IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 3, March 2017

Collaborative filtering in movie Recommendation System based on Rating and Genre

Ashwani Kumar Singh[1], P. Beaulah Soundarabai[2]

Christ University, Bangalore, Karnataka[1,2]

Abstract: Recommender system is one of the recent applications which help to recommend the items or products based on the user needs. Since the user needs varies from person to person we cannot generalize the recommender system. Movie recommender system also has the same issues since individuals have different expectations while watching a movie and recommendation is not possible based on the annotations given by the other users. To overcome this situation an affective recommender framework is proposed in this work. One of the most famous methodology is collaborative filtering and finding the some top similar users and then find the movies those users has viewed and given a good rating which has not been viewed by particular user, recommending movies to user has great benefit, but applying probability for particular movie based on genre is done, which will increase accuracy in recommendation engine.

Keywords: Recommender system, Movie recommender system, Rating and Genre.

I. INTRODUCTION

A recommendation system is an important component in many online services. The commendation system can help the user find interesting content or goods to consume which increases user satisfaction, and for the service provider this leads to the value of increased usage of the service. Movie recommendation systems have attracted significant interest in recent years. Giant players in industry like Netflix and Amazon used movie recommendation system to engage customer finds most relevant content for user and sell. In Netflix purpose of recommendation system is to recommend item to user which user has not seen it and good rating has been given to similar behaviour user. In Amazon it is based on item rating and customer buying pattern or customer search history. Both systems is relying on users rating and users search history, this paper has been taken rating behaviour and genre wise viewed for each particular users and finding similar users based on ratings and genre. In this we have found some interesting pattern which has helped system to find most appropriate similar user to recommend.

II. RELATED WORK

Many researchers have worked on movie recommendation system, recommendation system has been popular area of research from past decade, so many businesses has been built over it, many popular machine learning model and artificial intelligence has been used to study recommendation system for movie like in paper [1] K-Mean clustering and soft max regression has been used for movie recommendation from user rating, as Root-Means-Square Error (RMSE) has been used to evaluation of result, in paper [2] Collaborative filtering has been used to provide recommendation, it is very well known algorithm studied over decade for movie recommendation system, this techniques finds similar users who has seen same movie and their pattern rating is same as current user, finding similar user with same interest help to recommend other movie which similar user has seen and given good rating, which movie is not seem by current user. In paper [4] two versions has been proposed first which users rating to find similar users or user tags available in database or uses movie name to retrieve information about movies from IMDB server.

III. METHODOLOGY

In collaborative filtering, Pearson correlation algorithm has been used to find similar user from movie dataset, in this rating feature has been used to find correlation, algorithm has been used like:

Algorithm for finding similar users using rating	
1	load data into data structure
2	transform data into proper format
3	prepare users list data structure
4	for each user into user list
5	for each user into user list except same user
6	call function pearson correlation (Fig.2.)
7	collect user and correlation return by function put into data structure

Finding correlation between users based on rating will find that similar users whose behaviour of rating is similar but this does not find similarity whose viewing pattern based on genre is same for example in Fig 1: shown about high correlated users but their pattern of watching movies based on genre is different.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 3, March 2017

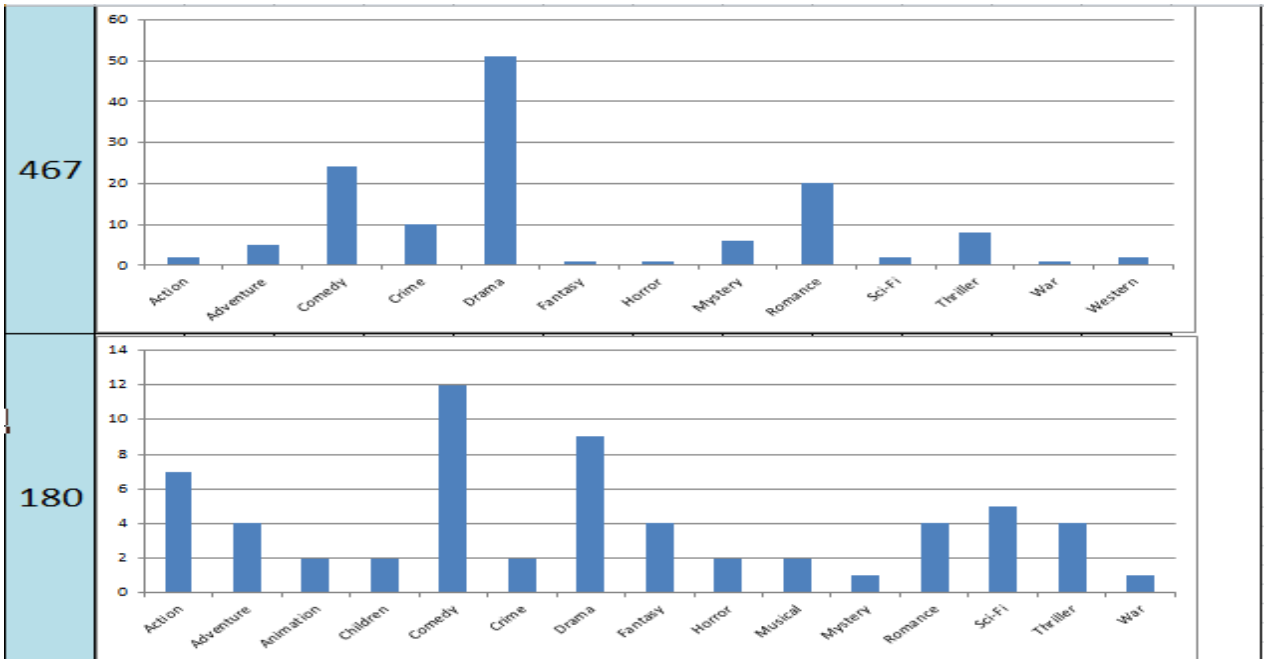

Fig. 1: Difference in Viewing based on Genre

In Fig. 1 we have seen depicted that users 180 and 467 have high correlation score but their viewing pattern is different, e.g. user 467 has watched more movies in adventure than action, where user 180 has watched more action movies than adventure, same for sci-fi and thriller, so if reply on correlation based on rating chances that recommendation will be based on similar user viewing pattern which might be less interesting for user, so correlation between users based on rating and using genre also, and then select K on based on both correlation score to find similar user and in and algorithm for finding similar user based on genre is given below.

Algorithm for finding similar users based on genre	
1	load data into data structure
2	transform data into proper format and aggregate data on genre based using summation
3	prepare users list data structure
4	for each user into user list
5	for each user into user list except same user
6	call function pearson_correlation (Fig.1.)
7	collect user and correlation value return by function put into data structure

Pearson correlation [6] is function is written using formula mentioned in Fig.2.

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

Fig.2: Pearson correlation function.

If we check once again the value of in below table user_id 180 correlation score 1 based on rating pattern with user 467, but if we see the viewing pattern based on genre, correlation score drop down to 0.625176828 which is average score but not very good although.

User_id	similar user	Rating Correlation	Genre Correlation
180	467	1	0.625176828

DOI10.17148/IJARCCE.2017.63107

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 3, March 2017

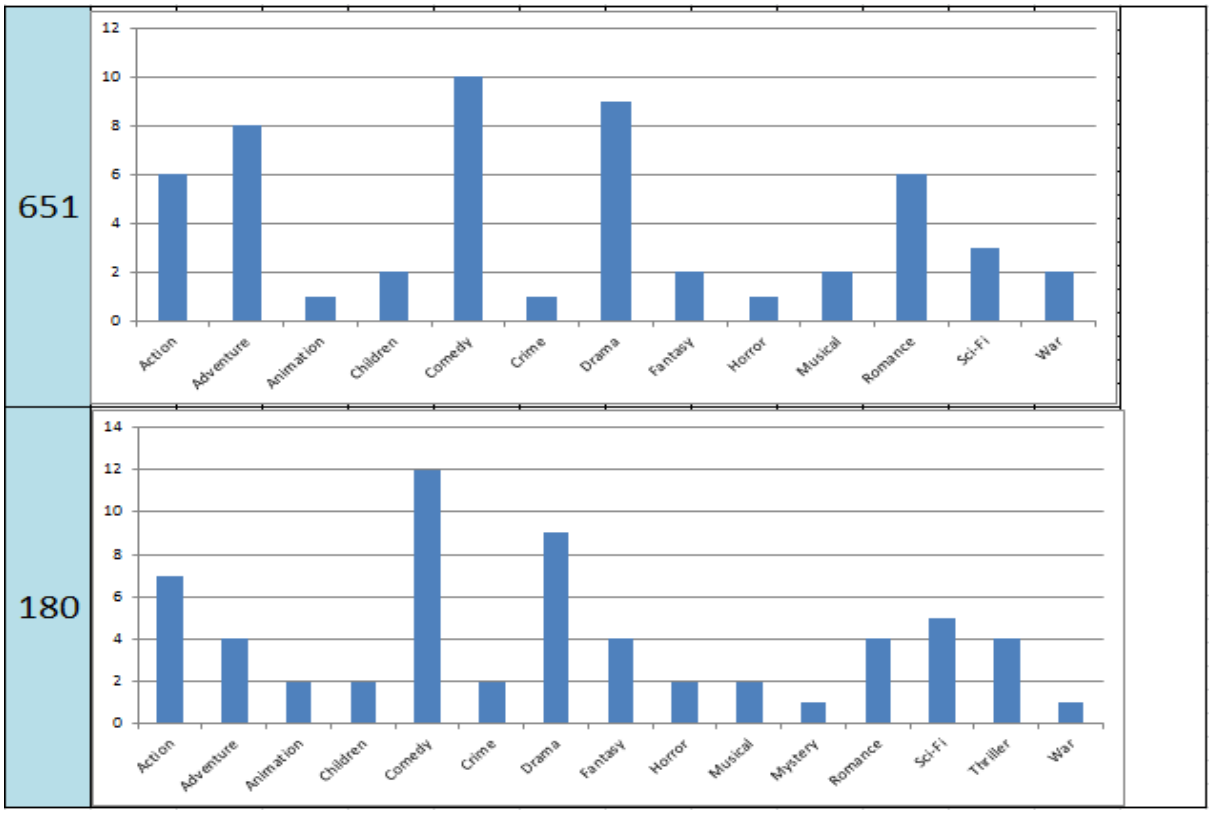

Fig. 2: After applying genre level correlation.

After applying correlation based on rating and genre we founds which are highly correlated on both mentioned in below table.

User_id	similar user	Rating Correlation	Genre Correlation
180	651	1	0.858032849

In this table we can see that user 651 is highly correlated with 180 by genre wise and also rating wise also, in Fig. 2 it shown.

IV. CONCLUSION AND FUTURE WORK

Recommendation that are generated using rating and genre helps in finding best similar users, which will be able to find best movies for users, but finding correlation between Hugh amount of users need very good computing power, but accuracy is good as compared to item-based collaborative filtering when we have dense data, and it is also good for the scenario where data changes frequently and users viewing pattern changes. In future start cast, director and IMDB review of movies can be taken into consideration to obtain more relevant recommendation, complete movie database and user view from social media can be included to find interested of same gender age group.

REFERENCES

[1] Movie Recommendations from User Ratings, Hans Byström, Stanford University, 2013.

[2] Lakshmi Tharun Ponnam, "Movie Recommender System Using Item Based Collaborative FilteringTechnique", GCCT 2015.

[3] N.Sivakumar, N.Balaganesh, K.Muneeswaran, "Feature Selection For Recommendation Of Movies", GCCT 2015.

[4] Ana Stanescu, Swapnil Nagar, Doina Caragea, "A Hybrid Recommender System: User Profiling from Keywords and Ratings", GCCT 2015.

[5] https://www.analyticsvidhya.com/blog/2016/06/quick-guide-build-recommendation-engine-python/

[6] https://en.wikipedia.org/wiki/Pearson_correlation_coefficient.