

## Examen Bimestral – Diseño de un Sistema Básico de Recuperación de Información

### Instrucciones:

En este examen, los estudiantes deberán diseñar e implementar un sistema básico de recuperación de información utilizando la base de datos **Rotten Tomatoes movies and critic reviews** disponible en Kaggle. El objetivo es responder consultas relacionadas con la temática de las películas y sus características.

La base de datos está compuesta por dos archivos:

1. **rotten\_tomatoes\_movies.csv**: Contiene información estructurada sobre las películas, incluyendo:
  - Títulos de películas, Géneros, Sinopsis o descripción breve, Puntuaciones de críticos y audiencia, y Fecha de lanzamiento.
2. **rotten\_tomatoes\_critic\_reviews.csv**: Contiene comentarios realizados por críticos sobre las películas, incluyendo:
  - Identificador de la película, Comentarios textuales de críticos, Valoraciones ("fresh" o "rotten").

El examen consiste en construir un sistema de recuperación de información que sea capaz de responder consultas textuales de ejemplo como:

- "Películas sobre viajes espaciales."
- "Películas para ver en familia."

### Requerimientos:

1. Preprocesamiento de Datos
2. Construcción del Sistema
3. Simulación de Consultas: El sistema debe mostrar un listado de las películas más relevantes para cada consulta, ordenadas según su relevancia.
4. Análisis de Resultados:

### Criterios de Evaluación:

- **Preprocesamiento (25%)**: Calidad y completitud en la preparación de los datos.
- **Diseño del Sistema (30%)**: Implementación del sistema y estrategia de búsqueda.
- **Simulación de Consultas (25%)**: Capacidad del sistema para responder adecuadamente a las consultas propuestas.
- **Análisis de Resultados (20%)**: Claridad y profundidad del análisis sobre la relevancia de los resultados y posibles mejoras.

### Formato de Entrega:

- Jupyter Notebook: Debe incluir ejemplos de las consultas realizadas y las respuestas generadas por el sistema.