

PRIMERA ENTREGA DE PROYECTO

POR:

SANTIAGO LORDUY GIL
MIGUEL ÁNGEL RIVAS ACEVEDO

MATERIA:

INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL

PROFESOR:

RAÚL RAMOS POLLÁN

UNIVERSIDAD DE ANTIOQUIA
FACULTAD DE INGENIERÍA
MEDELLÍN 2022

Store Sales- Time Series Forecasting

1.PLANTEAMIENTO DEL PROBLEMA

Se utilizará pronósticos de serie de tiempo para pronosticar ventas en tiendas con datos de corporación Favorita, para un minorista de comestibles en Ecuador. Se desea construir un modelo que prediga con alta precisión las ventas unitarias de miles de artículos vendidos en diferentes tiendas Favorita.

2. DATASET

El dataset seleccionado proviene de Kaggle y cuenta con siete archivos .csv donde está contenida toda la información compilada necesaria para resolver este reto.

train.csv

- aquí se encuentran los datos de entrenamiento, que comprenden series temporales de características store_nbr, family y onpromotion, así como las ventas objetivo.
- store_nbr identifica la tienda en la que se venden los productos.
- family identifica el tipo de producto vendido.
- sales de las ventas totales de una familia de productos en una tienda en particular en una fecha determinada. Los valores fraccionarios son posibles ya que los productos se pueden vender en unidades fraccionarias (1,5 kg de queso, por ejemplo, en lugar de 1 bolsa de papas fritas).
- onpromotion proporciona el número total de artículos de una familia de productos que se estaban promocionando en una tienda en una fecha determinada.

test.csv

- Los datos de prueba, que tienen las mismas características que los datos de entrenamiento. Pronosticará las ventas objetivo para las fechas en este archivo.
- Las fechas de los datos de prueba corresponden a los 15 días posteriores a la última fecha de los datos de entrenamiento.

sample_submission.csv

- Un archivo de envío de muestra en el formato correcto.

stores.csv

- Almacena metadatos, incluidos ciudad, estado, tipo y clúster.

- cluster es una agrupación de tiendas similares.

oil.csv

Precio diario del petróleo. Incluye valores durante los períodos de tiempo de los datos del tren y de la prueba. (Ecuador es un país dependiente del petróleo y su salud económica es altamente vulnerable a los impactos en los precios del petróleo).

holydays_events.csv

- Vacaciones y Eventos, con metadatos
- NOTA: Preste especial atención a la columna transferida. Un feriado que se transfiere oficialmente cae en ese día calendario, pero el gobierno lo movió a otra fecha. Un día transferido se parece más a un día normal que a un día festivo. Para encontrar el día en que realmente se celebró, busque la fila correspondiente donde el tipo es Transferencia. Por ejemplo, el feriado Independencia de Guayaquil se transfirió del 2012-10-09 al 2012-10-12, lo que significa que se celebró el 2012-10-12. Los días que son de tipo puente son días adicionales que se agregan a un feriado (por ejemplo, para extender el descanso durante un fin de semana largo). Con frecuencia, estos están compuestos por el tipo Día de trabajo, que es un día que normalmente no está programado para trabajar (por ejemplo, el sábado) y que está destinado a pagar el Puente.
- Los feriados adicionales son días que se agregan a un feriado regular del calendario, por ejemplo, como suele suceder alrededor de Navidad (haciendo que la Nochebuena sea un feriado).

3.MÉTRICAS

La métrica de evaluación a utilizar será el error logarítmico medio cuadrático (RMSLE), el cual se calcula mediante la siguiente ecuación:

$$RMSLE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(\hat{y}_i + 1) - \log(y_i + 1))^2}$$

donde n es el número total de instancias, \hat{y}_i es la predicción de la variable objetivo y y_i es el valor real de la variable objetivo.

4. DESEMPEÑO

Si se logra el objetivo se esperaría obtener para este distribuidor minorista una previsión más precisa pretendiendo reducir el desperdicio de alimentos por exceso de existencias y mejorar la satisfacción del cliente.

5.BIBLIOGRAFÍA

Store Sales - Time Series Forecasting | Kaggle. (s. f.). Kaggle. Recuperado 4 de julio de 2022, de <https://www.kaggle.com/competitions/store-sales-time-series-forecasting/overview/description>