

Unidad 2

Ejercicio 1:

Crea un dataset con las sinopsis y otros datos relevantes de las biografías de la página:

<https://ww3.lectulandia.com/genero/biografia/>

Luego, desarrolla un programa que recomiende biografías basándose en un input del usuario. El programa deberá tener dos enfoques:

1. **Coincidencia exacta:** recomendar biografías basadas métodos frecuentistas
2. **Modelo de embeddings:** utilizar un modelo de embeddings para analizar las sinopsis y ofrecer recomendaciones a partir de un input más general.

Ejemplo: Si el usuario ingresa "quiero leer una biografía de un deportista", el programa debe devolver biografías relacionadas con el deporte.

Además se debe realizar un análisis de cuál indicador de similitud semántica (distancia de coseno, distancia euclidiana, similitud de coseno, entre otras) es mejor para este ejercicio.

Ejercicio 2:

Realizar una nueva mejora en el programa de lunfardo de la unidad anterior utilizando la potencia semántica de los embedding de oraciones.