

Introducción a la Ciencia de Datos (Optativa)
Licenciatura en Informática
2025
Trabajo práctico 4

Problema:

Se desea hacer predicciones a partir de los datos de la serie de tiempo del [índice geomagnético SYM-H](#)

Para este ejercicio se desea:

1. Analizar los datos de la serie de tiempo (máximo, mínimo, existencia de datos faltantes, media, mediana, etc.)
2. Plotear la serie de tiempo para su análisis visual.
3. Desarrollar una función que permita la selección de un intervalo de la serie de tiempo (esto puede servir para tomar una parte de la serie de tiempo para su análisis).
4. Desarrollar una función que permita dividir los datos en datos de entrenamiento y datos de testeo.
5. Desarrollar una función sliding window que permite acomodar la serie de tiempo en cantidad de pasos de entrada y pasos de salida variables (esto es un requisito para el acomodamiento de datos antes de ingresar al modelado).
6. Desarrollar una función walking forward que permita predecir de acuerdo al sliding window definido sobre los datos de testeo. Calcular el error del modelo sobre los pasos que se puedan determinar según el walking forward.
7. Desarrollar una función grid search (teniendo en cuenta al menos 2 hiperparámetros) que permitan generar diferentes modelos y elegir el mejor.