

# Trabajo Práctico N°2

## Integrantes

- Milla, Andrés - 14934/6
- Mandarinino, Leonel - 15379/5

## Notas

- Este archivo representa el archivo **src.txt** mencionado en el enunciado.
- Tanto las respuestas, como la estructura del modelo se encuentran en el archivo **doc/notas.pdf**.
- En el directorio **data** se encuentran los siguientes archivos:
  - **data/scrap**: Información "cruda" recolectada de cada sitio, sin normalizar, ni unificar.
  - **movies.json**: Unificación de la información obtenida de los sitios, se utilizó una estrategia de mezcla para evitar inconsistencias y unificar la información (Para más información leer el archivo **doc/notas.pdf**)

## Requisitos

- *Python3.8* o versiones posteriores.
- Se recomienda un entorno virtual como *virtualenv* para instalar las dependencias

## Ejecución

1. Crear y activar el entorno virtual:

```
virtualenv -p python3 venv  
source venv/bin/activate
```

2. Instalar las dependencias:

```
pip3 install -r requirements.txt
```

3. Ir al directorio fuente:

```
cd src
```

4. Ejecutar el script:

```
python3 main.py
```

## • Opciones de ejecución

`--verbose` : Permite visualizar todos los logs.

`--offline` : Ejecuta solo el merge, es útil para no scrapear de nuevo.

`--extra` : Ejecuta el scrapeo de películas extras.