

TUIA NLP 2024 TRABAJO PRÁCTICO 1

Pautas generales:

- El trabajo deberá ser realizado en grupos de dos (2) integrantes.
- Deberá informar cómo está conformado su equipo de trabajo y cuál es la url del repositorio con el que van a trabajar en el siguiente formulario:
https://docs.google.com/forms/d/e/1FAIpQLSeSD2E0vMn6x5nqedEfdQB2IGP6LboVaY4teQlocXDfg5i0IQ/viewform?usp=pp_url
- Considerar que todos los integrantes deben conocer todos los aspectos del trabajo entregado, ya que será evaluado luego de la entrega de forma oral.
- Las librerías que utilicen deben estar (la primera vez que se cargan) sobre el código donde se las utiliza, no todas juntas al inicio.
- Se debe entregar un informe en el cual se incluya las justificaciones y un vínculo a los archivos que permitan reproducir el proyecto. Recomendamos **gitlab** o **github** para tal fin. Debe realizarse en **colab** y ser entregado en el formato de Jupyter Notebook **.ipynb**, dentro de un repositorio. Guardar una vez ejecutado.
- Para la solución del ejercicio puede utilizar todas las herramientas presentadas en las primeras 3 unidades de la materia.

EJERCICIO:

Mediante un proceso de web scraping genere un dataset de datos sobre libros de la siguiente web: [Lectulandia](https://lectulandia.com/)

El dataset debe tener de un mínimo de 100 libros con los siguientes datos:

- género literario (mínimo 10)¹
- autor
- título
- síntesis de cada libro

Propósito del Trabajo Práctico

El objetivo de este proyecto es desarrollar un programa que interactúe con el usuario para recomendar lecturas. El programa ofrecerá tres opciones principales:

Recomendación Directa: El programa preguntará al usuario "**¿Qué tienes ganas de leer hoy?**" y, mediante la clasificación de la respuesta, propondrá una lista de tres libros acordes a las temáticas mencionadas. Además, se detallará el autor, género y una breve reseña de cada libro que se recomienda descargar.

Elección por Autor: Si el usuario prefiere buscar por autor, el programa ofrecerá una lista de libros del autor especificado. En caso de que haya múltiples resultados, se basará en la similitud de los dos primeros resultados más relevantes, retornando dos títulos (con sus

¹ Nota: tener en cuenta que un libro puede pertenecer a más de un género.

respectivas reseñas) del resultado más cercano y uno del segundo que se recomiendan para descargar. Si existen varios libros del mismo autor aleatorizar.

Elección por Género Literario: Similar a la búsqueda por autor, si el usuario elige buscar por género, el programa ofrecerá una lista de libros del género especificado. Aplicará la misma lógica de similitud para seleccionar y presentar los resultados.