International Journal of Research in Advent Technology, Special Issue, March 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Movie Recommendation System Using Content and Collaborative Based Filtering

Divyansh Shandilya[1], Harshit Kaushik[2], Gaurav Raj[3]
[*]*Department of Computer Science and Engineering*
Jaypee University, Anoopshahr
Ghaziabad, India
divyanshshandilya24@gmail.com[1], hkaushik28@yahoo.com[2], *gaurav05501@gmail.com*[3]

Abstract— A movie recommendation is important in our social life due to its strength in providing enhanced entertainment. We propose a movie recommendation system based on genre correlations such a system can suggest a set of movies to users based on their interest, or the popularities of the movies. In this paper we propose a movie recommendation system that has the ability to recommend movies to a user. In this project, we attempt to understand the 'Movie Recommendation systems' and compare their performance (Movies Ratings) on the MovieLens datasets which we are implementing on our 'Movie Recommendation System' with the help of Content and Collaborative Based Filtering.

Keywords— Data sets, Content based filtering, Collaborative based filtering, Movies, Recommendation system

1. INTRODUCTION

In today's world as information technology has been developed very much, we can easily obtain the information from the Internet. Since there are large amount of movies available on the Internet, it is difficult to watch all of them efficiently. Thus we should choose which movies we have to watch and which movies we don't . This means that we need to know which movies are entertaining depending on different users and which are not, for better recommendation.

Personalized recommendation system tries to know the characteristics and preferences of the users by collecting and analysing past behavior to know what kind of movies does the user like to watch, what kind of behavior preference the user has, what kind of movies the user like to share and so on, and finally understand that user characteristics and preferences based on the rules of the platform and recommends the movies in which the users are interested.

We revisit the previous research papers with respect to movies recommendation systems. Then, we propose our approach that applies recommendation technique on data sets for movie recommendation systems . We conclude with future works of this research later.

2. RELATED WORK

There are lots of recommendation techniques investigated by many researchers . Recommendation system attempts to recommend information items that are likely to be of interest to the user. There are two kinds of recommendation techniques used in this i.e content-based and collaborative based filtering.

Content-based filtering method uses item-to-item similarity. If a user like B, we recommend A that is similar to B. Collaborative method uses correlation between the users.

Basically, our system uses item-based method. The detailed explanation for our recommending algorithm will be covered later.

3. RECOMMENDATION SYSTEM

Recommendation engines are nothing but an automated form of a "shop counter guy". You ask him for a product.Not only he shows that product, but also the related ones which you could buy. They are well trained in cross selling and up selling. So, thus our recommendation engines.

Recommendation systems are among the most popular applications of data science today.

Recommendation system helps in addressing the information overload problem by recieving the information desired by the user based on his or similar user's preferences and interests.

Some benefits of recommendation systems are :-
-Customer satisfaction
-Discovery
-Revenues
-Personalisation

Example:- Netfilx(2/3 rented movies are from recommendation), google news(38% more click-through are due to recommendation), Amazon (35% sales are from recommendation).

4. TYPES OF FILTERING USED IN RECOMMENDATION SYSTEMS

A. Content Based Filtering

Content filter is a set of programs that help to what content is acceptable for viewing and accessing and what is not.

It recommend items similar to those a user has liked in the past or Recommendations are based on the content of items rather than other user's opinion.

Content filters are majorly used by organisations and home users to block access unsuitable, undesirable, dangerous websites including social networking websites, auction websites, etc.
Due to expansion of new technology and internet , need of content filters are important to face challenges.
Example:- IMDB, Rotten Tomatoes

a) Advantages of Content Based Filtering:

- No need for data on other users.No cold start and sparsity
- Able to recommend users with unique taste
- Able to recommend new and unpopular items
- Increased productivity level
- To secure privacy
- Prevention of liablity issues
- Increased network protection

b) Disadvantages of Content Based Filtering:

- Data should be in structured format
- Unable to use quality judgements from other users
- Accidental access
- Non-Banned sites
- Alternative methods
- Existence of large no. of websites

B. Collaborative Based Filtering

The goal of collaborative filtering is to predict how well a user will like an item that he has not rated given a set of historical preference judgements for a community of users.

Basic Assumptions :-
-Users with similar interests have comman preferences
-sufficiently large number of user preferences are available
Main Approaches:-
-User based
-Item base

a) Advantages of Collaborative Based Filtering:

- CF engines are more versatile, in the sense that they can be applied to any domain, and with sum care could also provide cross-domain recommendations.
- CF engines work best when the user space is large.
- CF engines may be better at helping the user escpe from the "filter bubble" problem, thanks to users drawing bridges across subspaces in the item space.
- Collaborative filtering does not require content analysis and extraction.
- Works well for complex objects(or multimedia) such as music, pictures and movies.

b) Disadvantages Of Collaborative Based Filtering:

- Need to generate user's preferences(time consuming + may not be perfect).
- Collaborative filtering methods recommends item using user preferences, new user will need to rate sufficient number of items to enable the system to capture their preference accurately and thus provide reliable recommendations.

5. A STATISTICAL MODEL FOR COLLABORATIVE MODEL

We propose the following model of collaborative filtering: People watch movies and rate it. For example, movies are action, foreign or classic (can take any). People also generate a genre in their mind e.g., Intellectual , mind twisting or fun. These ratings are unknown, and must be derived as part of the model estimation process.We will eventually use a range of information to derive these ratings, but initially, let us ask how far we can get just using links indicating who liked what.
To see the form of the classes more concretely, rearrange the person x movie table-

Step 1 – Initialize The Movie Ratings

	Whispers	Star Wars
USER 1	9.1	7.5
USER 2	7.2	
USER 3	6.5	
USER 4	6.0	7.5
USER 5	8.2	5.5

The above insight can be made into a formal generative model of collaborative filtering. It is useful to think first of how the data are generated, and then later of how one might best estimate the parameters in the model. The generative model assures a clean, well-specified model.
We assume the following model:
-randomly assign each person to a class k
-randomly assign each movie to a class 1

for each person/movie pair, assign a link with probability Pkl. The model contains three sets of parameters:

-Pk = probability a (random) person is in class
-Pl = probability a (random) movie is in class
-Pkl = probability a person in class k is linked to a movie in class l.

The first two are just the base rates for the classes: what fraction of people are in a given class. The latter, Pkl are the numbers estimated in the above table.
Here's what the ratings matrix looks like:

```
[[ 9.1   7.5   0    8.2   0]
 [  0    0    7.2   0     0 ]
 [  0   6.5  8.5   0     0 ]
 [  0    0    0   6.0   7.5]
 [  0   8.2   0   5.5   0]]
```

Step 2 – Determine Whether a User Rated a Movie
To make our life easier, let's also declare a binary matrix (0's and 1's) to denote whether a user rated a movie.

1 = the user rated the movie.
0 = the user did not rate the movie.

Let's call this matrix 'did_rate'. Note it has the same dimensions as 'ratings',

did_rate = (ratings != 0) * 1;

This above command should give you the following binary matrix:

```
[[1  1  0  1   0]
 [0  0  1  0  0]
 [0  1   1  0  0]
 [0  0  0  1  1 ]
 [0  1   0  1  0 ]]
```

Step 3-Model estimation

Expectation (Assignment)
Find the expected rating for each person and movie.
Maximization (Model estimation)
Find the most likely Pk , Pl , Pkl

To understand this, we will first review the EM on a standard problem, that of estimating Gaussian Mixtures. We will then show how constraints complicate problem formulation, and finally show what constraints are implicit in the collaborative filtering problem.

The model for Gaussian mixtures is simple. Consider K classes, each generated from a normal distribution with mean $\mu k: x \sim gk(\mu k, \sigma 2)$. All have the same variance. The x's are the observed data, and the model parameters and class labels for x are unknown.

The EM iterates between two steps:

E step
estimate class assignments

$$Pik=P(xi \text{ in } k) \sim e-(xi-\mu k)*(xi-\mu k)/2\sigma$$

M step
estimate model parameters

$$\mu k=(\sum iPikxk) \div (\sum iPik)$$

6. CONCLUSION

Under the condition of massive information, the requirements of movie recommendation system from film amateur are increasing. This article designs and implements a complete movie recommendation system prototype based on the Content filtering algorithm, collaborative filtering algorithm and recommendation system technology. We give a detailed design and development process, and test the stability and high efficiency of experiment system through professional test. This paper has reference significance for the development of personalized recommendation technology.

REFERENCES

[1] Al-Shamri, M.Y.H.. "Fuzzy-genetic approach to recommender systems based on a novel hybrid user model", Expert Systems With Applications, 200810.
[2] Hu Jinming. "Application and research of collaborative filtering in e-commerce recommendation system", 2010 3rd International Conference on Computer Science and Information Technology, 07/2010.
[3] Xu, Qingzhen Wu, Jiayong Chen, Qiang. "A novel mobile personalized recommended method based on money flow model for stock exchange.(Research", Mathematical Problems in Engineering"), Annual 2014 Issue.
[4] Yan, Bo, and Guanling Chen. "AppJoy : personalized mobile application discovery", Proceedings of the 9th international conference on Mobile systems applications and services - MobiSys 11 MobiSys 11, 2011.