

ANÁLISIS DE DATOS MASIVOS

MINI PROYECTO 1

13 de febrero de 2025

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).

- **Ismael:** IMDB Dataset of 50K Movie Reviews
- **Alex:** UCI ML Drug Review dataset
- **Gabriel:** People Wikipedia Data
- **Juan Carlos:** Hotel Reviews
- **Victoria:** Tweet Sentiment Extraction
- **Yordi:** Top 100 Bestsellers reviews
- **Duniel:** Spanish poetry dataset

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: **jueves 20 de febrero, 2025.**
- Forma de entrega: el miniproyecto se presentará durante la clase.
- Trabajo individual
- Recuerda enviar la libreta por correo antes de la fecha límite.