

ANÁLISIS DE DATOS MASIVOS

MINI PROYECTO 1

22 de agosto de 2024

El objetivo de este miniproyecto es identificar los tópicos a partir de un conjunto de documentos usando un método de reducción de dimensionalidad (selecciona uno de los métodos vistos en clases).

- Abigail: IMDB Dataset of 50K Movie Reviews
- Isaac: UCI ML Drug Review dataset
- Bruno: People Wikipedia Data
- Alfonso: Hotel Reviews

El proyecto se evaluará con los siguientes puntos:

- Análisis exploratorio/ pre-procesamiento de los datos: 2 puntos
- Justificación e implementación del algoritmo: 3 puntos
- Visualización de los resultados: 2 puntos
- Análisis y redacción de los resultados: 2 puntos
- Documentación del código: 1 punto

- Fecha límite: **lunes 2 de septiembre**
- Forma de entrega: el miniproyecto se presentará durante la clase (puedes usar Databricks o Google Colab).
- Trabajo individual
- Recuerda enviar la libreta por correo antes de la fecha límite.