Multi-Disciplinary Journal
ISSN No- 2581-9879 (Online), 0076-2571 (Print)
www.mahratta.org, editor@mahratta.org

Towards A Hybrid Personalized Movie Recommender System

Gayatri U.Gadikar
gayatrigadikar04@gmail.com

Asst. Prof. Reena Bhati
reena4bhati@gmail.com
Department of Computer Science TMV, Pune-37

Abstract
Recommendation systems represent a powerful method to enabling users for filter through a wide variety of information. Research in this area is moving in the direction of a richer understanding of how recommendation technology may be embedded in specific domains. A recommendation system for movies is important in our social life to provide the enhanced entertainment. Two major recommendation techniques are Collaborative filtering and Content-based filtering. But these filtering techniques are having some limitations thus a hybrid approach is the best solution. The proposed movie recommender system can recommend user-specific movies by using their social networking site like Facebook data. This provides the generalized framework for personalized movie recommendation.

Keywords
Movie Recommendation System, Collaborative filtering, content based filtering, hybrid,Cold-Start problem

I) INTRODUCTION
Recommendation systems work on predictive analytics of Big Data. Recommendation approaches are content based, collaborative filtering and Hybrid recommendation. Recommendation systems provides the facility to understand an indivisual's taste and suggests new items by understanding his like and dislike pattern of different items. 'Recommendation systems' are the services that recommend new items such as movies, brand-wise products, news articles, books, and music to the user. Recommendation Systems based on AI technology have been explored especially around the 1990s when the WWW and internet services grew explosively all around the world. Recommender systems have been proved beneficial to both the product sellers and the users, as they reduce the excess costs required for finding the items that are most likely to sell.

Why recommendation systems?
- Recommendation systems help to accelerate sales
- increase user retention
- enable a better e-commerce experience

Challenges for recommendation systems:
- Processing large datasets using Big-data
- Using minimum computational power by using user's personal data only
- Giving the best possible recommendations to enable sales growth

I) LITERATURE SURVEY

Recommender systems have become an important research area in industry and academia.It provides us various ways to get the best fit to user's interest. Various applications of Recommendation systems are recommending products like books, CDs , selecting various items on Amazon.com, MovieLens…and many more.Moreover, some of the vendors have incorporated recommendation capabilities into their commerce servers [1].Another music streaming service, Pandora uses content-based filtering for recommending songs.[3]. User preferences are dynamic and by understanding those the product has to be recommended to the user. Both content based or collaborative based approach have some limitations like cold start problem.[6]. To avoid this limitation hybrid recommendation is the solution [5]. To build an efficient recommender system a hybrid combination of different methods ofrecommendation is must.[8]. For Hybrid recommender system we can use traditional matrix factorization methods or using deep learning[7].
Our proposed system makes use of hybrid recommendation torecommend a movie as per the interest of end user

I) RESEARCH PROBLEM

Creating a high-quality personalized movie recommendation system for a better e-commerce experience

Challenges

a) Getting complete user data through Facebook API or other social networking sights is became difficult now.
b) Processing large datasets using Big-data
c) Using minimum computational power by using user's personal data only

I) METHODOLOGY: FILTERING ALGORITHM

A. Collaborative based Recommendation systems:

Collaborative filtering Algorithm recommendation system became well known researched techniques of recommendation systems. The interest of user in his past will be similar to his interest in the future. Let see X and Y are having same purchase history and if recently X has purchased some item which is not selected by Y, then we can propose that item to Y. The collaborative filtering technique select items based on a user-based filtering and an item-based filtering [2].

1) User-Based Approach:

In the User-based approach, the user plays an important role. If a certain majority of the customer have the same taste, then they join into one group. Items are recommended to the user as per the evaluation of the items selected by the group of people sharing the common preference with the user. If the item was positively rated by that community, it will be recommended to the user.

2) Item-Based Approach:

Here in the Item-Based Approach, the items play an important role. Recommendations are based on the evaluation of items the system generates recommendations with items in the neighborhood that a user would prefer.[9]

B. Content-Based Recommendation System

Here in Content-based recommender systems deal with profiles of users that are created at the beginning. A profile has information about a user's taste which is based on how the user rates various items. In the process of recommendation, the engine comparesselected and unselected items and looks for similarities [2].

C. Hybrid Based Recommendation System

It's a combination of content-based filtering and collaborative filtering. Hybrid Recommendation helps to solve problem like ColdStart problem [2].

I) PROPOSEDMOVIE RECOMMENDERSYSTEM USER INTERFACE AND LIFE CYCLE

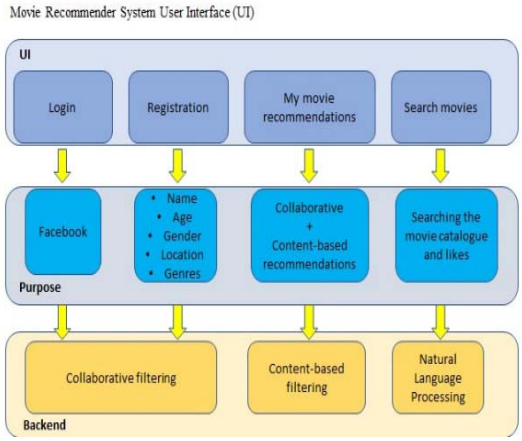

Fig. 1: Interface of Proposed Movie Recommender System

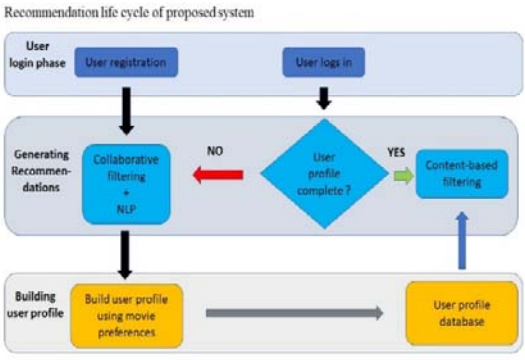

Fig. 2: Proposed Movie Recommender System life cycle

II) RESULTSAND DISCUSSION

Interface of Proposed Movie Recommender System:

In this subsection, the working of the proposed system is described using the architecture diagram in **Fig. 1**.

Proposed system is web based having functionality of **Login** (by registering on the site or by social networking site), **MyMovie recommendation tab** and **MySearch Movie bar**.

a) Login:

User can login eighter by using social networking site like Facebook or by registering on our web site.

By Using social networking site like Facebook:

We can catch the information name, age, gender, language, category and cast of movies can be extracted by using Facebook Graph API [4]

By registering on Personalized movie recommendation web site:

User has to register by providing his details like name, age, gender, profession and optionally select a few genres that the user likes[4].

b) My movie recommendations:

The movie scroller will scroll by providing the movies which are rated most for the age of targeted user, along with the newly launched movies.This will help to resolve cold start problem. Movie scroller facility providesthe user to see the trailer or to read/ write the comments and share his feedback.

c) MySearch movie bar:

The tab will make the provisionto search various movies.

MySearch movie bar will catch differentkeywords used by end user to find different movies.

Proposed Movie Recommender System life cycle:

In this subsection, the life cycle of the proposed movie recommender system is described using the architecture diagram in **Fig. 2**.

*Applying Natural Language Processing***:** Facebook login will help to extract the information of the user's like or dislike related to movies. The posts containing the tagged movie pages along with implicit and explicit feedback are considered for NLP. The weightage to every term (word, phrase, and emoticons) is given which is used to perform NLP and calculate the approximate rating (disliking and liking) for a movie.

Collaborative filtering: This method finds the sub-set of users who have similar testsand preferences to the targeted user and uses this sub-setfor presenting recommendations.

There are two main approaches, namely - 1) User-based and 2) Item Based.

1) *User-Based collaborative filtering*: By using Pearson correlation coefficient and Prediction

In this method, the system predict user's behavior against a certain item using the weighted sum ofdeviations from mean ratings of users that previously rated this item and the user mean rate. First, we calculate the user mean rate using the following formula [4].

$$\overline{v_u} = \frac{\sum_{i \in S_u} v_{ui}}{|S_u|}$$

The weight that we previously mentioned can be calculated using Pearson correlation according to the following formula [4].

$$sim(i,j) = \frac{\sum_{\{u \in U | i \in S_u \& j \in S_u\}} (v_{ui} - \overline{v_u})(v_{uj} - \overline{v_u})}{\sqrt{\sum_{\{u \in U | i \in S_u \& j \in S_u\}} (v_{ui} - \overline{v_u})^2 \sum_{\{u \in U | i \in S_u \& j \in S_u\}} (v_{uj} - \overline{v_u})^2}}$$

Prediction formula is given bellow[4].

$$p_{ai} = \frac{\sum_{\{j \in S_a | j \neq i\}} sim(i,j) \times v_{aj}}{\sum_{\{j \in S_a | j \neq i\}} |sim(i,j)|}$$

For example, the similarity matrix(Table 1) is attached for the reference. Where i and j are the movies and users respectively.

Item → / User ↓	I1	I2	I3	I4	I5
U1	5	8		7	8
U2	10		1		
U3	2	2	10	9	9
U4		2	9	9	10
U5	1	5			1
User a	2		9	10	

Table 1: Matrix for similarity between user and items to enable **user-based collaborative filtering**.

2) Item-based collaborative filtering:

I decided to apply item-based collaborative filtering to display similar items for the user once he or she has selected a particular item using the adjusted cosine formula

$$sim(i,j) = \frac{\sum_{\{u \in U | i \in S_u \& j \in S_u\}} (v_{ui} - \overline{v_u})(v_{uj} - \overline{v_u})}{\sqrt{\sum_{\{u \in U | i \in S_u \& j \in S_u\}} (v_{ui} - \overline{v_u})^2 \sum_{\{u \in U | i \in S_u \& j \in S_u\}} (v_{uj} - \overline{v_u})^2}}$$

Additionally, we can predict how the user will rate this item using the previous similarity:

$$p_{ai} = \frac{\sum_{\{j \in S_a | j \neq i\}} sim(i,j) \times v_{aj}}{\sum_{\{j \in S_a | j \neq i\}} |sim(i,j)|}$$

For example, the similarity matrix (Table 2) is attached for the reference. Where i and j are the movies and users respectively.

Item User	I1	I2	I3	I4	I5
U1	5	8		7	8
U2	10		1		
U3	2	2	10	9	9
U4		2	9	9	10
U5	1	5			1
User	2		9	10	

Table 1: Matrix for similarity between different items to enable **item-based collaborative filtering**.

Collaborative filtering requires all the rated movies tobe considered for ranking. If only a few rated items are available, the recommendation will be done for the small set.

Applying KNN:
We will apply k-nearest neighbors (KNN) algorithm to find the nearest neighbor vector.

Content-based recommendation: Create a user profile based on his like and dislike movies and by the keywords which were used by him to search. User's profile comprising his name, age, movie, genre, and his likes and dislikes. Also, the properties of movies i.e. (similarity and overlapping of the genres, overlapping of the actors and directors, similarity of the movie plots) are considered.Firstly, the movies are filtered by the genres liked by the user followed by filtering these movies by the favorite actors and directors of the user, thereby giving the user recommendations based on what the user has liked before [4].

Hybrid recommendation: Hybrid recommender can apply traditional matrix factorization methods or use deep learning[7]. In our case we have used Matrix factorization method.Aftersuccessful registration the system will generate recommendationsby combining collaborative filtering and NLP(Natural Language Processing). Then the system will build user profile by using movie preferences and create user profile database.Once the user Profile database is created then the system starts recommending content-based filtering. If user logs-in and if his profile is incomplete then the system will recommend by applying collaborative filtering algorithm. As recommendations by combining collaborative filtering and NLP(Natural Language Processing) are hard to process and require high computational power, the personalized movie recommendation will focus on makingstrong user profile to recommend personalized movie as per user's teste.Once the profile is being created,itwill monitor whether the movies have been selected by the user.

Datasets and libraries

- **Datasets:**
  - Ant Movie Catalog is the best movie catalog software. It lets you create and maintain your own personal movie database. You can import list of movies from various sources like CSV/Excel, MS Access Database, DVD Profiler (XML Report), Local Media Files etc.
  - OMDB API

- **Natural Language Processing:**
  - Scikit learn library for feature extraction and processing
  - PyTorch library for Natural Language Understanding
- **Collaborative filtering:**
  - Pandas library for efficient data processing
- **Content-based filtering:**
  - JSON format for storing and retrieving movie catalogue
  - SQL database for storage of JSON file

Benefits of proposed system

- **New user cold start problem**

  Using collaborative filtering at the start and content-based filtering it minimizes cold start problem
- **Data sparsity problem**

  Using NLP + collaborative filtering techniques for personalized recommendations and analyze the user tendencies to create data for content-based filtering
- **Gray sheep**

  We wait for sufficient data to be available for each user to understand the user preferences and only then use content-based approach to overcome this problem

III) CONCLUSION

This paper explains how personalized quality movie recommendations can be constructed by taking minimum inputs from the user and by using the hybrid filtration technique. We proposed a detailed architecture of the recommendation system to enable quality recommendation. We presented the possible datasets and useful libraries for implementation of the recommendation system.
In the future, we can implement this recommendation system as a web application and provide Recommendation As A Service (RAAS)
Film directors / producers can also check the overall mob interested in their film using our web application.

III) REFERENCES

[1] Mohammad Hamidi Esfahani, Farid Khosh Alhan - "New Hybrid Recommendation System Based On C-Means Clustering Method " - 5th conference of information Knowledge. IEEE-2013

[2] Khyati Aggarwal , Yashowardhan Soni "Movie Recommendations using Hybrid Recommendation Systems" International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 4 Issue: 12 206 – 209

[3] Ala Alluhaidan Grand Valley State University "Recommender System Using Collaborative Filtering Algorithm" Grand Valley State University ScholarWorks@GVSU Technical Library School of Computing and Information Systems 2013

[4] A Survey of Various Hybrid based Recommendation Method Jaimeel Shah Computer Engg Parul Institute of Engg & Tech Vadodara, Gujarat, India Lokesh Sahu Assistant Professor Parul Institute of Engg & Tech Vadodara, Gujarat, India

[5] "Content-Based Movie Recommendation System Using Genre Correlation "
SRS Reddy, Sravani Nalluri, Subramanyam Kunisetti, S. Ashok and B. Venkatesh

[6] Creating a Hybrid Content-Collaborative Movie Recommender Using Deep Learning
Adam Lineberry, Sep 10, 2018·11 .Written by Adam Lineberry and Claire Longo

[7] A Review of Movie Recommendation System: Limitations, Survey and Challenges .Mahesh Goyani ∗ and Neha Chaurasiya

[8] Jaimeel Shah , Lokesh Sahu " A Survey of Various Hybrid based Recommendation Method" International Journal of Advanced Research in Computer Science and Software Engineering . Volume 4, Issue 11, N

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

[10] https://grouplens.org/datasets/movielens/100k/

[11] Dongre, M. R., Singh, M. D., & Tilak, G. (2021). An Analytical Study on Influence of Streaming Media on Indian Television Industry. *International Journal of Future Generation Communication and Networking*, *14*(1), 2630-2635.