Find Your Next Read!

Muhammed Hisham T

21BDA43

About The Project

- This is a streamlit web application that can recommend various kinds of similar books based on a user's interest or based on user-defined books.

Introduction

- A book recommendation system is a type of recommendation system where we have to recommend similar books to the reader based on his interest.
- A book search recommendation system is needed to reduce large data output so that book search will be more effective compared to a book search system with a syntax method. The book search recommendation system uses a user-based collaborative filtering method.
- The main purpose of a recommendation system is that it will suggest items to users easily making their life easier. Today the number of facts on the net increased very hastily and those want few instruments to seek out and access appropriate data. The objective of recommender systems is to provide recommendations based on recorded information on the users' preferences. These systems use information filtering techniques to process information and provide the user with potentially more relevant items.

Dataset

- Dataset is collected from Kaggle.
- It contains three files Name of the books, Ratings of the books, and the Users of the books.
- Attributes in the Name of the Books are ISBN(International Standard Book Number), Book-Title, Book-Author, Year Of Publication, Publisher, Image-URL-S, Image-URL-M, Image-URL-L,(i.e., small, medium, large). It contains 271360 rows × 6 columns.
- Attributes in the Ratings of the Books are User-ID, ISBN (International Standard Book Number), and Book-Ratings. It contains 1149780 rows × 3 columns.
- Attributes in the Users of the books are User-ID, Location, and Age. It contains 278858 rows × 3 columns.

Methodology

- Importing Required Modules : Pandas, Numpy, Matplotlib, Seaborn,SciPy,sklearn,pickle,Streamlit.
- Data Preprocessing: Duplicates were dropped. Lines with too many fields (e.g. a csv line with too many commas) were also dropped.
- Spicy sparse was also used. Sparse data is data that has mostly unused elements (elements that don't carry any information). It can be an array like this one: [1, 0, 2, 0, 0, 3, 0, 0, 0, 0, 0, 0] Sparse Data: is a data set where most of the item values are zero.
- Model evaluation: Used KNN to do classification because it gives the best result when we have to classify a new data element to the nearest group.
- Interpreted the result: Gives the recommended book.

Conclusion

- The streamlit website created, uses KNN algorithm to recommend a book to an individual. Say if a person types the name of a romantic novel, the suggestions that will pop up will also belong to the same category. The purpose of this project is to mainly save time and to reduce the domain for an individual so that he/she will be less confused and can get better options to choose from.