

1. Resumen Ejecutivo

Este informe presenta el desarrollo de un modelo predictivo para **estimar ventas totales** en un entorno de comercio electrónico. El análisis proporciona **insights clave** sobre los factores más importantes que impactan las ventas, ayudando a optimizar estrategias de precios, inversión en marketing y segmentación de clientes.

2. Objetivo del Proyecto

- Predecir con precisión las **ventas totales** a partir de características del producto y del mercado.
- Identificar los **factores clave** que influyen en las ventas.
- Evaluar el impacto de cambios en **precios** y **gasto de marketing** para maximizar ingresos.

3. Metodología

1. Preprocesamiento de Datos:

- Creación de la variable objetivo: `total_sales = Units_Sold * Price`.
- Codificación de variables categóricas mediante **One-Hot Encoding**.
- Extracción de componentes temporales de la fecha.

2. Selección y Entrenamiento de Modelos:

- Modelos utilizados:
 - **Regresión Lineal**
 - **Random Forest**
 - **XGBoost**
- División de los datos en entrenamiento (80%) y prueba (20%).
- Métricas de evaluación:
 - **MAE** (Error Absoluto Medio)
 - **RMSE** (Error Cuadrático Medio)
 - **R²** (Coeficiente de Determinación)

3. Análisis Complementario:

- Identificación de **importancia de características**.
- Relación entre precio y ventas totales.
- Simulación del impacto del gasto en marketing.

4. Resultados Clave

4.1 Comparación de Modelos

La siguiente figura muestra las métricas de desempeño:

- El modelo **XGBoost** superó a los demás con un menor error y mayor capacidad predictiva:
 - **MAE**: 432.12
 - **RMSE**: 745.34
 - **R²**: 0.94

Decisión Empresarial:

Adoptar el modelo **XGBoost** como herramienta principal para predecir ventas, asegurando alta precisión en las proyecciones.

4.2 Importancia de las Características

- **Precio**: Factor más influyente en las ventas totales (~75%).
- **Unidades Vendidas**: Segunda variable más relevante (~25%).
- **Marketing Spend** y **Descuentos** tienen un impacto menor, pero aún significativo.

Decisión Empresarial:

1. Priorizar estrategias de **ajuste de precios** para maximizar ingresos.
2. Focalizar el gasto en marketing en productos con mayor elasticidad de ventas.

4.3 Predicciones del Mejor Modelo

- Las predicciones del modelo **XGBoost** muestran un ajuste sólido con las ventas reales.
- El modelo es fiable para pronosticar tendencias de ventas y planificar inventarios.

Decisión Empresarial:

Utilizar estas predicciones para optimizar la **gestión de inventario** y anticipar picos de demanda.

4.4 Relación entre Precio y Ventas Totales

- Existe una **relación positiva** entre el precio del producto y las ventas totales.
- Sin embargo, la dispersión sugiere que otros factores, como descuentos y promociones, también juegan un rol importante.

Decisión Empresarial:

1. Realizar **análisis de sensibilidad** de precios para identificar puntos óptimos.
2. Implementar **estrategias de descuentos selectivos** para productos menos elásticos.

5. Insights y Recomendaciones

1. **Optimización de Precios:**
 - El **precio** es el principal motor de las ventas. Implementar análisis de elasticidad ayudará a encontrar el equilibrio entre precio y demanda.
2. **Inversión en Marketing:**
 - Simulaciones indican que un aumento del **20% en el gasto de marketing** genera una mejora moderada en las ventas, pero no es el principal impulsor.
3. **Segmentación de Clientes:**
 - Análisis por segmentos (Premium y Regular) permite ajustar estrategias de marketing y precios según el comportamiento de cada grupo.
4. **Gestión de Inventario:**
 - Utilizar el modelo predictivo para anticipar tendencias de ventas y mejorar la planificación de inventarios.

6. Conclusión

Este proyecto demuestra cómo los modelos de Machine Learning, específicamente **XGBoost**, pueden utilizarse para predecir ventas con alta precisión y generar **insights accionables**. Las recomendaciones extraídas tienen un impacto directo en la toma de decisiones estratégicas, como la optimización de precios, la inversión eficiente en marketing y la planificación de inventarios.

Próximos Pasos:

- Desplegar el modelo en un entorno de producción.
- Integrar análisis de **elasticidad de precios** y segmentación avanzada de clientes.
- Explorar el impacto de factores externos (temporada, competencia).

7. Anexos

- **Dataset:** E-commerce Sales Prediction Dataset (Kaggle).
- **Tecnologías Utilizadas:** Python (pandas, numpy, matplotlib, seaborn, scikit-learn, XGBoost).