de todo el equipo de ventas.

Figura 13*Aporte de Ventas por Vendedor(%)*

Se presenta el movimiento por mes de la empresa en el periodo de 2021 a 2024, con el monto de ventas generado. Evidenciando un aumento de ventas a partir del mes de octubre a diciembre de todos los años, con excepción del 2023, donde se nos informa que el departamento de Nariño sufrió un paro por parte de los camioneros, perjudicando notoriamente las ventas de esta época de la empresa.

Figura 14

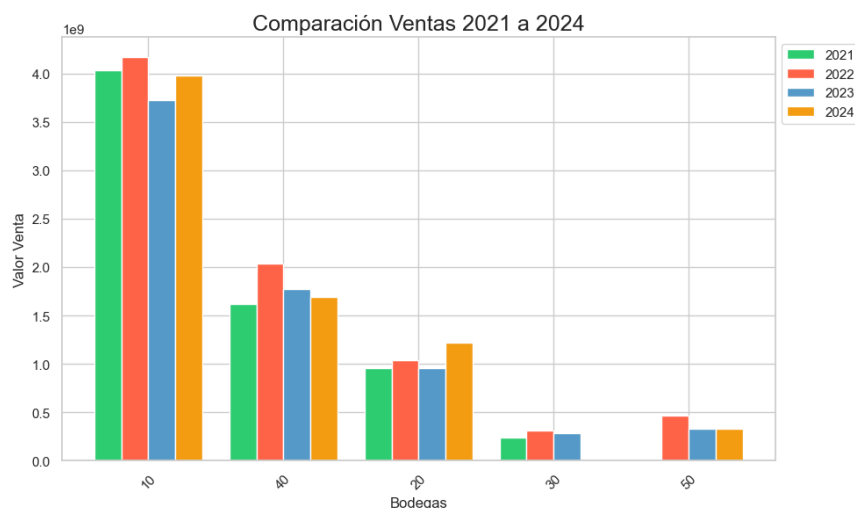
Comparación Mensual de Monto de Ventas Desde 2021 a 2024



Se realiza una comparación de ventas desde 2021 a 2024 por sedes de la empresa. Nos centraremos en el crecimiento o decrecimiento del año 2024 por sedes, se compara el año 2024 con relación al 2023. Se observa que la sede principal 10 tuvo un crecimiento del 6.93%, la sede 40 tuvo un decrecimiento del 4.54%, la sede 20 tuvo un importante crecimiento del 27.9% y la sede 50 tuvo un pequeño decrecimiento del 1.7%

Figura 15

Comparación Ventas Anual por Sede



Se encuentra una disminución de afluencia de clientes, en tabla se compara 2024 con 2023 tenemos una disminución de 5.774 facturas efectuadas que corresponde a un 4.1%, y comparando 2023 con 2022 se tuvo una disminución del 6% correspondiente a 9.041 facturas emitidas. El único aumento registrado en los cuatro años fue de 2022 con 6.182 más facturas. Por lo tanto, desde el 2022 se evidencia una disminución del 9.89% correspondiente a 14.815 facturas emitidas. Por lo tanto, se recomienda revisar por qué la pérdida de estos posibles clientes puede deberse a factores como: atención al cliente, falta de atención, falta del producto, crisis económica, deficiencia en facturación, calidad del producto, migración de clientes a canales digitales, entre otras.

Tabla 1

Comparación de Facturas Emitidas Desde 2021 a 2024

Año	Número de facturas	Porcentaje%	Diferencia
2021	143.566		
2022	149.750	4.3	6.182
2023	140.709	-6.03	9.041
2024	134.935	-4.1	5.774

Nota. Comparación de facturas emitidas desde 2021 a 2024

En este análisis exploratorio se identifica un campo llamado grupo el cual contiene una línea específica de productos, por ejemplo el grupo 48 están ingresados toda la línea de micrófonos y estos están segmentados por códigos únicos de productos, Se considera importante conocer el monto que aporta cada grupo a las ventas totales de la empresa, por eso se obtiene a los cinco primeros grupos con mayores montos de ventas del año 2024.

Tabla 2*Top 5 de Grupos de Productos del Año 2024*

ítem	Grupo	Cantidad	Valor
1	64	1788.00	\$1.178.865.000
2	15	100172.75	\$875.570.400
3	12	193104.91	\$695.624.900
4	65	2108.00	\$477.769.100
5	48	4529.00	\$412.051.900

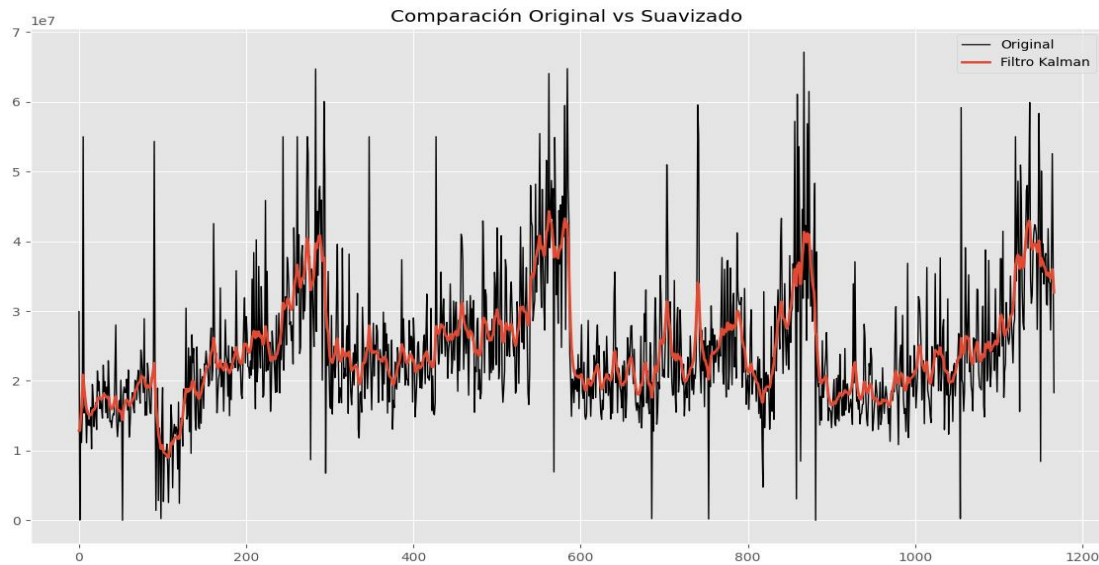
Nota. Grupos con mayores ventas registradas para la empresa.

Con esta información se segmenta la base de datos, una base con todos los datos generales, otra con los datos del grupo número uno el 64, correspondiente a ‘sonido profesional’ y otra con los datos de los grupos número 12 y 15 correspondiente a ‘iluminación’.

Se realiza un suavizado de los datos a cada segmentación con el método Kalman, el más aplicado en campo como las finanzas y el comercio. Nos ayuda a reducir el ruido y hacer que las series de tiempo se han más claras y fáciles de analizar. (Chainika Thankar, 2024)

Figura 16

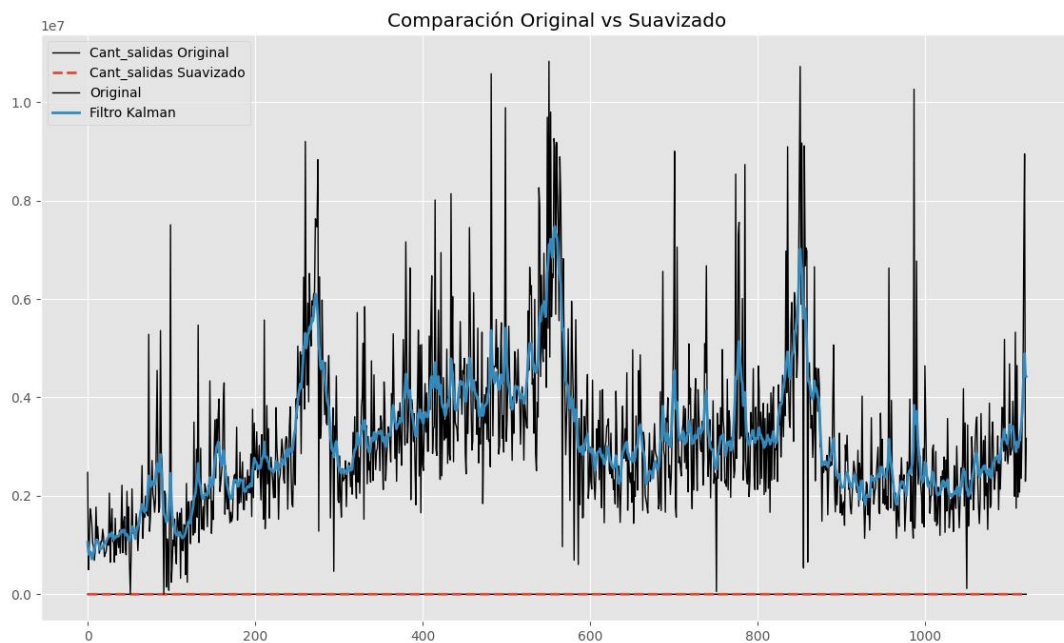
Comparación de Datos Original y Suavizado con Kalman Segmentación Uno



En la imagen podemos ver como a pesar de haber realizado el debido proceso con los valores extremos, siguen apareciendo valores superiores a la media y se opta por el suavizado en todas las tres redes neuronales para lograr tener un control más óptimo de los datos.

Figura 17

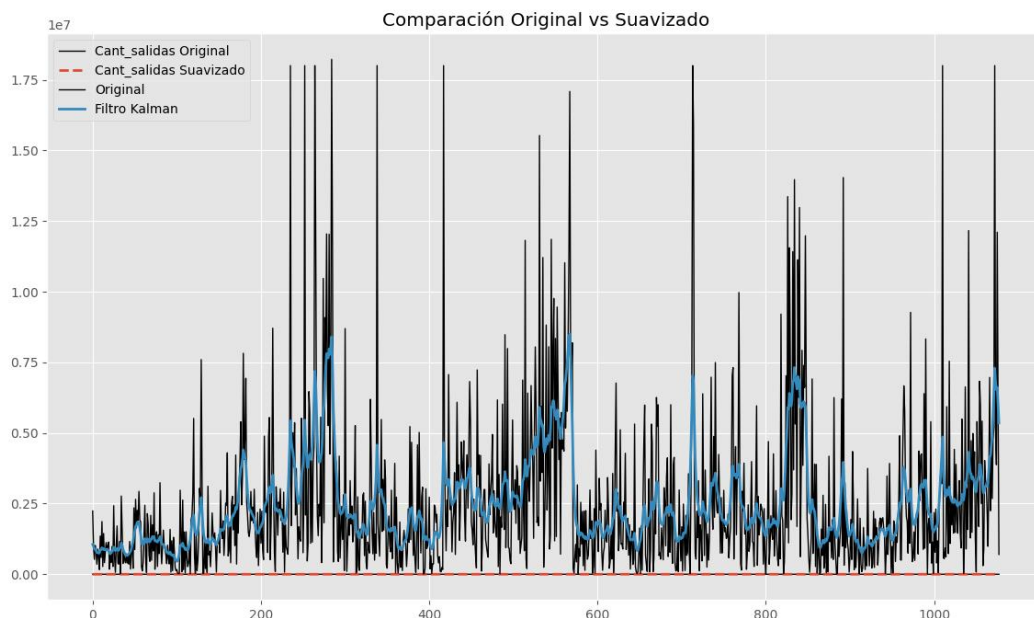
Comparación de Datos Original y Suavizado con Kalman Segmentación Dos



En la siguiente imagen se ve la diferencia entre los valores sin realizar la técnica del suavizado de Kalman y en línea azul los valores con el escalado pertenecientes a los datos del grupo de iluminación.

Figura 18

Comparación de Datos Original y Suavizado con Kalman Segmentación Tres



Con el objetivo de aprovechar los grupos de productos que generan mayores ventas se ha decidió desarrollar tres redes neuronales específicas, cada una diseñada para abordar distintos aspectos del análisis de ventas.

Primer Red: Contendrá el modelo global, estará orientada a predecir las ventas a nivel general, utilizando como principales características el valor de la venta y la fecha. Para la fecha se desglosa en varias variables temporales clave: el día de la semana, el mes y la temporada, esta última representa el último trimestre de cada año.

En este punto se divide los datos en train, validación y test, con un 80%, 10% y 10%, respectivamente. Los datos de entrenamiento van desde enero 2021 a marzo de 2024 con un total de 926 registros, la validación desde abril de 2024 a agosto de 2024 con un total de registro se 111. Y el test desde agosto de 2024 a diciembre de 2024 con un total de 109 registros.

Segunda Red: Contendrá a la segmentación dos que contiene al grupo 64 que representa al sonido profesional y pertenece en su mayoría a la sede principal. En este caso, además de las características mencionadas en el modelo global, se añade una característica importante la cantidad vendida, esto permitirá que el modelo aprenda no solo sobre el valor de la venta, sino también sobre el volumen de ventas. Aquí se divide los datos en train, validación y test, con un 80%, 10% y 10%, respectivamente. Los datos de entrenamiento van desde enero 2021 a febrero de 2024 con un total de 890 registros, la validación desde febrero de 2024 a junio de 2024 con un total de registro se 106. Y el test desde junio de 2024 a noviembre de 2024 con un total de 105 registros

Tercera red: Tendrá a la segmentación tres que contiene los códigos 12 y 15, los cuales representan productos de iluminación, como paneles led, bombillos entre otros. Al igual que la segunda red neuronal, las características utilizadas serán cantidad vendida, junto con las variables temporales de día de la semana, mes y temporada. Aquí se divide los datos en train, validación y test, con un 80%, 10% y 10%, respectivamente. Los datos de entrenamiento van desde enero 2021 a febrero de 2024 con un total de 855 registros, la validación desde febrero de 2024 a junio de 2024 con un total de registro se 102. Y el test desde junio de 2024 a noviembre de 2024 con un total de 100 registros

Para las tres redes neuronales se aplica el mismo proceso de datos, en primera medida se escalan los datos para esto usamos el método MinMaxScaler para tener los datos en un mismo rango de 0 y 1, esto es importante para que las redes neuronales puedan aprender de manera más eficiente, evitando que valores muy altos o muy bajos afecten su entrenamiento y ayudando a estabilizar el entrenamiento.

Dado que la red neuronal recibe una estructura de array se realiza el proceso para enviar los datos de la manera correcta.

Modelamiento De Datos

En la construcción del modelo se hizo uso de la librería Keras, una herramienta muy popular para la implementación de redes neuronales en Python. La arquitectura del modelo se basa en una red neuronal secuencial Sequential que nos permite añadir capas de manera progresiva, facilitando el diseño y ajustando los componentes de la red.

Se inicia el proceso con la definición de los datos de entrada, para esto se utiliza un tensor que ayuda a enviar la información desde la capa de entrada hacia las capas ocultas. Estos datos se organizan en forma de vectores y se selecciona una ventana temporal de 7 pasos (Windows size), este valor se determina tras ensayar iterativamente otros tamaños y se elige porque permite capturar patrones de corto plazo, en este caso semanales, además, se utiliza un tamaño de lote (batch size) de 1 ayudando a que el modelo se entrene más rápidamente y permite que la red se adapte los cambios en los patrones y fluctaciones de los datos, también nos permitió optimizar los recursos computacionales disponibles. Además se incorpora un conjunto de características que definen el comportamiento de los datos. En la arquitectura de los datos LSTM tenemos:

Primera capa LSTM: Se configura el tamaño de estado oculto en 32 unidades permitiendo captar patrones más complejos y abstractos. Se utiliza la función de activación ReLU (Rectified Linear Unit) que ayuda en el aprendizaje de datos no lineales y facilitando el aprendizaje. Se añade el parámetro `return_sequences= True` lo que nos asegura que la capa devuelva la secuencia de salida completa en lugar de solo un paso de tiempo. Dado que tenemos la segunda capa y es necesario que se aproveche toda la secuencia.

Segunda capa LSTM : Añadimos 12 unidades, permitiendo una mayor reducción de dimensionalidad y extracción de características más relevantes. Se coloca la función de activación ReLu

Capa de Salida (Dense): Finalmente se añade una capa completamente conecta con una única salida. Y es la encargada de realizar la predicción final.

Después de haber definido la arquitectura del modelo, pasamos a realizar la compilación del modelo para esto se usó:

Optimizador Adam: Adam ajusta de manera dinámica la tasa de aprendizaje, acelerando la convergencia y mejorando el rendimiento, es común usarlo en redes neuronales debido a ser robusto y de fácil implementación.

Función de pérdida MSE (Error cuadrático medio): mide el error promedio cuadrático entre las predicciones del modelo y los valores reales.

Métrica adicional MAE (Error absoluto medio): Se la añade como un cálculo complementario durante el entrenamiento, este mide el error promedio en términos absolutos.

Figura 19

Construcción del Modelo LSTM Ejecutado del Proyecto

```
model_1= Sequential()  
# Capa LSTM  
model_1.add(InputLayer(batch_input_shape=(batch_sizes>window_size, num_features)))  
model_1.add(LSTM(32, activation='relu', stateful= True, return_sequences=True ))  
model_1.add(LSTM(12, activation='relu', stateful= True))  
model_1.add(Dense(1))  
model_1.compile( loss= 'mse', optimizer= 'adam', metrics=['mae'])
```

Entrenamiento

Se realiza el entrenamiento del modelo, para lo cual se utiliza el método `model.fit`, función utilizada para el desarrollo de redes neuronales y otros modelos predictivos. Para este proceso se emplea el conjunto de datos de entrenamiento y sus etiquetas, expuestos en la preparación de datos, esto con el fin de que el modelo aprenda a mapear los patrones de entrada con los valores de salidas.

En el entrenamiento se busca que se realice un aprendizaje completo y robusto, por eso se configura al modelo para ejecutar 150 épocas, lo que significa que el modelo pasara 150 veces ajustando los pesos de la red. También se establece el batch size que determina cuantos ejemplos procesará en cada paso antes de actualizar los parámetros del modelo.

Se incluye el parámetro `verbose=1`, para que nos permite tener una visualización más detallada del entrenamiento, con esta configuración nos muestra la barra de progreso de cada época y nos muestra la pérdida y las métricas antes configuradas.

Se añade el parámetro `shuffle= True`, el cual hace que los datos de entrenamiento se reordenen de manera aleatoria al inicio de cada época. Funciona en casos donde existe patrones repetitivos, obteniendo una reducción de riesgo de que el modelo se ajuste demasiado y genere una mejor generalización a nuevos datos aumentando la capacidad predictiva del modelo.

Se incluye `validation_split = 0.2` para evaluar el modelo en problemas de sobreajuste, con lo cual hacemos que el modelo sea evaluado con datos que no ha visto durante el entrenamiento. Si la pérdida en los datos de validación comienza a aumentar mientras la pérdida en los datos disminuye será una clara señal de sobreajuste.

Estos datos pertenecen a la red neuronal numero uno en la siguiente imagen se observa el entrenamiento que se le realizo en el lenguaje de programación Python.

Figura 20*Entrenamiento de Red Neuronal Global*

```
red_1= model_1.fit(train_X, train_y, epochs=150, batch_size=batch_sizes, verbose=1, shuffle=True, validation_split=0.2)
```

En cada red neuronal se trabajó con épocas distintas, para conocerlas se realizó el proceso de manera iterativa y se escogió las que mejores métricas arrojaban, para la red neuronal dos perteneciente a los productos de sonido profesional, en donde las mejores métricas las obtuvimos al llegar a las 200 épocas.

Figura 21*Entrenamiento de Red Neuronal Dos*

```
red_2= model_1.fit(train_X, train_y, epochs=200, batch_size=batch_sizes, verbose=1, shuffle=True, validation_split=0.2)
```

Para la red neuronal número tres fue la red con menos épocas usadas un total de 100, para obtener buenos resultados, se muestra su entrenamiento en la siguiente imagen.

Figura 22*Entrenamiento de Red Neuronal Tres*

```
red_12= model_12.fit(train_X12, train_y12, epochs=100, batch_size=batch_sizes, verbose=1, shuffle=True, validation_split=0.2)
```

Se especifican y se comparan los parámetros que se utilizaron para lograr entrenar a cada red neuronal en la siguiente tabla:

Tabla 3*Tabla de Parámetros de Redes Neuronales*

<i>Parámetro</i>	<i>Red neuronal uno</i>	<i>Red neuronal dos</i>	<i>Red neuronal tres</i>
	<i>Valor</i>		
<i>Epochs</i>	<i>150</i>	<i>200</i>	<i>100</i>
<i>Batch size</i>	<i>1</i>	<i>1</i>	<i>1</i>
<i>Optimizer</i>	<i>Adam</i>	<i>Adam</i>	<i>Adam</i>
<i>Loos</i>	<i>mse</i>	<i>mse</i>	<i>mse</i>
<i>LSTM UNIT 1</i>	<i>32</i>	<i>32</i>	<i>30</i>
<i>LSTM UNIT 2</i>	<i>12</i>	<i>12</i>	<i>12</i>

Nota. Se indica los parámetros utilizados en cada red neuronal.

Resultados

Para cada red neuronal se realiza las predicciones, comprendiendo que se dividió cada base de datos en tres train con los cuales la red es entrenada y contamos con dos adicionales para realizar las respectivas evaluaciones.

Figura 23

Código de Realización de Predicciones

```

### Realización de predicciones
predic_train= model_1.predict(train_X, batch_size= batch_sizes)
predic_testp = model_1.predict(test_X, batch_size= batch_sizes)
predic_valp = model_1.predict(val_x, batch_size= batch_sizes)

```

Python

```

926/926 ————— 10s 10ms/step
111/111 ————— 1s 11ms/step
109/109 ————— 1s 11ms/step

```

Primera Red Neuronal

Se emite el resumen del modelo, en el cual podemos interpretar que en la primera parte nos emite las características con las cuales entrenamos a la red neuronal, nos informa que es una red Sequential, es decir las capas se apilan de forma lineal, en la primera capa lstm_6 tiene una salida de (1,7,32) el primer valor es el tamaño del lote, el segundo representa la longitud de la ventana, y el ultimo son el número de unidades LSTM que tiene la capa. Y los parámetros 4.864 incluyen los pesos para las entradas, pesos recurrentes y sesgos.

En la segunda capa lstm_7 tiene una salida de (1,12) donde 12 son el número de unidades en la capa y tiene 2.160 parámetros.

En la capa densa está completamente conectada con una única neurona de salida y tiene 13 parámetros de pesos y sesgo.

También nos informa que el total de los parámetros son 7.037, que es la suma de los parámetros de las capas y que todos se entrenaron durante el ajuste del modelo.

Tabla 4

Resumen de Modelo de Red Neuronal General

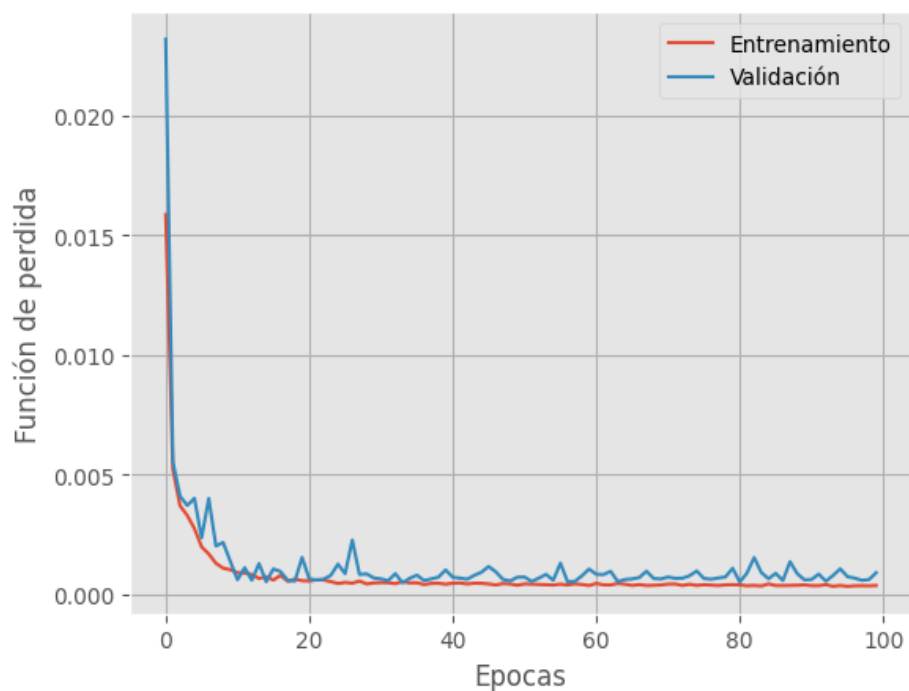
Red Neuronal uno general		
Layer	Output Shape	Parámetros
lstm_6 (LSTM)	1, 7, 32	4864
lstm_6 (LSTM)	1,12	2160
dense_3 (Dense)	1,1	13
Total, parámetros	7037	
Trainable Params	7037	

Nota. Nos muestra los parámetros del resumen del modelo de la red neuronal general.

Después de correr la red neuronal se obtiene la gráfica de aprendizaje, en la figura 25 podemos observar la función de pérdida calculada del conjunto de entrenamiento y de validación en cada época. Esta gráfica se puede observar que la curva de entrenamiento disminuye la pérdida a medida que el modelo aprende de los datos. Y en la curva de validación se observa como descende de manera similar a la de entrenamiento, pero cuenta con algunos picos suaves, pero demostrando que el modelo es capaz de generalizar bien.

Figura 24

Gráfica de Aprendizaje Red Uno

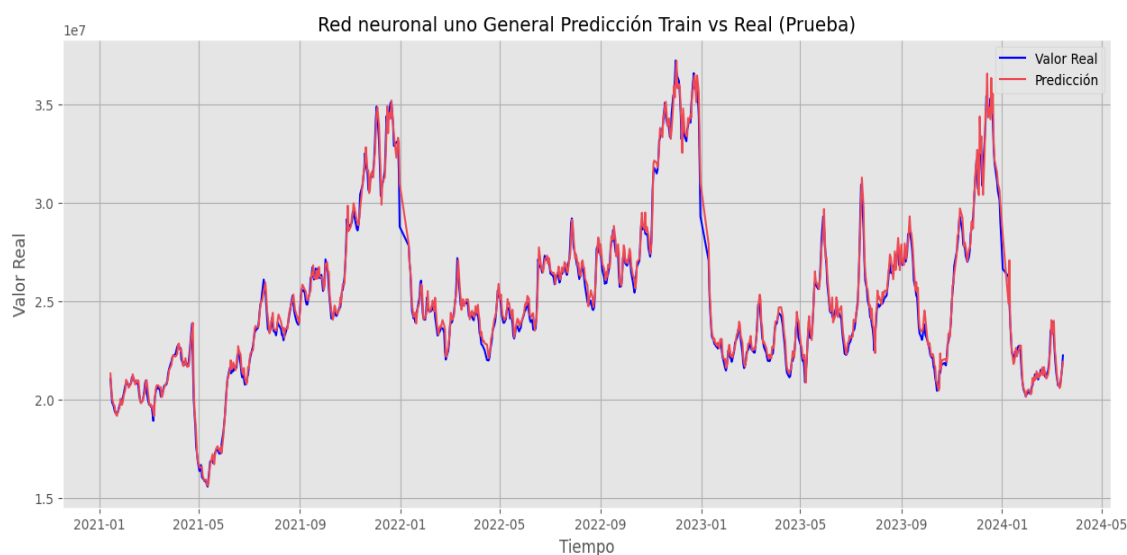


Se examina los datos obtenidos en el entrenamiento, vemos con claridad que la red está proporcionando datos correctos y muy similares a los originales. Se puede decir que el modelo está generalizando de manera correcta, además de que gracias a la gráfica de aprendizaje

sabemos que el modelo está aprendiendo bien con los datos suministrados, y lo vemos en la gráfica.

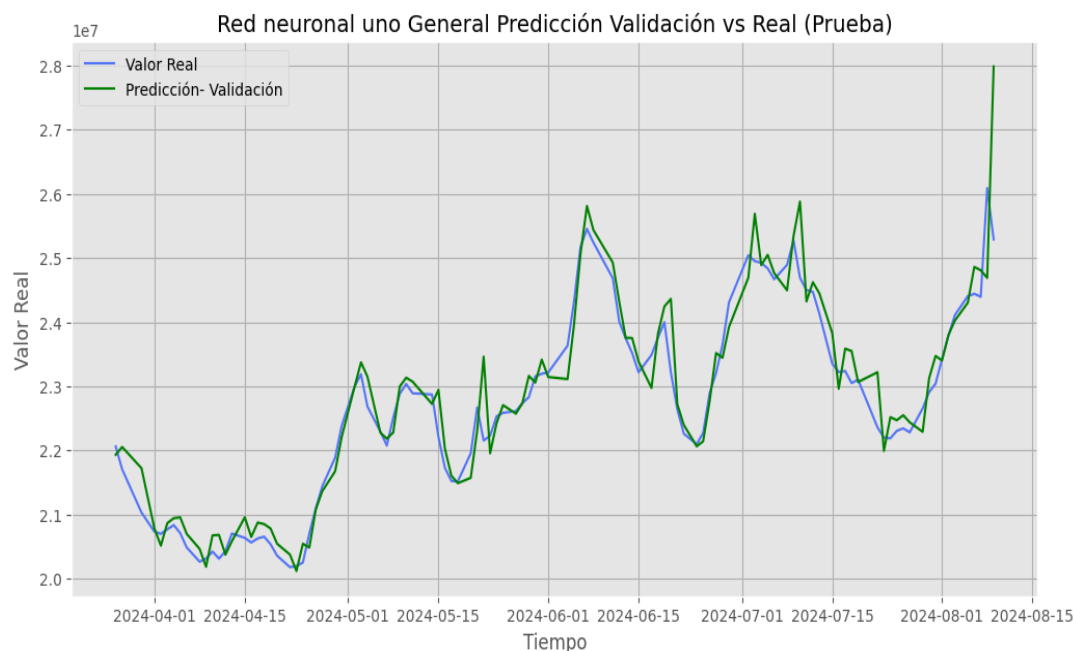
Figura 25

Datos Reales Vs Datos Predichos de Entrenamiento

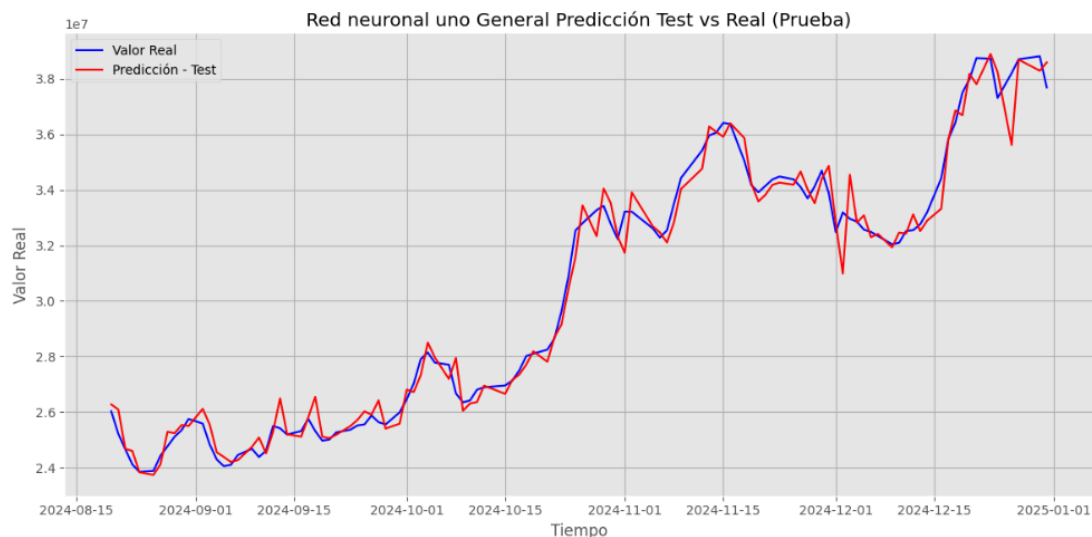


Se obtiene la gráfica de comparación entre valores reales y valores predichos, del conjunto de datos de prueba, los datos reales están de color azul y los predichos de verde, se puede observar cómo en su mayoría las líneas de prueba coinciden con los valores reales, lo que significa que el modelo logra capturar las tendencias generales del conjunto de validación.

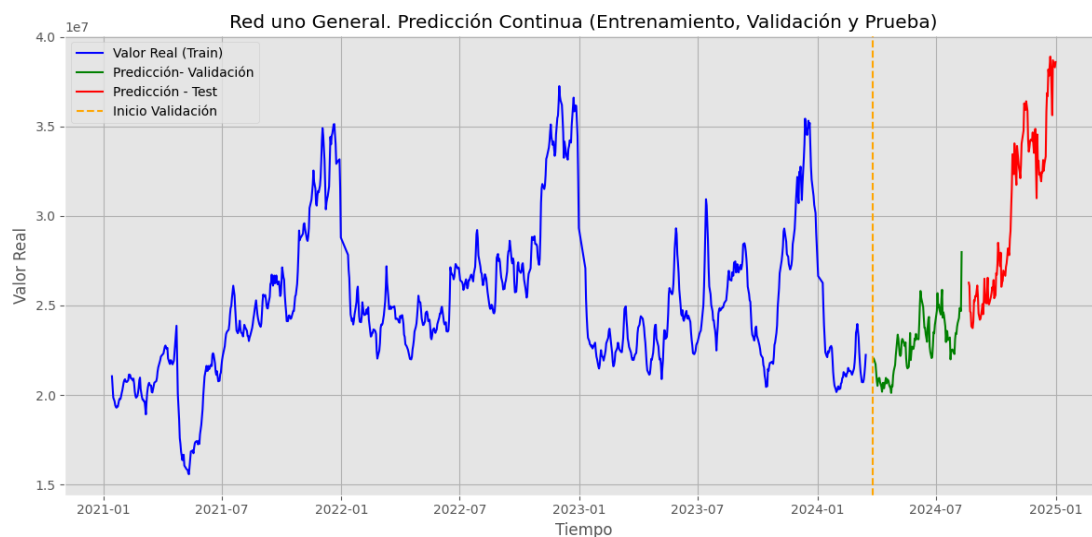
En algunas zonas las líneas divergen ligeramente, estas discrepancias parecen ser más pronunciadas en picos o valles pronunciados, A pesar de estas diferencias el modelo sigue siendo consistente en seguir la dirección general de las tendencias.

Figura 26*Datos Reales Vs Datos de Validación*

En esta gráfica tenemos la comparación del conjunto de datos de Test y los datos originales y nos sigue corroborando que el modelo de manera general es capaz de predecir y seguir el modelo de manera general, lo que nos indica que capta bien las tendencias principales y los valles y picos, a pesar de que tiene algunas discrepancias no son tan evidentes. Se estima que la gráfica indica un ajuste más preciso que con los datos de validación.

Figura 27*Datos Reales Vs Datos de Prueba*

Se une los dataset de datos originales de train y los datos predichos de validación y test para obtener la siguiente gráfica, se observa como a pesar de las pequeñas diferencias el modelo sigue su tendencia.

Figura 28*Gráfica Completa Entrenamiento Real Más Predicciones de Validación Y Test*

Se observa en la predicción una primera alza considerable para junio y julio, luego se sigue creciendo pero de manera más estable, sin embargo para incios de octubre se observa un notable aumento de monto de ventas , sin embargo se genera un ultimo pico para finales de diciembre.

Evaluación Red Neuronal General Uno

Se obtiene las métricas de RMSE, MAE y R^2 para determinar si coincide con lo observado en las gráficas.

RMSE: Nos indica que el modelo tiene errores relativamente pequeños, con valores inferiores al 3.5% en todas las etapas. El valor más alto se presenta en la validación con un 3.44% lo que puede corresponder a un ligero cambio de las características en este rango de datos. Lo que nos confirma lo visto en las gráficas donde el conjunto de validación mostraba mayor discrepancia.

MAE: Error absoluto del modelo son inferiores al 2% en todas las etapas del proceso, lo que nos indica un ajuste preciso.

R^2 : Se observa que el coeficiente de determinación en el conjunto de datos de entrenamiento y prueba son superiores a 0.98 lo que indica que el modelo captura muy bien la variabilidad de los datos, el conjunto de prueba confirma el desempeño del modelo, sin embargo, en el conjunto de validación se más bajo de 0.91 aunque sigue siendo aceptable, pero presentando diferencias entre los datos de validación y entrenamiento.

Tabla 5*Métricas RMSE, MAE y R² de Red Neuronal Uno*

Métrica red neuronal uno			
Nombre	Entrenamiento	Validación	Prueba
RMSE	2.77%	3.44%	3.16%
MAE	1.48%	1.92%	1.72%
R ²	0.989	0.91	0.980

Nota: Se relaciona las métricas de la red neuronal para poder validar la información obtenida visualmente.

Las métricas emitidas por la red neuronal uno, nos muestra que puede ser muy útil para el negocio, dado que se podría basar las compras a un presupuesto estable, gestionar de manera estratégica el inventario, lograr una planeación de ventas de manera más precisa.

Predicciones Red General Uno

Se realiza las predicciones de fechas futuras únicamente a esta red neuronal de los primeros 90 días del año para determinar su efectividad a lo largo del tiempo.

Figura 29

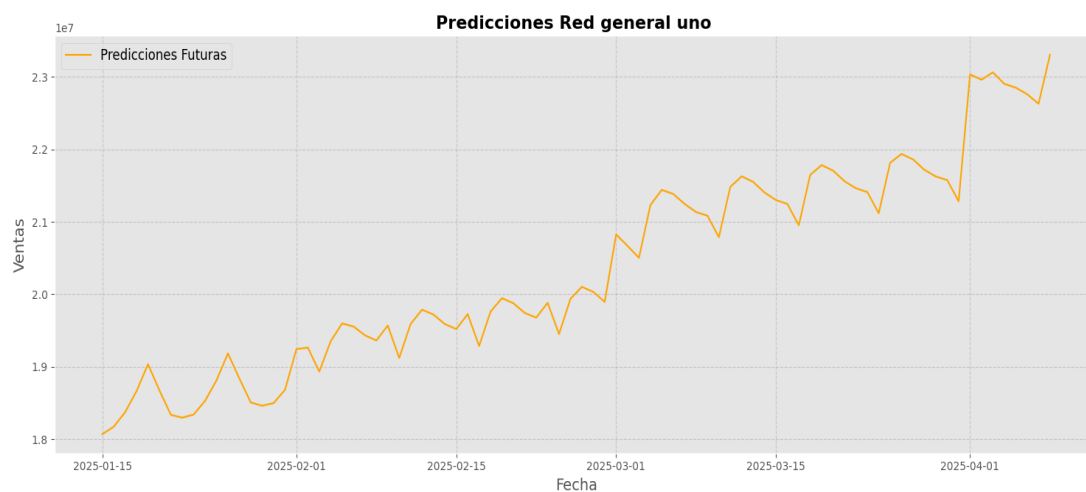
Predicción de Fechas Futuras Red General

```

# Generar predicciones
future_predictions = model_1.predict(future_X, batch_size=batch_sizes)
183] ✓ 3.5s Python
.. 84/84 ————— 2s 12ms/step

```

Se observa que los primeros días el modelo pronostica unas ventas bajas y luego va aumentando, dado que solo se esta realizando hasta el mes de abril.

Figura 30*Predicción Futura*

Se obtiene la media y la desviación estandar de estas predicciones y obtenemos que:

- La media de los valores predichos diarios es \$ 20.359.820,33
- La desviación estandar es: \$ 1.396.287,52
- Porcentaje de la desviación estándar de: 6.86%

Lo cual se considera que al tener una variabilidad del 6.86% las predicciones son consistentes y estables, con una desviación relativamente baja y corresponde a lo que en promedio diarios se vende en los primeros meses del año.

Segunda Red Neuronal

Como se menciona esta segunda red se encarga únicamente de predecir valores de venta del grupo 64, que corresponde a Sonido profesional, lo que significa cabinas activas, pasivas, parlantes entre otros, y este código representa mayores ganancias para la empresa.

El resumen del modelo describe las características empleadas para entrenar la red neuronal, se utiliza una red secuencial es decir que las capas están organizadas en un orden lineal, una después de la otras y contiene:

En la primera capa se tiene una salida con forma (1,7,30) donde la ventana es 7, dado que un número mayor nos emitía más margen de error. Y para esta capa se utilizaron 12 unidades LSTM en la capa.

En la segunda capa LSTM genera una salida (1,12) donde 12 son las unidades en la capa y esta recibe información de la primera capa y cuenta con 2064 parámetros.

Por último, la capa densa la cual esta completamente conectada con la única neurona de salida y contiene 13 parámetros. Con un total de 6517 parámetros, también nos muestra los parámetros del optimizador de 13.036 los cuales son relevantes para el proceso de ajuste del modelo.

Tabla 6

Resumen de Modelo de Red Neuronal Dos

Red neuronal dos		
Layer	Output Shape	Parámetros
lstm_2 (LSTM)	1, 7, 30	4440
lstm_3 (LSTM)	1,12	2064
dense_1 (Dense)	1,1	13
Total, parametros	19553	
Trainable Params	6517	
Non- Trainable	0	
Optimizer Params	13036	

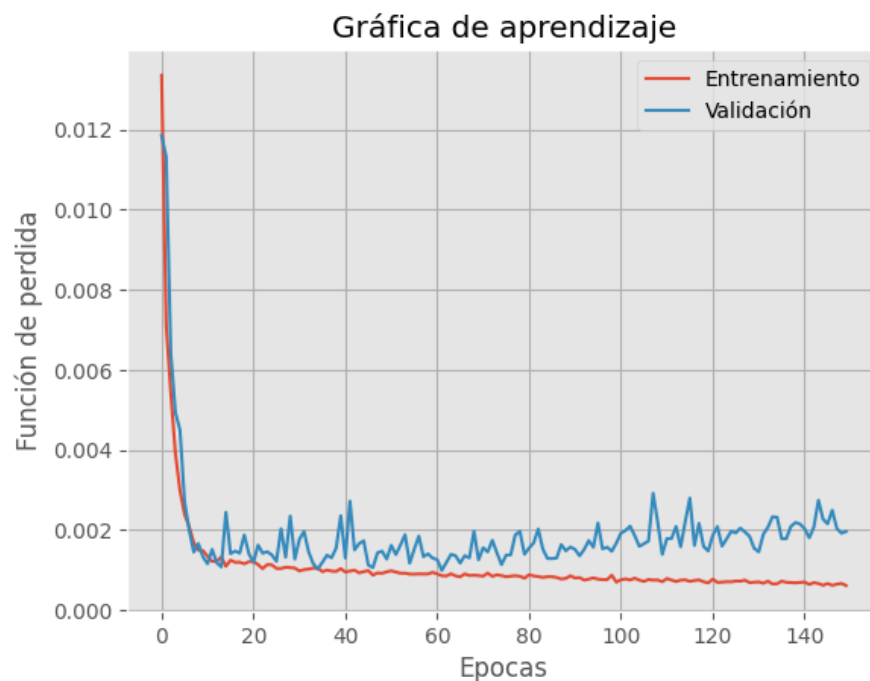
Nota: Parámetros del resumen del modelo neuronal generado

Al correr el modelo de la red neuronal se emite el gráfico de aprendizaje de entrenamiento y validación, en cual se observa que el entrenamiento cae de manera acelerada y mantiene esa disminución, lo que indica que el modelo está aprendiendo de los datos y ajustando los parámetros para minimizar el error, se estabiliza y sugiere que el modelo está convergiendo de manera efectiva, sin embargo, el aprendizaje de validación a pesar que inicia con una disminución considerable, lo que refleja que el modelo está generalizando bien en las primeras etapas del entrenamiento, a lo largo del entrenamiento presenta pequeñas oscilaciones o picos que puede ser un indicio de que el modelo enfrenta ligeros problemas de generalización en ciertos momentos.

En general el modelo está aprendiendo, pero se observa una margen para optimizar el proceso de validación y lograr reducir los picos observados.

Figura 31

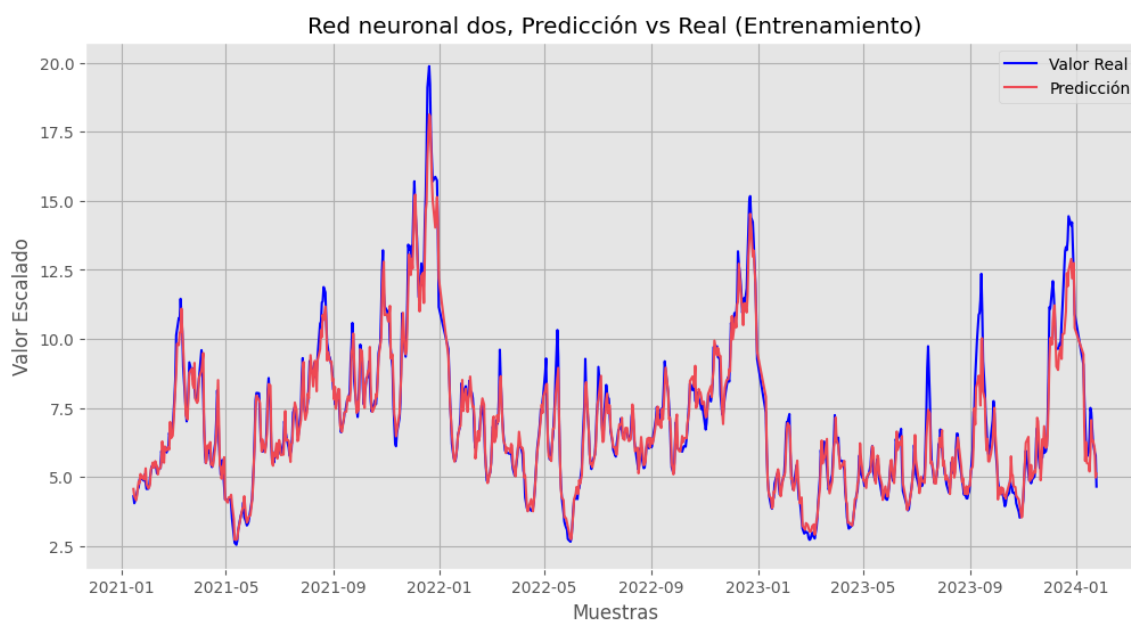
Gráfica de Aprendizaje Red Dos



Se examina la gráfica de entrenamiento y predicción, el color azul son los valores reales y el rojo la predicción están alineadas en gran parte, lo que indica que el modelo es capaz de capturar la tendencia general de los datos, reflejando que el modelo está aprendiendo correctamente los patrones de este conjunto de datos. Sin embargo, existen pequeñas discrepancias entre las líneas, especialmente en los picos más pronunciados.

Figura 32

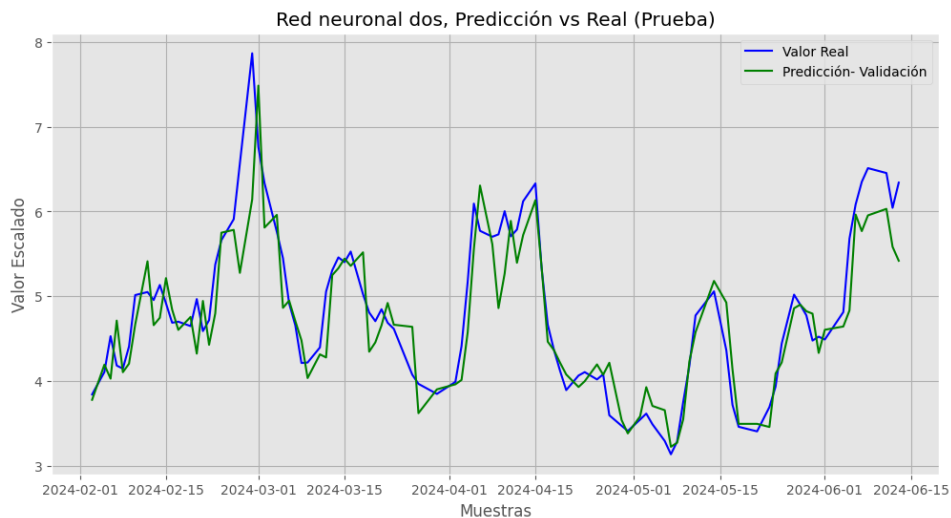
Datos Reales Vs Datos Predichos de Entrenamiento



En el gráfico de validación el color azul son los reales y los verdes la predicción para diferenciar cada fragmento de datos, en este se logra observar que de manera general el modelo se ajusta a los datos reales, lo que supone que el modelo es capaz de generalizar parcialmente los patrones aprendidos del conjunto de datos de entrenamiento, además, se observa diferencias más pronunciadas en comparación al conjunto de datos de entrenamiento.

Figura 33

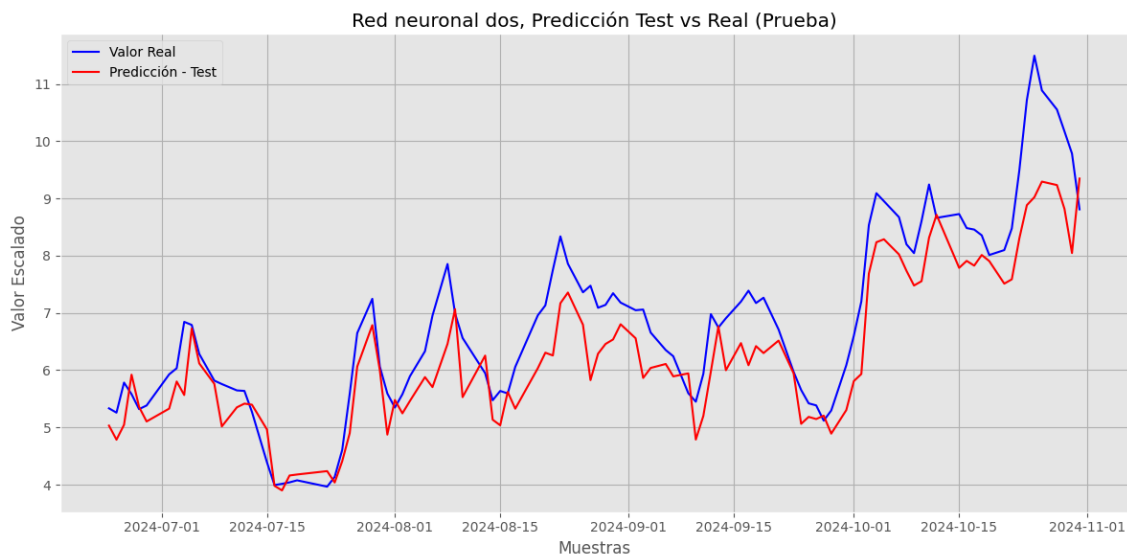
Comparación Conjunto de Validación Vs Valores Reales



El modelo sigue confirmando que tiene más dificultad para capturar las fluctuaciones más abruptas o inusuales presentes en los datos de test, dado que las diferencias se presentan en mayor medida en los picos de los datos, entonces nos muestra que al ser datos altos menos frecuentes, podría necesitar más características, como relacionar las elecciones o la semana santa, donde se observa en la gráfica un aumento en la demanda.

Figura 34

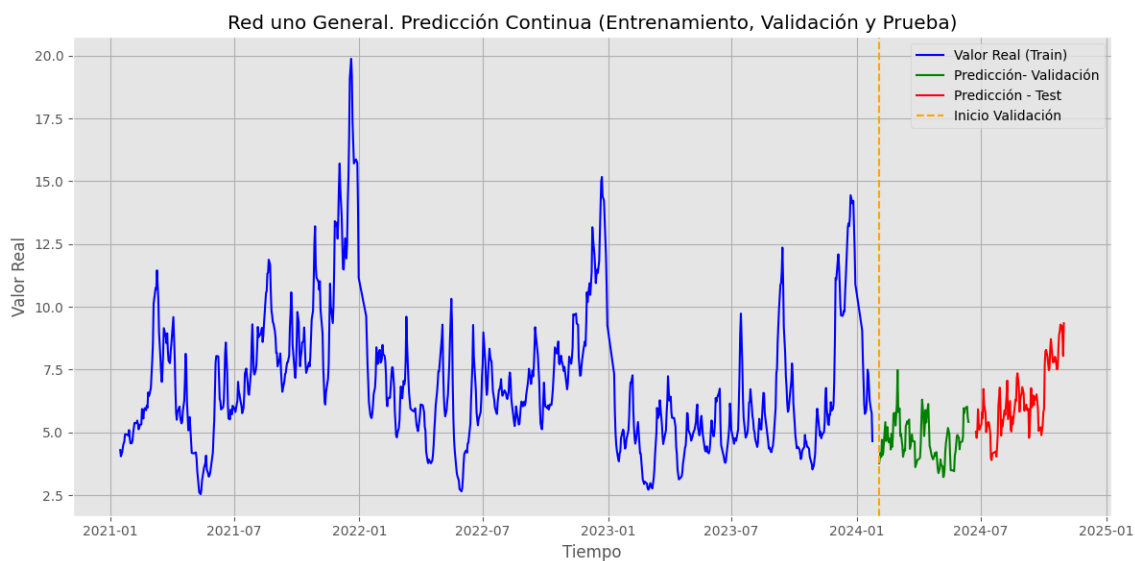
Comparación Base de Datos Test Vs Valores Originales



Se une de manera visual los datos originales de entrenamiento y los valores de predicción en validación y test.

Figura 35

Datos Continuos Entrenamiento Reales Más Predicción Validación y Test



Se observa en los datos pronosticados una pequeña alza en las ventas aproximadamente desde el mes de octubre, sin embargo para los meses de enero a septiembre se mantiene en una media estable, generando picos muy sutiles.

Evaluación Red Neuronal Dos

Análisis de RMSE, MAE y R^2 para determinar si coincide con lo observado en las gráficas.

RMSE: El error en el conjunto de entrenamiento es de 11.05% relativamente bajo, lo que nos indica que el modelo tiene un buen ajuste con los datos entrenados. Sin embargo, el error aumenta en la validación a 17.98% lo que nos confirma lo observado en las gráficas, a pesar de esto el error en el conjunto de prueba es de 14.35% lo cual se considera razonable, lo cual nos indica que el modelo generaliza bien y tiene margen de mejora.

MAE: Con esta métrica reforzamos la interpretación del RMSE y lo visualizado en las gráficas, que el modelo no captura completamente ciertas fluctuaciones, sin embargo, esta diferencia es moderada y en general tiene un buen desempeño.

R^2 : En el conjunto de entrenamiento tiene un valor de 0.96 lo que nos indica que el modelo explica casi toda la variabilidad en este conjunto. Sin embargo, esta disminuye a 0.80 en el conjunto de validación, lo cual constituye a menor capacidad para explicar los datos. Y por último el conjunto de prueba con un 0.854 el cual es un valor sólido e indica un buen equilibrio entre el ajuste y generalización.

Tabla 7*Métricas RMSE, MAE y R² de Red Neuronal Dos*

Métricas red neuronal dos			
Nombre	Entrenamiento	Validación	Prueba
RMSE	11.05%	17.98%	14.35%
MAE	5.39%	11.81%	8.48%
R ²	0.96	0.80	0.854

Nota: Se relaciona las métricas de la red neuronal para poder validar la información obtenida visualmente.

Con estas métricas nos muestra que el modelo debe ser segmentado por producto para mejorar la precisión y/o agregar más características, sin embargo nos ayuda a comprender el movimiento del grupo 64 en el almacén, se observa en las figuras el aumento en total de ventas en los últimos meses del año.

Tercer Red Neuronal

Para esta tercera red se utiliza los grupos siguientes en valor de mayor vendidos los cuales son el 12 y 15 correspondiente a paneles led e iluminación de la sede principal de iluminación. Para esta red neuronal el resumen nos emite las características utilizadas para entrenar la red neuronal y se observa que pertenece a un orden lineal.

Se tiene en la primera capa una salida de (1, 7, 30) suficientes para entrenar los datos, además de tener una segunda capa con 12 unidades y en esta se recibe los 204 parámetros. Y tiene una sola capa de salida que nos dará los valores de las predicciones.

Tabla 8*Resumen de Modelo de Red Neuronal Tres*

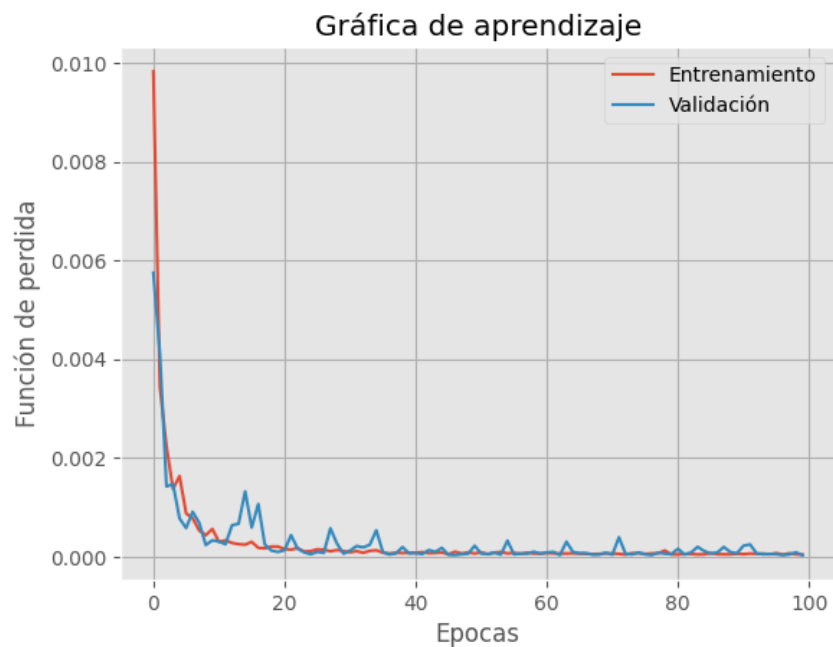
Red neuronal tres		
Layer	Output Shape	Parámetros
lstm_2 (LSTM)	1, 7, 30	4440
lstm_3 (LSTM)	1,12	2064
dense_1 (Dense)	1,1	13
Total, parametros	19553	
Trainable Params	6517	
Non- Trainable	0	
Optimizer Params	13036	

Nota. Nos muestra los parámetros del resumen del modelo de la red neuronal tres.

Al final del entrenamiento de la red neuronal se emite la gráfica de aprendizaje para verificar como trabajo y aprendió la red. Lo que se observa es que tanto el entrenamiento como la validación caen de manera abrupta, la validación presenta picos muy pequeños, lo que nos indica que el modelo está generalizando de manera correcta y no está siendo sobre ajustado el modelo.

Figura 36

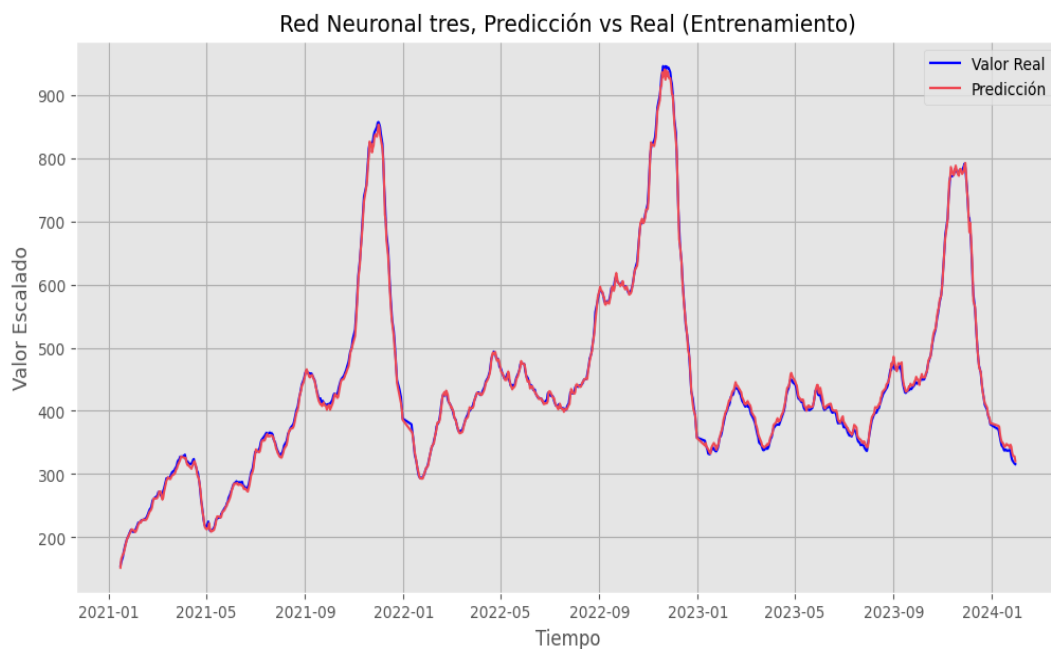
Gráfica de Aprendizaje Red Neuronal Tres



Se da paso a verificar la exactitud del modelo primero de manera visual con el conjunto de datos de entrenamiento, en este se observa que el modelo aprende de manera adecuada y eficiente, dado que no se nota diferencia entre una y otra base, demostrando que el modelo con la base entrenada es muy similar.

Figura 37

Datos Reales Vs Datos Predichos de Entrenamiento

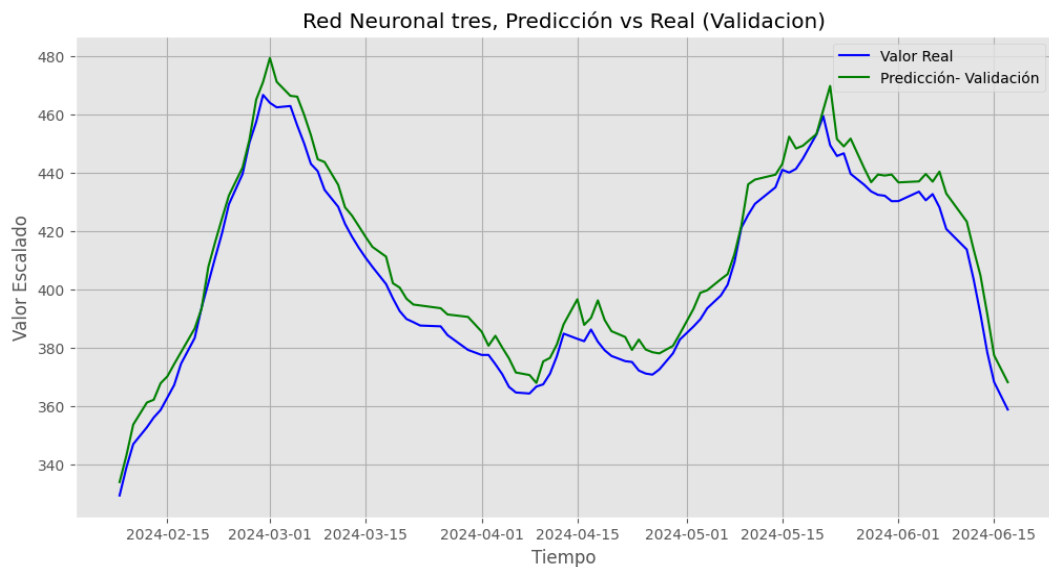


Se añaden las gráficas de validación y prueba o test para verificar como el modelo está prediciendo, se observa que en las dos gráficas las líneas de predicción se ajustan muy bien a los datos reales, además de que estos datos no presentan subidas tan grandes como las redes anteriores, se logra apreciar como la red generaliza los datos de manera más adecuada.

En esta gráfica de validación de la red neuronal tres se observa como los valores de predicción están un poco por encima de los valores reales, pero sus diferencias no son tan fuertes.

Figura 38

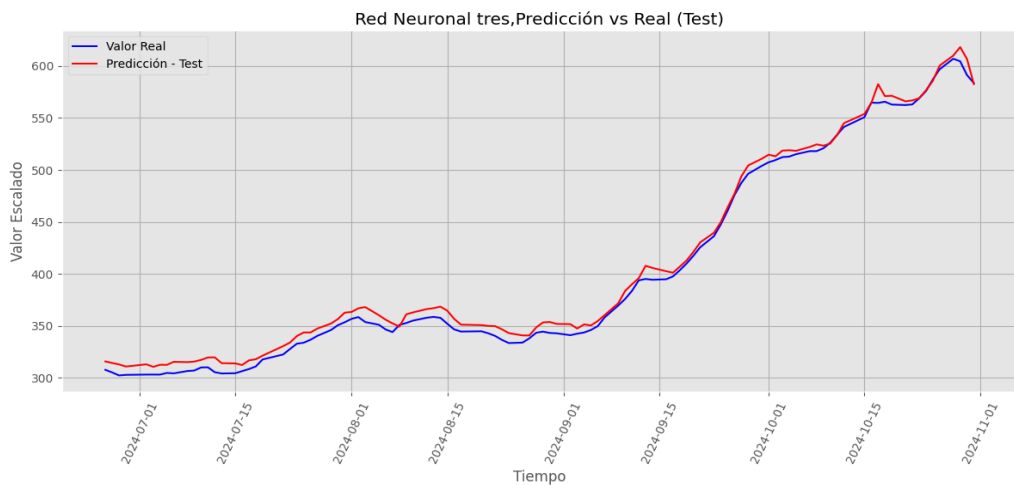
Comparación de Validación con Datos Reales



Se observa la gráfica de comparación de los valores de predicción y test de la última etapa de la red, en donde se observa mucha similitud entre ambas con algunas diferencias no muy marcadas.

Figura 39

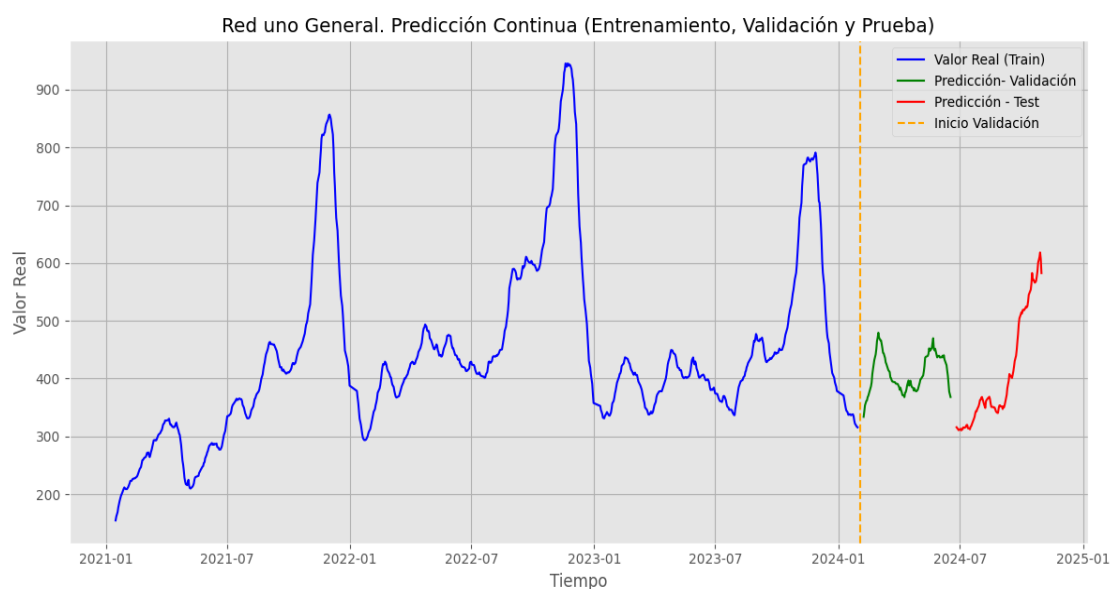
Comparación de Prueba con Datos Reales



Se une las gráficas de entrenamiento de datos originales más las predicciones de test y validación para tener un esquema completo de como estaría establecido con las predicciones.

Figura 40

Comparación de Validación con Datos Reales



Se observa que para el mes de julio van a disminuir significativamente las ventas, sin embargo desde mediados de septiembre aproximadamente se observa una alza en la demanda de los productos de iluminación.

Evaluación Red Neuronal Tres

Análisis de RMSE, MAE y R^2 para determinar si coincide con lo observado en las gráficas.

RMSE: En el conjunto de entrenamiento se obtiene 1.09% lo cual es un valor muy bajo, lo cual el modelo se ajusta muy bien a los datos que fue entrenado. Con los datos de validación están ligeramente superiores con 1.77% sigue demostrando que el modelo tiene un buen ajuste, y en el conjunto de datos de prueba se obtuvo 3.22% casi el doble de lo obtenido en el

entrenamiento, sin embargo, esta métrica indica que el modelo generaliza muy bien y sin sobreajuste o sub-ajuste.

MAE: En el error absoluto medio en el conjunto de entrenamiento su métrica es muy baja de 0.71% lo que nos sugiere que el modelo ha aprendido de manera correcta. Ya en el conjunto de prueba estos datos suben a 2.5% lo cual es consistente con un modelo que generaliza bien.

R²: Se observa de manera general que el coeficiente de determinación es superior al 0,95% lo cual nos indica que tanto las gráficas como las métricas corresponde a que el modelo generaliza de manera correcta, el modelo es robusto en datos de entrenamiento y nuevos.

Tabla 9

Métricas RMSE, MAE y R² de Red Neuronal Tres

Métricas red neuronal tres			
Nombre	Entrenamiento	Validación	Prueba
RMSE	1.09%	1.77%	3.22%
MAE	0.71%	1.3%	2.5%
R ²	0.98%	0.97%	0.96%

Nota: Se relaciona las métricas de la red neuronal para poder validar la información obtenida visualmente.

Con estas métricas obtenidas nos muestra que el modelo realiza una predicción muy precisa, lo que ayudaría a planificar los inventarios de este grupo de manera optimizada, logrando obtener una demanda más predecible, y la empresa pueda ajustar sus niveles de stock con mayor certeza.

Conclusiones

Se diseñó modelos predictivos basados en el total de las ventas de la empresa de comercio tecnológico utilizando la red neuronal recurrente LSTM, los resultados que se obtuvieron en la tabla 4, 5 y 6 nos muestra que los modelos son óptimos basándonos en la métrica de MAE, demostrando que son óptimos y tienen una buena capacidad de generalización y precisión de los modelos diseñados.

Durante el análisis exploratorio de los datos, se identifica información clave que ayuda a la toma de decisiones en la empresa. Entre los hallazgos más relevantes para la empresa:

Rango de vendedores por facturas emitidas y total de ventas en el año 2024, dejando en evidencia el rendimiento de los trabajadores y mostrando los mejores vendedores y los que deben mejorar.

Movimiento del total de ventas por sedes identificando su comportamiento a lo largo del tiempo. Se detecta un decrecimiento del 4% en retención de clientes en comparación al año anterior. Se identifican grupos de productos que generan mayores ingresos a la empresa, lo cual permitió enfocarnos en ellos en los modelos de predicción.

Además, se logra identificar patrones de estacionalidad permitiendo añadir factores temporales como características del modelo tales como día de la semana y temporada navideña en los meses de octubre a diciembre, enriqueciendo al modelo con información contextual valiosa.

El primer modelo realizado corresponde a un modelo general con todos los grupos de productos, pero no se considera cantidades vendidas como característica, este modelo mostró un alto nivel de precisión en la estimación de las ventas, logrando capturar los datos de manera efectiva. Las métricas de error absoluto (MAE) obtenidas fueron en entrenamiento 1.48% en

validación de 1.92% y en test de 1.72%, estos resultados demuestran que el modelo está correctamente ajustado, con una capacidad predictiva precisa y consistente y se pudo verificar en la gráfica 7 obtenida en comparación con los valores reales y los predichos.

Para el modelo de red neuronal dos de grupo 64 y grupos 12 y 15 se incluyó una característica adicional la cantidad vendida, En el segundo modelo grupo 64 o sonido profesional tuvo un error absoluto más elevado en comparación con el modelo general, es importante destacar la capacidad al ajustarse a los patrones específicos de este grupo. Se obtuvo la métrica de error medio en conjunto de entrenamiento de 5.39%, en validación 11.8% y en prueba 8.48%. podemos concluir que este modelo analizo el grupo con comportamiento más volátil, donde la cantidad si juega un papel importante en la predicción y deja espacio para analizar otros tipos de parametros.

En el modelo de grupos 12 y 15 correspondiente a iluminación demuestra que el modelo es altamente efectivo, con un ajuste óptimo y una capacidad predictiva superior, su métrica de error en conjunto de entrenamiento de 0.71%, en el conjunto de validación de 1.3% y en conjunto de prueba de 2.5%, En términos de métricas este modelo obtuvo los mejores resultados de los diseñados seguido por el modelo general de la empresa.

Se concluye que al usar las redes neuronales LSTM para predecir el valor total de las ventas fue una buena elección dado que las métricas obtenidas son óptimas y corresponden a buen aprendizaje de los datos con fluctuación y temporadas.

Recomendaciones

Los modelos de redes neuronales al tener buena capacidad para resolver problemas robustos, se puede analizar que otros tipos de variables exógenas se puede incluir como temporada escolar, dado que gran parte de la electrónica es impulsado por los estudiantes, además de poder incluir variaciones climáticas, lluvia o verano, y lograr incluir los periodos electorales, dado que todas estas excepciones hacen que la demanda de algunos productos aumente.

De manera general se recomienda limitar compras los primeros meses del año, y programarse para un pequeño incremento en el mes de abril y un mayor incremento para los meses de junio e inicio de julio, además que desde estos meses se observa el considerable aumento que se puede presentarse hasta final de año, sin embargo se espera picos más altos de ventas para inicios de noviembre y finales de diciembre, se recomienda prepararse con anticipación para estos periodos.

En la red neuronal dos se recomienda poder realizar un análisis predictivo más específico con productos concretos en el grupo de 'sonido', esto ayudará a tener una idea más clara de la demanda que genera y lograr obtener una buena gestión de inventarios. Sin embargo, se recomienda programarse para una posible alza de demanda en estos productos para el inicios de marzo y el mes de octubre en adelante, sin embargo se aconseja ser más cautelosos para los meses de mayo y julio, dado que se espera una baja en las ventas.

En la red neuronal tres se recomienda aumentar el presupuesto para el mes de septiembre, sin embargo se debe considerar que los montos de ventas aportados son menores al otro grupo.

Resultaría beneficioso extender la realización modelos de predicción con redes neuronales LSTM a otras áreas, como las compras, gastos, ventas por sede, siempre priorizando aquellos que tengan un mayor impacto en la eficiencia operativa. Además, dado el análisis exploratorio obtenido se recomienda realizarlo con regularidad para comprender de mejor manera el direccionamiento de la empresa y lograr tomar decisiones informadas a tiempo, logrando redireccionar objetivos y si es necesario replantear estrategias que se adapten a la necesidad.

Referencias Bibliográficas

- Aguilar, R. (1999). *Red Neuronal de Topología Flexible*. VI Congreso Nacional de Ciencias de la Computación, La Paz, Bolivia.
- Awan, A. (2024). *Explicación del ciclo de vida del machine Learning*. Datacamp
<https://www.datacamp.com/es/blog/machine-learning-lifecycle-explained>
- Bishop, C. M. (1995) *Neuronal Networks for Pattern Recognition*, Clarendon Press, Oxford
- Briones, J. (2023). *Construcción de un modelo de predicción para activos financieros*. Trabajo de fin de máster. Universidad da Coruña.
https://ruc.udc.es/dspace/bitstream/handle/2183/36522/Briones_Viqueira_Jorge_TFM_2023_Construccion%20de%20un%20modelo%20de%20prediccion%20de%20activos%20financieros.pdf?sequence=2&isAllowed=y
- CAAD. (2020), *Machine Learning en retail ¡Impulsa tus ventas!* <https://caad-design.com/formas-de-impulsar-las-ventas-con-machine-learning-en-retail/>
- Caio, R. (2022). *Crisp-DM: las 6 etapas de la metodología del futuro*. MBA USP.
<https://blog.mbauspesalq.com/es/2022/05/31/crisp-dm-las-6-etapas-de-la-metodologia-del-futuro/>
- Calvo, D. (2017). *Clasificación de redes neuronales artificiales. Aprendizaje Automático*, Redes neuronales. <https://www.diegocalvo.es/clasificacion-de-redes-neuronales-artificiales/>
- Camarrillo, M. (2022). *Pronósticos de venta*. Infor Blog
<https://www.infor.com/latam/blog/sales-forecasts>
- Camones, F. (2002). *Control de calidad en los procesos estadísticos*. Centro de investigación y desarrollo (CIDE).

https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0516/Libro.pdf

Castañeda Rojas, Miguel Angel David, (2020), *Machine Learning para la gestión de ventas en la Empresa Vértice Empresarial S.A.C.*

https://alicia.concytec.gob.pe/vufind/Record/UCVV_77f39fccfd154ed8c3bba6c542338c00/Details

Castro Morales, Uriel Mauricio, (2024), *Pronóstico de ventas a corto plazo en el sector de consumo masivo a través de modelos de machine Learning.*

<https://repositorio.uniandes.edu.co/entities/publication/4a60c5ef-2a11-4f1a-91e2-d64ebe91591f>

Carretero, J. (2023). *Desarrollo de un sistema de predicción de ventas para tiendas físicas, Trabajo de fin de grado*, Universidad Complutense de Madrid.

<https://docta.ucm.es/rest/api/core/bitstreams/e5a44991-dc04-4ed4-8a80-866165f98d4f/content>

Cazares Baldeón, Cristian Geovanny, (2023), *Software de predicción de ventas en una PYME basado en técnicas de aprendizaje automático.*

<https://dspace.ups.edu.ec/handle/123456789/26429>

Conceptos básicos sobre redes neuronales. Grupo US

<https://grupo.us.es/gtocom/pid/pid10/RedesNeuronales.htm>

Correa Loaiza, (2023), *Análisis de modelo de pronóstico para la predicción de la demanda de productos en la empresa Dyna y Cía.*, Trabajo de grado especialización, Especialización en Analítica y Ciencia de Datos, Universidad de Antioquia, Medellín,

https://bibliotecadigital.udea.edu.co/bitstream/10495/35745/3/CorreaAlejandro_2023_ModelosPrediccionDemanda.pdf

EGOS BI. (2022). *¿El machine Learning en ventas brindar mayor competitividad comercial?*

<https://www.egosbi.com/machine-learning-en-ventas-competitividad/>

Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). *Time-series forecasting of seasonal items sales using machine learning – A comparative analysis*. International Journal of

Information Management Data Insights, Volumen 2(1), 100058.

<https://doi.org/10.1016/j.rineng.2022.100027>

Fernandez Salguero, R. A (2021). *Series Temporales Avanzadas: Aplicación de Redes Neuronales para el Pronóstico de Series de Tiempo*. (Trabajo de Fin de Máster). Universidad de Granada, Departamento de Estadística e Investigación Operativa.

https://masteres.ugr.es/estadistica-aplicada/sites/master/moea/public/inline-files/TFM_Fernandez%20SalgueroRicardo%20Alonzo.pdf

Filtrado y suavizado de Kalman. (2024). *Fiveable*. [https://library-fiveable-](https://library-fiveable-me.translate.google.com/advanced-signal-processing/unit-7/kalman-filtering-smoothing/study-guide/RBIDWSAyiTyoqvE4?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sge&_x_tr_hist=true)

[me.translate.google.com/advanced-signal-processing/unit-7/kalman-filtering-smoothing/study-guide/RBIDWSAyiTyoqvE4?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sge&_x_tr_hist=true](https://library-fiveable-me.translate.google.com/advanced-signal-processing/unit-7/kalman-filtering-smoothing/study-guide/RBIDWSAyiTyoqvE4?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sge&_x_tr_hist=true)

Frank, La La (2019) *¿Cómo aprenden las redes neuronales?* Microsoft Learn. Vol 34 N° 4.

<https://learn.microsoft.com/es-es/archive/msdn-magazine/2019/april/artificially-intelligent-how-do-neural-networks-learn>

Hayes, Adam. (2024). *Autoregressive Integrated Moving Average (ARIMA) Prediction Mode*,

Investopedia. <https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp>

- Infante, F. A. (2018). *Propuesta de una Metodología para el Pronóstico de la Demanda Utilizando Método de Series de Tiempo en Laboratorios La Santé S.A.*
<http://hdl.handle.net/11349/13623>
- Intuit Mailchimp. *Aprendizaje automático: revoluciona tu marketing.*
<https://mailchimp.com/es/resources/machine-learning-in-marketing-guide/>
- Jorge Botana, Guillermo de. 2024. *Redes neuronales recurrentes y Transformers para modelos cognitivos del lenguaje.* Madrid: Ediciones Complutense
- Llunitasig, M. (2021). *Simulación de pronósticos de ventas en la empresa impactex mediante redes neuronales.* Trabajo de Magíster. Universidad Técnica de Ambato.
<https://repositorio.uta.edu.ec/server/api/core/bitstreams/8a2e5a4d-b2e8-48d8-ac15-cb405921dcfe/content>
- Lopera Hernández, W. (2023). *Comparación de metodologías basadas en una red neuronal artificial y un modelo GARCH para el pronóstico de la volatilidad del precio de las acciones cotizados en la BVC.* Universidad Nacional de Colombia.
- Mesquita, Renato. (2018). *¿Qué es Marketing? Una guía completa del concepto, tipos, objetivos y estrategias.* https://rockcontent.com/es/blog/marketing-2/#google_vignette
- McCulloch y Pitts . (2022). *La equivalencia formal en el lenguaje de las neuronas lógicas de Epistemología E Historia De La Ciencia*, 7(1), 22-40. <https://doi.org/10.61377/ehc.34428>
- Nalda, Victor. (s.f). *Machine Learning: Los orígenes y la evolución.*
<https://www.futurespace.es/machine-learning-los-origenes-y-la-evolucion/>
- Negrete Godoy, Gabriela Paz; (2022), *Diseño de un modelo Machine Learning para mejorar la venta de rutas diarias en la Startup de logística Wareclouds.*
<https://repositorio.uchile.cl/handle/2250/191396>

Neurons and Nerves. *Ask a Biologist*. Arizona State University.

<https://askabiologist.asu.edu/neuron-anatomy>

Noble Joshua, (2024), *¿Qué son los modelos ARIMA*, IBM <https://www.ibm.com/mx-es/topics/arima-model>

Optimizador Adam. (s.f) *Ultralytics*. <https://www.ultralytics.com/es/glossary/adam-optimizer#:~:text=El%20Optimizador%20Adam%20es%20un,del%20entrenamiento%20de%20redes%20neuronales>.

Outliersctliers. *5 ejemplos de Machine Learning y cómo ayuda a tus ventas*.

<https://www.outliersconsulting.com/general/5-ejemplos-de-machine-learning-y-como-ayuda-a-tus-ventas/>

Pacheco, V. (2019). *Una Breve Historia del Machine Learning*. Blog Telefonica Tech.

<https://telefonicatech.com/blog/una-breve-historia-del-machine-learning>

Palma, R. (2022), *Arima Estacionalidad*. Uncu, Argentina

<https://themys.sid.uncu.edu.ar/rpalma/MBA/Evaluaciones%202022/Arima/Arima2.html>

Ponce, G. (2022). *Modelo basado en machine learning para optimizar el pronóstico de ventas de la empresa ricos pan, año 2020 – 2021*. Trabajo de fin de grado. Universidad Nacional del Antiplano.

https://repositorio.unap.edu.pe/bitstream/handle/20.500.14082/19132/Ponce_Illacutipa_Gerson_Manuel.pdf?sequence=3&isAllowed=y

Ramírez, U. Tello-Leal, E. Ríos Alvarado, A. (2018), *Modelo basado en redes neuronales recurrentes lstm para la predicción de la siguiente actividad en procesos de negocio*

(lstm recurrent neural network based-model for the prediction of the next activity in

- business processes). *Pistas Educativas*, Vol. 40. Núm. 130
<https://pistaseducativas.celaya.tecnm.mx/index.php/pistas/article/view/1736>
- Redes Neuronales LSTM. (s.f) *CiberseguridadMax*. <https://ciberseguridadmax.com/lstm/>
- Rodrigo Cadena Martínez, (2023), *Modelo de Ciencia de Datos para predecir ventas en una Empresa*, *Ciencia Latina, Revista científica Multidisciplinar* 7 (4): 9374-9393
https://www.researchgate.net/publication/374225173_Modelo_de_Ciencia_de_Datos
- Rodriguez, Ruben Abril (2018), *Neuronas de McCulloch y Pitts*.
<https://lamaquinaoraculo.com/deep-learning/el-modelo-neuronal-de-mcculloch-y-pitts/>
- Rojas-Jiménez, K. 2022. *Ciencia de Datos para Ciencias Naturales*.
https://bookdown.org/keilor_rojas/CienciaDatos/
- Ruelas, S., & Laguna, G. (2014), *Comparación de predicción basada en redes neuronales contra métodos estadísticos en el pronóstico de ventas. Actualidad y nuevas tendencias*, Vol. IV, N°12, 91-105 <http://servicio.bc.uc.edu.ve/ingenieria/revista/Industrial/voliv12/art07.pdf>
- Russell, S. J., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson Education. https://people.engr.tamu.edu/guni/csce421/files/AI_Russell_Norvig.pdf
- Salazar, M. y Cabrera, M. (2007), *Pronóstico de demanda por medio de redes neuronales artificiales*. ResearchGate. División de Posgrado en Ingeniería de Sistemas, FIME-UANL
https://www.researchgate.net/publication/236028213_Pronostico_de_demanda_por_medio_de_redes_neuronales_artificiales
- Saadeddin, Z. (2024). *ARIMA para la previsión de series temporales: Guía completa*, Datacamp.
<https://www.datacamp.com/es/tutorial/arima>

Silva, D.(2022). *¿Para qué sirve un pronóstico de ventas? [4 USOS VENTAJOSOS]*. Blog de Zendesk <https://www.zendesk.com.mx/blog/pronostico-de-ventas/>

Sotaquirá, M. *La función de activación*. Creando Bits. <https://codificandobits.com/blog/funcion-de-activacion/>

Tecnología Bi. *Potenciando los pronósticos de ventas con Machine Learning*.

<https://tecnologiabi.com/potenciando-los-pronosticos-de-ventas-con-machine-learning/>

Tema 6: *Modelización con datos de series temporales*. (2023). Universidad Complutense de Madrid https://www.ucm.es/data/cont/docs/518-2013-10-25-Tema_6_EctrGrado.pdf

Thkar, C (2024). *Kalman Filter Python: Tutorial and Strategies*. Quant Insti.

<https://blog.quantinsti.com/kalman-filter/>

Toro, Nicolas. (2024). *Modelos de Forecast en Python: Una Guía para Expertos*.

[https://medium.com/@nicolastoro7/modelos-de-forecast-en-python-una-gu%C3%ADa-para-expertos-](https://medium.com/@nicolastoro7/modelos-de-forecast-en-python-una-gu%C3%ADa-para-expertos-ebb8f57dfc83#:~:text=El%20modelo%20ARI0en%20su%20propio%20historial)

[ebb8f57dfc83#:~:text=El%20modelo%20ARI0en%20su%20propio%20historial](https://medium.com/@nicolastoro7/modelos-de-forecast-en-python-una-gu%C3%ADa-para-expertos-ebb8f57dfc83#:~:text=El%20modelo%20ARI0en%20su%20propio%20historial)

Zhao, K. Wang, C. (2017), *Sales Forecast in E-commerce using Convolutional Neural Network*.

Cornell University. College of Computer Science Zhejiang University Hangzhou,

<https://arxiv.org/abs/1708.07946>