

## **Trabajo Práctico 1 : Clasificador de Recomendaciones Recreativas utilizando Procesamiento de Lenguaje Natural (NLP)**

**Contexto:** Una persona dentro de un mes, se tomará 15 días de vacaciones en la playa. Sin embargo, se estima que durante al menos cuatro de esos días habrá lluvias, lo que podría limitar las actividades al aire libre. Para esos días de mal clima, se propone una solución que facilite la recreación en función del estado de ánimo del día.

**Objetivo:** Desarrollar un programa de Procesamiento de Lenguaje Natural que, según el estado de ánimo del usuario, recomiende entre ver una película, jugar un juego de mesa o leer un libro (o varias opciones para cada caso). Para ello, deberá construir un clasificador que categorice el estado de ánimo del usuario. Luego sugerir el conjunto de recomendaciones basada en una frase de preferencia ingresada por el usuario.

### **Pasos para la construcción del proyecto:**

#### **1. Clasificación del Estado de Ánimo:**

- Utilice los conocimientos aprendidos en la Unidad 3 para desarrollar un clasificador a partir de un prompt con el que determine el estado de ánimo del usuario, el cual deberá categorizarse por ejemplo: "Alegre", "Melancólico" o "Ni fu ni fa".

#### **2. Ingreso de Preferencias:**

- Una vez determinado el estado de ánimo, el usuario deberá ingresar una frase que describa la temática que le gustaría explorar. Por ejemplo: "una historia de amor en la selva".

#### **3. Búsqueda de Opciones:**

- El programa deberá comparar la frase ingresada por el usuario con diversas estructuras de texto provenientes de diferentes fuentes de datos utilizando los métodos aprendidos en clase.
- Disponga de los siguientes datasets:
  - [bgg\\_database.csv](#): Base de datos de juegos de mesa.
  - [IMDB-Movie-Data.csv](#): Base de datos de películas.
  - **Libros del Proyecto Gutenberg**: Realice web scraping para conformar un dataset con información sobre los 1000 libros más populares del Proyecto Gutenberg. El enlace a utilizar es el siguiente: <https://www.gutenberg.org/browse/scores/top1000.php#books-last1>.

#### **4. Recomendaciones:**

- Con base en el estado de ánimo del usuario y la frase ingresada, el programa deberá ofrecer recomendaciones pertinentes entre películas, juegos de mesa o libros. Utilice las herramientas de NLP aprendidas en las tres primeras unidades para lograr resultados coherentes y personalizados.

### **Requerimientos mínimos:**

- Utilice clasificadores para determinar el estado de ánimo (por ejemplo, métodos de clasificación supervisada).
- Aplique técnicas de embeddings y comparación de similitud semántica para encontrar las mejores coincidencias en los datasets.
- Utilice la potencia de reconocimiento de entidades nombradas, (NER, modelo Gliner) con el objetivo de obtener los mejores resultados buscados.

### **Entrega:**

- Se debe entregar un informe donde se documente cómo se implementa cada parte del programa, incluyendo explicaciones de cómo funcionan los algoritmos utilizados.
- Realice pruebas con diferentes ejemplos para mostrar la efectividad del clasificador y del sistema de recomendación.
- El código debe estar bien comentado y seguir buenas prácticas de programación. Debe utilizar entregar el o los colab utilizados en formato de notebook.
- Se debe entregar el dataset generado con el web scraping.

**Nota:** Las fuentes de datos se encuentran en inglés, la aplicación debe comunicarse con el usuario en español.

**Opcional:** Puede presentar una aplicación para el programa desarrollado, utilizando Google Colab con una interfaz sencilla basada en widgets como los proporcionados por la librería ip widgets.