

Proyecto #1: Regresión
Prof. Cristian López Del Alamo
T.A. Ariana Villegas

Actividades

1. Descargar el dataset desde la página de Kaggle: Presupuesto y Ejecución de Gastos 2023. En el enlace pueden encontrar el dataset y la descripción de las columnas en el diccionario de datos.
2. Analizar las características del dataset. La columna **MTO_PIA** será la variable objetivo de este proyecto. Elegir qué otras columnas utilizarán para realizar la predicción, utilicen todas las que consideren pertinentes, excepto las columnas después de **MTO_PIA**. Justifiquen su elección.
3. Implementar un modelo de regresión, ya sea lineal o no lineal, según lo consideren adecuado para este problema. Justifiquen su elección.
4. Realizar experimentos sin regularización, con regularización L1 y con regularización L2. Comenta y analiza los resultados obtenidos de cada experimento.
5. Participar en esta competencia de **Kaggle** con las predicciones realizadas para el subset de datos proporcionado en la competencia.
6. Use esta librería para analizar visualmente los datos: **Seaborn**

NOTA: Cada equipo deberá subir un solo documento resolviendo las actividades propuestas en formato de reporte, de preferencia seguir la siguiente estructura:

1. Introducción: Descripción del proyecto.
2. Dataset: Exploración y análisis del dataset.
3. Metodología: Explicación del modelo, funciones de pérdida y técnicas de regularización.
4. Implementación: Incluir el enlace de Colab o GitHub donde se encuentre la implementación, evitar colocar el código directamente en el informe. [Opcional] También pueden incluir detalles relevantes de la implementación (manejo de errores, paralelización, etc.)
5. Experimentación: Presentar resultados con gráficos y/o tablas, evitar capturas de pantalla de la terminal.
6. Discusión: Interpretar los resultados obtenidos y su relación con la teoría aprendida.
7. Conclusiones: Resumen de resultados, limitaciones y recomendaciones.

Finalmente, anexar el link del GitHub o el Colab para verificar el código fuente. Así mismo, definir una **semilla** para replicar los resultados.