UNIVERSIDAD TECNOLÓGICA DE PANAMÁ

FACULTAD DE INGENIERÍA DE SISTEMAS DE COMPUTACIONALES

MAESTRÍA EN ANALÍTICA DE DATOS

PROYECTO FINAL MODELOS PREDICTIVOS

MODELO PREDICTIVO DE VENTAS POR PRODUCTO

ESTUDIANTE:

JEANCARLOS GUNTER 8-887-1969

PROFESOR

JUAN MARCOS CASTILLO, PhD

AÑO LECTIVO

2025

# Introducción

La predicción de ventas es un proceso fundamental en las estrategias comerciales de cualquier empresa del sector minorista, ya que tiene un impacto directo en las decisiones operativas y financieras. La capacidad de prever las ventas de productos no solo permite una asignación eficiente de inventarios, sino que también facilita la optimización de estrategias de marketing, la planificación de promociones y la distribución de productos dentro de las tiendas. Una predicción precisa de ventas es clave para reducir costos relacionados con el exceso o la falta de inventario, lo que puede evitar tanto pérdidas de ventas como costos innecesarios de almacenamiento.

En el caso específico de Big Mart, la empresa busca desarrollar un modelo predictivo avanzado en Python para anticipar las ventas de sus productos, utilizando una serie de variables clave. El modelo tomará en cuenta características como el tipo y peso del producto, su contenido de grasa, su visibilidad, el precio de venta, así como la información relacionada con la tienda, como el tamaño del establecimiento, el tipo de outlet y el año de establecimiento. Estas variables, al ser analizadas a partir de datos históricos de ventas, permitirán identificar patrones que mejoren la precisión de las predicciones.

El propósito final es construir un modelo que sea capaz de predecir las ventas de productos en los distintos puntos de venta de Big Mart, considerando la combinación de características de cada producto y tienda. Con esta predicción precisa, la empresa podrá optimizar la distribución de productos en las tiendas, ajustar de manera más efectiva sus estrategias de marketing y promociones, y en última instancia, mejorar la experiencia de compra de sus clientes. Al anticipar las demandas con mayor exactitud, Big Mart podrá no solo aumentar sus ventas, sino también fortalecer la percepción de la marca, asegurando que los productos más demandados estén siempre disponibles para sus clientes.

Contenido

[Introducción 2](#_Toc195026637)

[Justificación 4](#_Toc195026638)

[Antecedentes 4](#_Toc195026639)

[Definición del problema 6](#_Toc195026640)

[Análisis Predictivo 6](#_Toc195026641)

[Base de Datos a Utilizar 6](#_Toc195026642)

[Preprocesamiento y limpieza 7](#_Toc195026643)

[Eliminación de valores nulos o faltantes 7](#_Toc195026644)

[Análisis Descriptivo 9](#_Toc195026645)

[Graficas de Visualización 12](#_Toc195026646)

[Gráfico de Correlación 15](#_Toc195026647)

[Selección de variables 16](#_Toc195026648)

[Selección de Modelos 18](#_Toc195026649)

[Graficas Tras Aplicar los modelos 19](#_Toc195026650)

[Dispersion 19](#_Toc195026651)

[Gráfico de líneas Ventas Reales y Predichas por ItemMRP y Ventas 20](#_Toc195026652)

[Grafica de Residuos 21](#_Toc195026653)

[Ventas Reales Vs PRedichas 22](#_Toc195026654)

[Conclusiones 23](#_Toc195026655)

[Recomendaciones y futuros estudios 24](#_Toc195026656)

[Recomendaciones: 24](#_Toc195026657)

[Futuros Estudios: 24](#_Toc195026658)

[Bibliografía 26](#_Toc195026659)

# Justificación

La justificación de este trabajo radica en la creciente necesidad de las empresas del sector minorista de optimizar sus operaciones para mejorar su rentabilidad y eficiencia. La predicción precisa de las ventas es una herramienta clave que permite a las organizaciones anticipar las fluctuaciones de la demanda, ajustando sus estrategias de inventario, marketing y distribución en función de las expectativas de ventas. En este contexto, un modelo predictivo robusto puede ser crucial para lograr una toma de decisiones más informada y estratégica.

Big Mart, como empresa del sector minorista, se enfrenta al reto de gestionar un catálogo extenso de productos y un número elevado de puntos de venta. Este escenario implica un alto grado de complejidad en la gestión del inventario y la planificación de estrategias comerciales. La implementación de un modelo de predicción de ventas permitirá a Big Mart optimizar estos procesos y, por ende, mejorar su rentabilidad. Un modelo predictivo preciso puede ayudar a anticipar no solo las ventas futuras, sino también las necesidades específicas de cada tienda, lo que facilita una mejor asignación de inventarios y reduce el riesgo de tanto exceso como escasez de productos.

El análisis de datos históricos de ventas, considerando factores como las características de los productos (peso, contenido de grasa, visibilidad, precio) y las características de los puntos de venta (tamaño, tipo de establecimiento, ubicación), proporciona una visión más detallada y completa de las variables que influyen en las ventas. Estos factores, aunque en algunos casos pueden ser subestimados, juegan un papel esencial en la demanda de los productos y, por tanto, en las predicciones de ventas. Al integrar esta información en el modelo, Big Mart puede obtener un entendimiento más preciso de los patrones de compra de los clientes, lo que facilita la toma de decisiones más efectivas y adaptadas a las condiciones de cada tienda.

# Antecedentes

Los antecedentes de la predicción de ventas en el sector minorista han sido ampliamente investigados y aplicados en diversas empresas a lo largo de los años, utilizando diferentes enfoques y técnicas de modelado. El objetivo de estas investigaciones ha sido siempre mejorar la toma de decisiones estratégicas, optimizar la gestión del inventario, y ajustar las estrategias de marketing y ventas a las demandas del mercado. A continuación, se presentan algunos de los antecedentes más relevantes en este campo:

1. **Modelos de predicción de ventas tradicionales**: Durante décadas, los modelos de predicción de ventas se basaron principalmente en técnicas estadísticas tradicionales, como la regresión lineal, la descomposición de series temporales (como el modelo ARIMA) y los modelos de suavizamiento exponencial. Estas metodologías eran útiles para identificar tendencias y patrones estacionales en las ventas, pero carecían de la flexibilidad para incorporar múltiples factores explicativos, como las características del producto o del punto de venta. Sin embargo, fueron un primer paso importante en el análisis predictivo de ventas.
2. **Uso de redes neuronales y aprendizaje automático**: Con el avance de la tecnología y el crecimiento de la capacidad computacional, las técnicas de **aprendizaje automático** y **redes neuronales** han comenzado a jugar un papel crucial en la predicción de ventas. Modelos como los **árboles de decisión**, **máquinas de soporte vectorial (SVM)**, **redes neuronales artificiales (ANN)** y **bosques aleatorios** han permitido a las empresas manejar grandes volúmenes de datos y considerar un mayor número de variables para predecir las ventas de manera más precisa. Estos enfoques han demostrado ser especialmente efectivos en escenarios con múltiples variables complejas y no lineales, como los que se encuentran en el ámbito minorista.
3. **Modelos híbridos y sistemas de recomendación**: Otro avance significativo en la predicción de ventas en el sector minorista ha sido el uso de **modelos híbridos** que combinan enfoques estadísticos y de aprendizaje automático. Estos modelos permiten la integración de múltiples tipos de datos, como los históricos de ventas, las características de los productos, las promociones, las tendencias del mercado y el comportamiento de los consumidores. Además, los **sistemas de recomendación** basados en **filtrado colaborativo** y **filtrado basado en contenido** también han sido adaptados para prever la demanda de productos, ayudando a las empresas a ajustar sus inventarios y estrategias de marketing.
4. **Aplicaciones específicas en el sector minorista**: En el ámbito minorista, empresas como **Walmart** y **Target** han implementado modelos de predicción de ventas avanzados utilizando técnicas de big data y aprendizaje automático. Estos modelos no solo analizan las ventas históricas de productos, sino que también consideran factores como el clima, eventos especiales, promociones y el comportamiento de compra de los clientes. Este enfoque multidimensional ha permitido a estas empresas mejorar la precisión de sus previsiones de ventas, optimizar la distribución de productos y reducir los costos asociados con el exceso o la falta de inventario.
5. **Desarrollo de modelos específicos para cadenas de tiendas**: En cuanto a la implementación de modelos predictivos en cadenas de tiendas, diversas investigaciones han demostrado que considerar las características específicas de cada tienda (como el tamaño, el tipo de outlet, y el tipo de productos que venden) es esencial para obtener predicciones más precisas. Un estudio llevado a cabo por **Goh et al. (2019)** exploró cómo las variables de cada tienda y las características de los productos influían en la predicción de ventas, sugiriendo que modelos que integran esta información contextualizada ofrecen mejores resultados en términos de optimización de inventarios y planificación de la demanda.

# Definición del problema

El problema central de este análisis consiste en predecir las ventas de productos dentro de una cadena de tiendas, considerando diversas características tanto del producto como del punto de venta. Aunque factores como el precio, el tipo de producto y las características de la tienda son ampliamente reconocidos como influyentes en las ventas, una predicción precisa requiere un análisis integral de todas estas variables. En particular, el peso, el tipo de producto, el contenido de grasa, el precio, así como características de la tienda como el tamaño, la ubicación y el tipo de establecimiento, son determinantes clave para entender el comportamiento de compra de los consumidores.

El reto principal es construir un modelo predictivo que considere una combinación de estas variables de manera efectiva para predecir las ventas de cada producto en cada tienda. Este modelo debe ser capaz de integrar estos factores de manera que se logre una visión más completa de los patrones de compra y demanda. El objetivo final es proporcionar a la empresa herramientas que permitan optimizar la distribución de productos en las tiendas, ajustar las estrategias de marketing y promoción de manera más eficiente, y en última instancia, maximizar las ventas al anticipar las necesidades de inventario y las preferencias de los clientes.

# Análisis Predictivo

## Base de Datos a Utilizar

**Big Mart Sales (**[**Big Mart Sales Prediction Datasets**](https://www.kaggle.com/datasets/shivan118/big-mart-sales-prediction-datasets/data?select=train.csv)**) Train.Csv**

Esta Base recopila las ventas de productos en diferentes puntos de venta de la cadena BigMart , esta información fue tomada en 2013 con un total de 10 puntos de venta y un total de 1559 productos los cuales poseen características únicas que se utilizaran para realizar las predicciones.

Columnas:

1. Item\_Identifier ---- ID único del producto: Un identificador único para cada producto.
2. Item\_Weight ---- Peso del producto: El peso del producto.
3. Item\_Fat\_Content ---- Contenido de grasa del producto: Indica si el producto es bajo en grasa o no.
4. Item\_Visibility ---- Visibilidad del producto: El porcentaje del área total de exhibición de todos los productos en una tienda asignado a un producto en particular.
5. Item\_Type ---- Categoría del producto: La categoría a la que pertenece el producto (por ejemplo, bebidas, alimentos, etc.).
6. Item\_MRP ---- Precio máximo al por menor (precio de lista) del producto: El precio de venta recomendado del producto.
7. Outlet\_Identifier ---- ID único de la tienda: Un identificador único para cada tienda.
8. Outlet\_Establishment\_Year ---- Año de establecimiento de la tienda: El año en que se estableció la tienda.
9. Outlet\_Size ---- Tamaño de la tienda: El tamaño de la tienda en términos de área de terreno cubierta.
10. Outlet\_Location\_Type ---- Tipo de ubicación de la tienda: El tipo de ciudad donde se encuentra la tienda (por ejemplo, ciudad pequeña, ciudad grande, etc.).
11. Outlet\_Type ---- Tipo de tienda: Indica si la tienda es solo una tienda de comestibles o algún tipo de supermercado.
12. Item\_Outlet\_Sales ---- Ventas del producto en una tienda particular: Las ventas del producto en esa tienda en particular. Esta es la variable de resultado que se debe predecir.

## Preprocesamiento y limpieza

### Eliminación de valores nulos o faltantes

En esta etapa se realizo la verificación de columnas con datos en blanco en los que se encontró que las columnas de ItemWeight, OutletSize poseían valores nulos

Texto

El contenido generado por IA puede ser incorrecto.

Ademas se comprobó si existían columnas duplicadas

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

* Se remplazan los valores nulos encontrados en las columnas antes mencionadas con la media,
* Se remplazan los valores iguales a 0 para la columna ItemVisibility
* Se hace un ajuste a los datos de ItemFatContent ya que existían valores con la misma identificación pero escritas de maneras distintas

Por Ultimo se escriben los datos en un nuevo archivo de Excel para continuar trabajando el modelo

Captura de pantalla de computadora

El contenido generado por IA puede ser incorrecto.

## Análisis Descriptivo

Se describen las variables existentes en la fuente de datos, se observan 4 variables flotantes, 1 variable entera y 7 variables categóricas

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Variables categóricas vs Variables numéricas

Texto

El contenido generado por IA puede ser incorrecto.

Se verifica la frecuencia de las variables en las columnas categóricas

Texto

El contenido generado por IA puede ser incorrecto.

Se Utiliza el método describe() para calcular datos estadísticos como el percentil, la media y la desviación estándar de los valores numéricos de la serie o el DataFrame. Analiza tanto series numéricas como de objetos, así como conjuntos de columnas del DataFrame de tipos de datos mixtos.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

### Graficas de Visualización

Gráfico, Histograma

El contenido generado por IA puede ser incorrecto.

Gráfico, Histograma

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de barras, Histograma

El contenido generado por IA puede ser incorrecto.

Gráfico

El contenido generado por IA puede ser incorrecto.

Gráfico, Histograma

El contenido generado por IA puede ser incorrecto.

## Gráfico de Correlación

Se observa una relacion del 0.57 entre las variables ItemMRP y ItemOutletSales

Esta correlación de 0.57 sugiere que existe una relación moderadamente fuerte entre las dos variables, aunque no es perfecta, los cambios en una variable tienden a estar asociados con cambios en la otra de manera razonablemente consistente, pero no necesariamente de forma estricta.

Gráfico

El contenido generado por IA puede ser incorrecto.

En este grafico de dispersión se puede observar la relacion de las variables de ventas con el precio del producto

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

## Selección de variables

Para Probar los modelos se seleccionaron las variables del dataframe excluyendo el OutletEstablishmentYear y nuestra variable a Predecir ItemOutletSales para las x.

Para la y se utilizara la variable dependiente ItemOutletSales

Texto

El contenido generado por IA puede ser incorrecto.

Este código define una función train\_and\_evaluate\_model que entrena y evalúa un modelo de regresión, calculando dos métricas de desempeño: **RMSE** (Root Mean Squared Error) y **R2 Score**. El flujo de trabajo incluye los siguientes pasos:

1. **División de datos**: Se divide el conjunto de datos en entrenamiento y validación (80/20).
2. **Pipeline de preprocesamiento y modelo**: Se crea un pipeline que incluye la normalización de datos (usando StandardScaler) y el modelo que se pasa como argumento.
3. **Entrenamiento**: El modelo se entrena con los datos de entrenamiento.
4. **Predicción y evaluación**: Se hacen predicciones sobre los datos de validación y se calculan las métricas RMSE y R2 Score.
5. **Almacenaje de resultados**: Los resultados del modelo se almacenan en un DataFrame (model\_scores) junto con el nombre del modelo.
6. **Impresión de resultados**: Se imprimen las métricas de desempeño para cada modelo evaluado.

Este proceso se repite para diferentes modelos, proporcionando una forma de comparar su desempeño.

Texto

El contenido generado por IA puede ser incorrecto.

## Selección de Modelos

Se prueban diferentes modelos para obtener el rmse y el R2 para hacer comparación y evaluar que modelo ofrece un mejor rendimiento en cuanto a predicción tomando estos valores como referencia.

Se Obtienen los siguiente Valores

|  |  |  |
| --- | --- | --- |
| Model | RMSE | R2 |
| XGBRegressor | 1148.339527 | 0.530812 |
| Random Forest | 1103.357027 | 0.566850 |
| Linear Regression | 1107.729008 | 0.563411 |
| Ridge | 1107.734050 | 0.563407 |
| Lasso | 1107.784775 | 0.563367 |
| SVM | 1648.217447 | 0.033426 |
| Decision Tree | 1531.390792 | 0.165593 |
| Random Forest | 1099.905348 | 0.569556 |
| AdaBoost | 1160.904777 | 0.520488 |
| Decision Tree | 1519.678628 | 0.178307 |
| KNN | 1167.950036 | 0.514651 |
| LightGBM | 1082.279437 | 0.583241 |
| ElasticNet | 1184.748452 | 0.500589 |
| LightGBM | 1082.279437 | 0.583241 |
| Gradient Boosting | 1061.243099 | 0.599285 |
| Bayesian Ridge | 1107.789263 | 0.563363 |

El modelo que mejor rendimiento nos dío fue: Gradient Boosting con un RMSE de : 1061.243099 y un R2 de: 0.599285.

Se utiliza la función train\_test\_split de **scikit-learn** para dividir el conjunto de datos en dos subconjuntos: **entrenamiento** y **validación**. Captura de pantalla de un celular

El contenido generado por IA puede ser incorrecto.

Se crea un modelo de machine learning y lo ajusta a los datos para que aprenda a predecir la variable objetivo.

Texto

El contenido generado por IA puede ser incorrecto.

Texto

El contenido generado por IA puede ser incorrecto.

Utilizamos la función Describe() sobre la variable de valores predichos

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

## Graficas Tras Aplicar los modelos

### Dispersion

Podemos ver los valores predichos versus la línea ideal y se observa la distribución de las predicciones frente a los valores reales.

La relación ideal muestra las predicciones que coinciden con los valores reales. En donde se destacan los valores presentes en la línea lo que indica que el modelo esta haciendo de manera efectiva las relaciones.

Ademas se puede observar que el modelo posee valores predichos por debajo de la línea ideal lo que indica que el modelo tiene una subestimación en las ventas de un valor mas alto.

Gráfico, Gráfico de dispersión

El contenido generado por IA puede ser incorrecto.

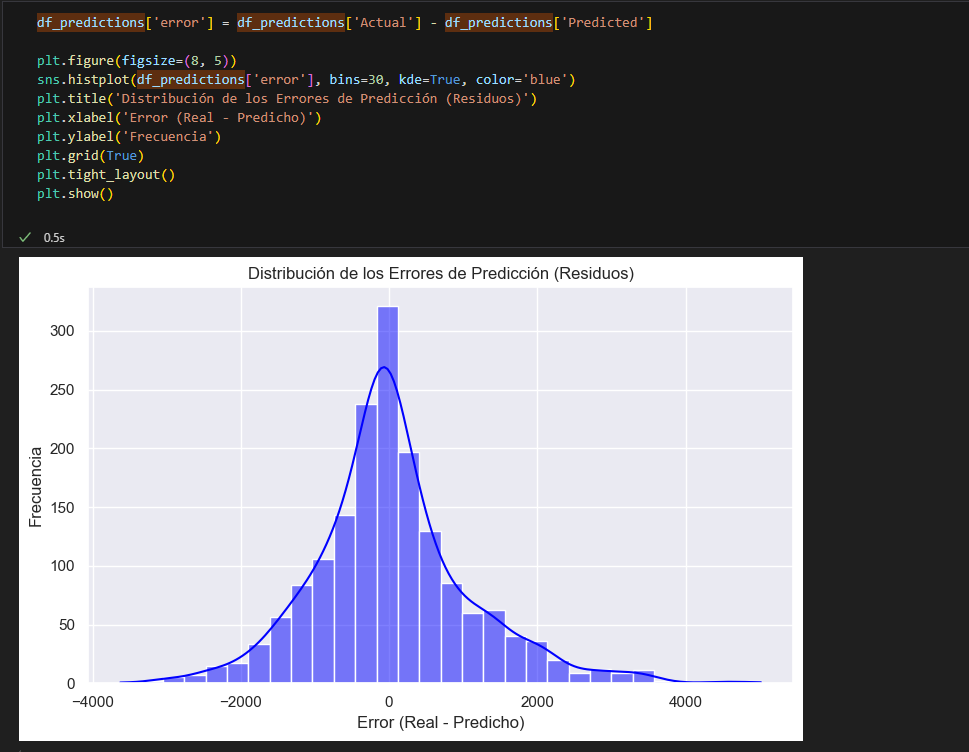
Se hicieron ajustes al modelo “Ajuste de Hiperparámetros” pero el resultado no ofrecía mayor mejora

### Gráfico de líneas Ventas Reales y Predichas por ItemMRP y Ventas

Gráfico, Gráfico de líneas

El contenido generado por IA puede ser incorrecto.

### Grafica de Residuos



Dado que la mayoría de los residuos son cercanos a cero, y la frecuencia disminuye a medida que te alejas de cero, es probable que el modelo esté ajustando bien a la mayoría de los datos. En este caso podemos considerar que el modelo tiene un buen rendimiento general.

### Ventas Reales Vs PRedichas

Gráfico

El contenido generado por IA puede ser incorrecto.

Este gráfico compara las ventas reales con las ventas predichas por el modelo, ordenadas por el valor real de venta.

Se observa poca dispersión entre los valores reales y los predichos.

Se observa una cercanía entre los resultados lo que significa que el modelo predice de manera eficiente

Las discrepancias entre las líneas muestran los errores de predicción.

# Conclusiones

En este análisis, hemos evaluado el desempeño de nuestro modelo predictivo de ventas utilizando diferentes métricas y visualizaciones. Al observar los valores predichos frente a la línea ideal, notamos que en su mayoría las predicciones se alinean estrechamente con los valores reales, lo que indica que el modelo está realizando relaciones efectivas y precisas para la mayoría de los casos. Los valores cercanos a la línea ideal reflejan un buen ajuste y precisión en las predicciones.

Sin embargo, también se identificaron algunas predicciones por debajo de la línea ideal, lo que sugiere que el modelo está subestimando ciertos valores de ventas más altas. A pesar de esto, la distribución de los residuos muestra que la mayoría de los errores son cercanos a cero y que la frecuencia disminuye a medida que nos alejamos de cero. Esto indica que, en general, el modelo ajusta bien la mayoría de los datos y tiene un rendimiento sólido en términos de predicción.

El modelo que mejor desempeño presentó fue **Gradient Boosting**, con un **RMSE** de **1061.24** y un **R2 de 0.59**, lo que sugiere que el modelo tiene una capacidad razonable para explicar la variabilidad de las ventas y realizar predicciones precisas. Al comparar las ventas reales con las predicciones del modelo, se observa una **baja dispersión** entre los resultados, lo que refuerza la idea de que el modelo predice de manera eficiente y precisa.

Las discrepancias observadas entre los valores reales y los predichos reflejan los errores de predicción, que son naturales en cualquier modelo. No obstante, considerando la cercanía entre las predicciones y los valores reales, podemos concluir que el modelo tiene un **buen rendimiento general** y puede ser utilizado con confianza para optimizar las decisiones de ventas y distribución en el futuro.

# Recomendaciones y futuros estudios

## Recomendaciones:

1. **Optimización del Modelo**: Si bien el modelo de **Gradient Boosting** ha mostrado un buen rendimiento, siempre existe espacio para mejoras. Se recomienda explorar técnicas de optimización adicionales, como la **ajuste de hiperparámetros** (por ejemplo, utilizando técnicas como grid search o random search) para mejorar aún más su precisión y reducir el error de predicción.
2. **Evaluación de Otros Modelos**: Aunque **Gradient Boosting** ha sido el modelo con el mejor rendimiento en este estudio, es valioso continuar explorando otros algoritmos de aprendizaje automático, como **XGBoost**, **LightGBM** o **Redes Neuronales**, para comparar su rendimiento en distintas condiciones y conjuntos de datos.
3. **Incorporación de Más Variables Externas**: Aunque se ha considerado un conjunto de características clave (precio, tipo de producto, tamaño de la tienda, etc.), la incorporación de **factores externos** como **tendencias estacionales**, **eventos de ventas especiales** (como promociones o campañas publicitarias), y **análisis de competencia** podría mejorar la precisión de las predicciones, al reflejar un mayor rango de influencias sobre las ventas.
4. **Monitoreo Continuo del Desempeño del Modelo**: A medida que nuevos datos de ventas y cambios en el mercado se generen, es crucial realizar un **seguimiento continuo** del desempeño del modelo y ajustar el modelo predictivo con regularidad para asegurar que las predicciones sigan siendo precisas y relevantes.
5. **Mejoras en la Gestión de Residuos**: Aunque el modelo ha demostrado un buen ajuste general, es importante evaluar y gestionar los residuos de manera más específica. Examinar los casos donde el modelo subestima las ventas (por debajo de la línea ideal) podría proporcionar información útil para mejorar las predicciones en esos escenarios.

## Futuros Estudios:

1. **Estudio de la Influencia de la Visibilidad en las Ventas**: Aunque la visibilidad no fue uno de los factores principales en este modelo, un estudio más profundo sobre cómo la **ubicación en el estante** y la **exposición visual** del producto afectan las ventas podría ofrecer una nueva perspectiva para mejorar las predicciones. En el futuro, se podría incorporar un análisis de visibilidad en la tienda y su impacto directo en el rendimiento de ventas.
2. **Análisis de Series Temporales**: Los datos históricos de ventas podrían ser mejor modelados mediante técnicas de **series temporales**, como **ARIMA** o **LSTM** (Long Short-Term Memory). Este enfoque permitiría capturar patrones y tendencias a lo largo del tiempo, proporcionando predicciones más precisas para los productos con estacionalidad o ciclos de ventas repetitivos.
3. **Análisis de Sentimientos**: Incorporar datos de **sentimiento de los consumidores** obtenidos de redes sociales o reseñas en línea podría enriquecer aún más el modelo. La percepción de los consumidores sobre ciertos productos puede influir significativamente en las ventas, y un análisis de sentimientos podría agregar valor adicional a las predicciones.
4. **Evaluación del Impacto de Promociones y Descuentos**: Un análisis más profundo de cómo las **promociones** y los **descuentos** afectan las ventas puede ser crucial. A medida que se introduzcan estrategias promocionales, se puede ajustar el modelo para tener en cuenta el impacto de estos eventos y mejorar la predicción de las ventas durante estos períodos.
5. **Estudio de Modelos de Deep Learning**: A medida que la capacidad computacional mejora, explorar modelos de **Deep Learning**, como **Redes Neuronales Artificiales** o **Redes Neuronales Convolucionales**, podría mejorar aún más la capacidad predictiva, especialmente cuando se tiene un volumen de datos muy grande y complejo

# Bibliografía

Torres, C. Á. F., Pérez, V. H. C., & Saucedo, C. I. T. (2022). Análisis comparativo de modelos tradicionales y modernos para pronóstico de la demanda: enfoques y características. *RIDE Revista Iberoamericana Para la Investigación y el Desarrollo Educativo*, *12*(24). https://doi.org/10.23913/ride.v12i24.1203

Datarock. (2023, 15 febrero). *Understanding Regression Metrics – Datarock*. <https://datarock.com.au/2023/02/15/understanding-regression-metrics/>

Shatby, S. E. (2022, 18 mayo). *How to Build a Predictive Model in Python?* 365 Data Science. https://365datascience.com/tutorials/python-tutorials/predictive-model-python/

Salazar, N. T. (2024, 5 febrero). Modelos de Forecast en Python: Una Guía para Expertos. *Medium*. https://medium.com/@nicolastoro7/modelos-de-forecast-en-python-una-gu%C3%ADa-para-expertos-ebb8f57dfc83

*GradientBoostingRegressor*. (s. f.). Scikit-learn. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html

*Gradient Boosting con python*. (s. f.). https://cienciadedatos.net/documentos/py09\_gradient\_boosting\_python

Joaquín, A. R. (2023). Cienciadedatos.net. *Zenodo*. https://doi.org/10.5281/zenodo.10006330

Singh, A. K. (2022, 6 enero). Gradient Boosting Regression Example in Python - Anup Kumar Singh - Medium. *Medium*. https://anupkrsingh.medium.com/gradient-boosting-regression-example-in-python-abcd36126614