

Primera entrega de proyecto

POR:

Jeffrey Sneider Rengifo Marin CC:1037659300

Elkin David Sanchez Velez CC:1007104359

MATERIA:

Introducción a la inteligencia artificial

PROFESOR:

Raul Ramos Pollan



UNIVERSIDAD DE ANTIOQUIA

FACULTAD DE INGENIERÍA

MEDELLÍN 2023

1. Planteamiento del problema

Predecir las ventas mensuales de un producto basado en características propias de un producto.

2. Dataset

El dataset usado es el siguiente:

<https://www.kaggle.com/competitions/online-sales/overview>. El dataset posee 795 mil filas y 558 columnas

El archivo que contiene los datos del edificio es nombrado como TrainingDataset.csv y contiene la siguiente información:

- **Outcome_M1 al Outcome_M12** – El numero de ventas en los primeros doce meses posteriores al lanzamiento del producto
- **Date_1** – Fecha de lanzamiento de la campaña publicitaria principal y la salida del producto al publico
- **Date_2** – Fecha en que el producto fue anunciado y se lanzó la primera campaña publicitaria.
- **Quan_X** – son variables cuantitativas
- **Cat_X** – son variables categóricas

3. Métricas

La métrica de evaluación principal para el modelo será el error logarítmico medio cuadrático (RMSLE), el cual se calcula mediante la siguiente expresión:

$$\epsilon = \sqrt{\frac{1}{n} \sum_{n=1}^i (\log(p_i H) - \log(a_i + 1))^2}$$

Donde ϵ es el valor del RMSLE, n es el número total de observaciones en el dataset, p_i es la predicción de la variable objetivo y a_i es el valor real de la variable objetivo

Por otra parte, en cuanto a la métrica de negocio, se tiene interés en que las predicciones sean lo suficientemente confiables para saber las ventas mensuales de un producto. Con esta

información se pueden realizar análisis financieros para determinar qué tan rentable puede ser un producto

4. Desempeño

El modelo de predicción será considerado exitoso en producción si logra una reducción del 20% o más en el RMSLE en comparación con las proyecciones históricas de ventas mensuales. Se realizará una evaluación continua y se permitirán ajustes para mejorar el rendimiento. Además, se evaluará el impacto comercial real en las ventas del producto.

5. Bibliografía

- David Chudzicki, Online Product Sales Host. (2012). Online Product Sales. Kaggle. <https://kaggle.com/competitions/online-sales>