

Universidad del Valle de Guatemala
Departamento de Matemática
Licenciatura en Matemática Aplicada

Estudiante: Rudik Roberto Rompich
Correo: rom19857@uvg.edu.gt
Carné: 19857

CC3066 - Data Science I - Catedrático: Luis Furlan
25 de septiembre de 2021

Laboratorio 7 - Series de tiempo

Instrucciones:

Para este ejercicio, trabajarán con 4 conjuntos de datos:

1. daily-total-female-births.csv
2. shampoo.csv
3. monthly-car-sales.csv
4. monthly-mean-temp.csv

Cada uno de estos conjuntos representan series de tiempo, pero muestran diferentes características relacionadas con la tendencias, estacionalidad, período, etc. (ver presentación de clase)

Hemos visto 5 métodos diferentes para predecir con series de tiempo, a decir:

1. Promedio (para usar como base de referencia)
2. SARIMAX
3. Alisamiento exponencial (Winter-Holt)
4. Red Neuronal
5. FB Prophet

En este laboratorio, deben ejecutar los 5 métodos para cada uno de los conjuntos de datos arriba indicados.

Problema 1. *Utilizando la prueba del RMSE (raíz cuadrada de la media de los errores alcuadrado). Generen gráficas para el caso que dió el mejor resultado de cada método en cada conjunto de datos.*

Problema 2. *¿Qué métodos logran captar mejor las tendencias y las variaciones estacionales?*

	Ninguna tendencia	Tendencia	Estacionalidad	Tendencia y estacionalidad
Método/Datos	Nacimiento	Shampoo	Temperatura	Carros
Promedio	6.37	113.15	5.14	3647
SARIMA	6.16 (Ganador)	62.83 (Ganador)	2.27	2600
Holt-Winter	6.74	97	4.58	3635
Red Neuronal	6.7	115	2.20 (Ganador)	2091 (Ganador)

Problema 3. *Generen una tabla comparativa que muestre los RMSE más bajos para cada método con cada conjunto de datos. En la tabla indiquen qué método fue mejor para cada caso de datos. Por ejemplo:*

Problema 4. *Ahora apliquen el FB Prophet con cada uno de los conjuntos de datos. Compare los resultados con los de los métodos anteriores. ¿Hay algún ganador claro, entre todos los métodos?*

Problema 5. *Escriban sus conclusiones sobre lo aprendido en este módulo sobre series de tiempo. ¿Cuál es el mejor procedimiento para resolver un problema de predicción de series de tiempo?*