

Introducción y Objetivos del Proyecto

- **Objetivo:** Desarrollar un modelo que prediga las ventas en una tienda de retail, utilizando el análisis de datos y técnicas de machine learning.
- **Importancia:** Optimizar el inventario y mejorar la toma de decisiones estratégicas en base a la demanda esperada.

Descripción del Conjunto de Datos

- **Fuente:** Datos históricos de ventas y características de productos.
- **Características:**
 - **Numéricas:** Precio por unidad, cantidad, monto total, edad, entre otros.
 - **Categorías:** ID de cliente, categoría del producto, género, etc.
- **Objetivo de Predicción:** Cantidad de productos que un cliente probablemente comprará.

Principales Análisis y Hallazgos

- **Análisis Exploratorio de Datos (EDA):**
 - Se observaron patrones significativos entre variables como el precio por unidad y el monto total.
 - Identificación de posibles correlaciones entre características de los productos y el comportamiento de compra.
- **Modelos Evaluados:** Regresión Lineal, K-Nearest Neighbors (KNN), Random Forest.
- **Evaluación de Modelos:**
 - **Rendimiento:** KNN y Random Forest tuvieron un rendimiento superior para la clasificación binaria.
 - **AUC y R^2 :** Estos modelos demostraron una mejor capacidad predictiva en comparación con la regresión lineal simple.

Visualizaciones Clave con Explicación

1. **Distribución de Precio por Unidad:** Histograma que muestra la frecuencia de diferentes rangos de precios, lo que ayuda a visualizar la estructura de precios del portafolio de productos.
2. **Relación entre Precio por Unidad y Monto Total:** Gráfico de dispersión para evidenciar la relación directa entre el costo y la cantidad comprada, resaltando tendencias y valores atípicos.
3. **Curva ROC para Modelos de Clasificación:** Muestra la capacidad de los modelos de clasificación binaria para distinguir entre clases, con el AUC como métrica clave.

Conclusiones y Recomendaciones

- **Conclusiones:**
 - El modelo Random Forest es el más adecuado para predecir la demanda, dado su rendimiento superior en métricas de evaluación.
 - Los modelos de machine learning permiten identificar patrones de comportamiento que pueden mejorar la precisión en la predicción de ventas.
- **Recomendaciones:**

- Continuar optimizando el modelo con datos adicionales y ajustes de hiperparámetros.
- Integrar estos resultados en el sistema de inventario para ajustar las compras y optimizar el almacenamiento.