Artificial Intelligence (CS411) Assignment 3 – 20/10/2018

Content Based Recommended Systems

Team
Preetham Dasari – 15XJ1A0539
Garaga Sai Chand – 15XJ1A0522
Bhanu Prasad Neerukonda – 15XJ1A0534

Recommended Systems

A **recommender system** or a **recommendation system** (sometimes replacing "system" with a synonym such as platform or engine) is a subclass of information filtering system that seeks to predict the "rating" or "preference" a user would give to an item.

Applications

Recommender systems are utilized in a variety of areas including movies, music, news, books, research articles, search queries, social tags, and products in general. There are also recommender systems for experts, collaborators, jokes, restaurants, garments, financial services, life insurance, romantic partners (online dating), and Twitter pages

Content Based Recommended Systems

In a content-based recommender system, keywords are used to describe the items and a user profile is built to indicate the type of item this user likes. In other words, these algorithms try to recommend items that are similar to those that a user liked in the past (or is examining in the present). In particular, various candidate items are compared with items previously rated by the user and the best-matching items are recommended. This approach has its roots in information retrieval and information filtering research.

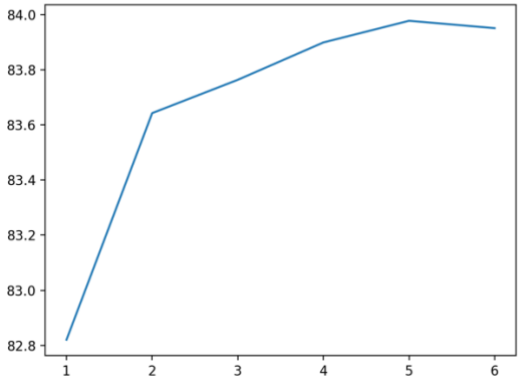
Assignment Approach

A dataset of 100k Movie ratings are given. In the data there are approximately 984 users and each user has rated around 20 movies, And there are 19 different genres. A decision tree is constructed for each user and the tree has been trained by splitting the data into Train (70% of the total data) set and Test (30% of the total data) set. The input to the data set is genre vector and the output is the rating The model is evaluated using two metrics precision and recall for top-k recommended system.

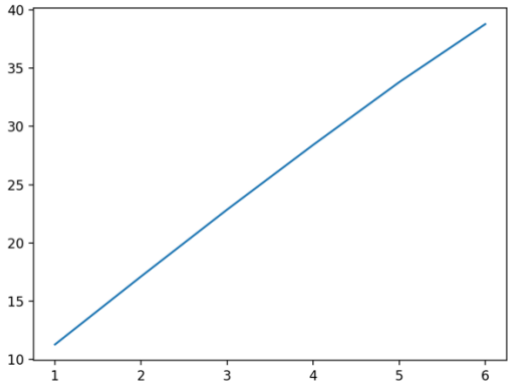
Results:

Taking K=6

Precision:



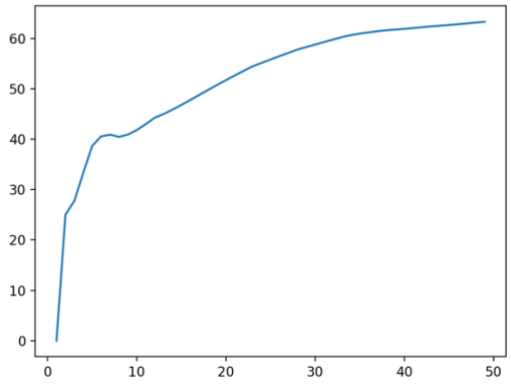
Recall:



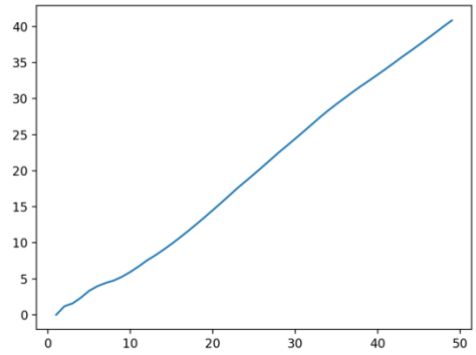
Taking value for k=6 for this data set is accurate as these metrics are applicable to every member in the data set as there are some users who has rated only 20 movies.

Taking K=50

Precision:



Recall:



These metrics are not applicable to every user in the data set as this metrics are calculated only for users who has rated more number of movies.

Implementation Details:

We choose Python programming language and used sklearn library(As the decision tree functions perform faster and produce better accuracy). All the tests were conducted on a 2015 Macbook Pro with 2.7 Ghz Intel i5 processor, 8Gb 1867 Mhz DDR3 Ram, and macOS Mojave operating system.

Web References:

[1] Recall and Precision at k for Recommender Systems**:** https://medium.com/@m_n_malaeb/recall-and-precision-at-k-for-recommender-systems-618483226c54

[2] sklearn library: http://scikit-learn.org/stable/

[3] Errors: https://stackoverflow.com/

[4]Literature: https://en.wikipedia.org/wiki/Recommender_system