Recommendation System

MINOR PROJECT SYNOPSIS

BACHELOR OF TECHNOLOGY

Information Technology

SUBMITTED BY:

VAIBHAV GUPTA(1921110/1905410)

SWAPLEEN KAUR ARORA (1921105/1905405)

NITISH SHARMA (1921076/1905373)

SUBMITTED TO:

Dr. KS Mann

JAN-JULY 2022

GURU NANAK DEV ENGINEERING COLLEGE

LUDHIANA-141006 INDIA

Contents

1 Introduction

Recommender systems are the systems that are designed to recommend things to the user based on many different factors. These systems predict the most likely product that the users are most likely to purchase and are of interest to. Companies like Netflix, Amazon, etc. use recommender systems to help their users to identify the correct product or movies for them. Technologies that we'll be using in this project will be Python, Pandas, numpy, Machine Learning, Streamlit.

A large number of companies are making use of recommendation systems to increase user interaction and enrich a user's shopping experience. Recommendation systems have several benefits, the most important being customer satisfaction and revenue. Movie Recommendation system is very powerful and important system. But, due to the problems associated with pure collaborative approach, movie recommendation systems also suffers with poor recommendation quality and scalability issues.

Pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.Pandas is mainly used for data analysis and associated manipulation of tabular data in Dataframes.

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

Streamlit is an open-source app framework for Machine Learning and Data Science teams,It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc.

Problem statement

The goal of the project is to recommend a movie or a book to the user. Providing related content out of relevant and irrelevant collection of items to users of online service providers.

2 Objectives

The key objectives of Recommender system are:

1. **Improving the Accuracy of the recommendation system** .
2. **Improve the Quality of the Recommendation system** .
3. **Improving the Scalability**.
4. **Enhancing the user experience** .

3 Scope

The goal of the project is to improve the quality of recommendation system, such as accuracy, quality and scalability of system than the pure approaches. This is done using content based filtering and collaborative filtering. To eradicate the overload of the data, recommendation system is used as information filtering tool in social networking sites .Hence, there is a huge scope of exploration in this field for improving scalability, accuracy and quality of recommendation systems. Recommendation system is very powerful and important system. recommendation quality and scalability issues. .

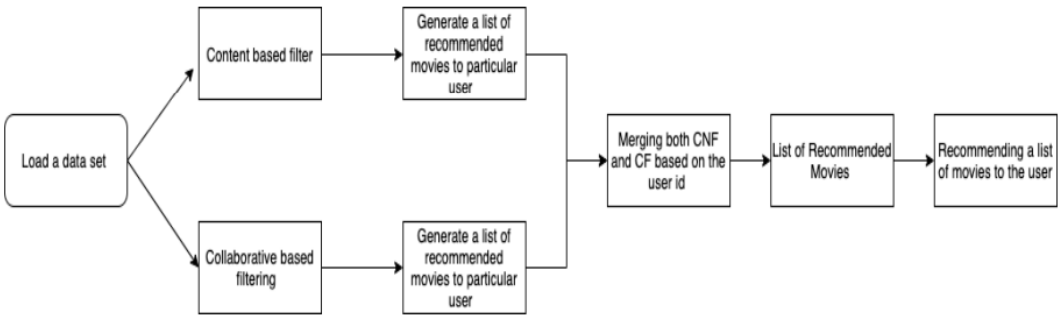
4 Feasibility Study

Our Recommendation system recommends movies and books similar to user's interest. We are using 2 types of reccomendator systems.

1. Content based filtering- Content based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback.

2. Collaborative based filtering- Collaborative filtering is a technique that can filter out items that a user might like on the basis of reactions by similar users. It works by searching a large group of people and finding a smaller set of users with tastes similar to a particular user. It looks at the items they like and combines them to create a ranked list of suggestions.

Our model consists of a very large dataset. So we can perform machine learning algorithm better as Machine Learning algorithms work much better using large dataset. We have used better algorithm for a better accuracy.Also the interface is interactive and easy to use for new users. Our major marketplace are teenagers and adults who wish to read books or movies of similar interest. Our platform is free to use and saves user the time to go through different sites to find similar books or movies.



5 System Requirements Specifications

4.1 Hardware Requirements

- A PC with Windows/Linux OS

- Processor with 1.7-2.4gHz speed

- Minimum of 8gb RAM

- 2gb Graphic card

4.2 Software Specification

- Text Editor (Jupyter notebook)

- Anaconda distribution package (PyCharm Editor))

4.2.1 Anaconda distribution: Anaconda is a free and open-source distribution of the Python programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management system and deployment. Package versions are managed by the package management system conda. The anaconda distribution includes data-science packages suitable for Windows, Linux and MacOS.3

4.2.2 Python libraries: For the computation and analysis we need certain python libraries which are used to perform analytics. Packages such as SKlearn, Numpy, pandas, Matplotlib, Streamlit framework, etc are needed.

SK learn: It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

Pandas: Pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.Pandas is mainly used for data analysis and associated manipulation of tabular data in Dataframes.

Streamlit: Streamlit is an open-source app framework for Machine Learning and Data Science teams,It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc.

6 References

1. Python documentaion

   Python

2. Pandas Tutorial

   Pandas tutorial

3. Numpy Tutorial

   Numpy basics

4. Youtube/Campusx Youtube