

Desafío - Prediciendo las ventas de una tienda

En este desafío tendrás la oportunidad de poner a prueba los conceptos aprendidos durante la sesión. Los ejercicios están diseñados para reforzar practicar lo explicado en clases y poder implementar un caso real.

Lee todo el documento antes de comenzar el desarrollo individual, para asegurarte de tener el máximo de puntaje y enfocar bien los esfuerzos. Asegúrate de seguir las instrucciones específicas en cada ejercicio y de completar los requerimientos adicionales, si los hubiera. ¡A disfrutar aprendiendo!

Tiempo destinado: 4 horas cronológicas



¡Mucho éxito!

Descripción

En esta tarea, se te presenta el desafiante escenario de predecir las ventas mensuales de una tienda utilizando técnicas de análisis de series de tiempo. El objetivo principal es aplicar un enfoque riguroso de análisis y modelado de series de tiempo para realizar predicciones precisas de las ventas futuras.

La tienda ha recopilado datos históricos de ventas mensuales durante varios años, en el archivo **Sales_Data.xlsx**. Tu tarea es desarrollar un modelo que sea capaz de predecir las ventas mensuales futuras basándose en estos datos.

Pasos a Realizar:

1. Descomposición de la Serie de Tiempo: Comienza descomponiendo la serie de tiempo en sus componentes principales: tendencia, estacionalidad y residuos. Utiliza métodos como el suavizamiento exponencial o la descomposición estacional para comprender la estructura subyacente de los datos.
2. Identificación de Parámetros ARIMA: Utiliza las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) para identificar los posibles valores de los parámetros p (orden del término autorregresivo) y q (orden del término de media móvil) en el

modelo ARIMA. Además, realiza el test de Dickey-Fuller para determinar si es necesario aplicar la diferenciación (valor d) para lograr la estacionaridad.

3. Modelado con ARIMA: Aplica el modelo ARIMA utilizando los parámetros identificados en el paso anterior. Ajusta el modelo a los datos históricos de ventas mensuales y realiza predicciones para los próximos meses.
4. Evaluación de Predicciones: Divide los datos históricos en un conjunto de entrenamiento y uno de prueba. Utiliza el conjunto de entrenamiento para ajustar el modelo ARIMA y luego evalúa las predicciones en el conjunto de prueba. Calcula métricas de evaluación como el error cuadrático medio (MSE) o el error absoluto medio (MAE) para medir el rendimiento del modelo.

Requerimientos

1. Importar las librerías importantes para el análisis, importar la data y descomponerla en sus 3 componentes principales.
(2 Puntos)
2. Buscar los parámetros más indicados para p, d y q de la serie de tiempo, donde utilizaremos el test de Dickey-Fuller para encontrar el parámetro d volviendo la serie estacionaria, luego utilizar la función de autocorrelación y la función de autocorrelación parcial para encontrar p y q.
(4 Puntos)
3. Dividir la data en entrenamiento y testeo e instanciar el modelo de ARIMA con los parámetros encontrados y la librería statsmodels y también probar auto_arima de la librería pmdarima, finalmente entrenar los parámetros en la data de entrenamiento.
(2 Puntos)
4. Evaluar los modelos en la data de testeo y generar un gráfico del ajuste, como también calcular la métrica.
(2 Puntos)



¡Mucho éxito!

Consideraciones y recomendaciones

- Debes entregar la solución en un archivo de Jupyter Notebook, con el código y las explicaciones necesarias