



Universidad
Internacional
Menéndez Pelayo

Máster Universitario en Investigación en Inteligencia Artificial

Curso 2020-2021

Sistemas de Recomendación

Recomendación para grupos en Python

23 de abril de 2021

Laura Rodríguez Navas

DNI: 43630508Z

e-mail: rodrigueznava@posgrado.uimp.es

Índice

1. Introducción	3
2. Conjunto de datos	3
Bibliografía	4

1. Introducción

2. Conjunto de datos

Para descargar el conjunto de datos: [MovieLens 100k](#).

Crearemos dos dataframes utilizando la librería pandas [1], a partir de los archivos descargados: *movies.csv* y *ratings.csv*.

```
movies_df = pd.read_csv('dataset/movies.csv')
ratings_df = pd.read_csv('dataset/ratings.csv')
```

Observamos el contenido de *movies_df* a continuación:

movieId	title	genres
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
2	Jumanji (1995)	Adventure Children Fantasy
3	Grumpier Old Men (1995)	Comedy Romance
4	Waiting to Exhale (1995)	Comedy Drama Romance
5	Father of the Bride Part II (1995)	Comedy

Tabla 1: Contenido del dataframe *movies_df*.

Definición 1. *Un DataFrame es una estructura de datos etiquetada bidimensional que acepta diferentes tipos de datos de entrada organizados en columnas. Se puede pensar en un DataFrame como una hoja de cálculo o una tabla SQL.*

Bibliografía

- [1] Jeff Reback, Wes McKinney, jbrockmendel, Joris Van den Bossche, Tom Augspurger, Phillip Cloud, gfyong, Sinhrks, Simon Hawkins, Matthew Roeschke, Adam Klein, Terji Petersen, Jeff Tratner, Chang She, William Ayd, Shahar Naveh, Marc Garcia, Jeremy Schendel, Andy Hayden, Daniel Saxton, Vytutas Jancauskas, Ali McMaster, Pietro Battiston, Skipper Seabold, patrick, Kaiqi Dong, chris b1, h vetinari, Stephan Hoyer, and Marco Gorelli. pandas-dev/pandas: Pandas 1.1.5, December 2020.