A Three Way Hybrid Movie Recommendation System

Karan Soni
UG Student,
Department of
Information
Technology,
Shah and Anchor
Kutchhi Engineering
College,
Mumbai-400 088,
India.

Rinky Goyal
UG Student,
Department of
Information Technology,
Shah and Anchor
Kutchhi Engineering
College,
Mumbai-400 088, India.

Bhagyashree Vadera
UG Student,
Department of
Information Technology,
Shah and Anchor
Kutchhi Engineering
College,
Mumbai-400 088, India.

Siddhi More
UG Student,
Department of Information
Technology,
Shah and Anchor Kutchhi
Engineering College,
Mumbai-400 088, India.

ABSTRACT

Recommendation Systems or Engines are found in many applications. These systems or Engines offer the user or service subscriber with a list of suggestions or recommendations that they might choose based on the user's already known preferences. In this paper, the focus is on combining a content-based algorithm, a User-based collaborative filtering algorithm, and review based text mining algorithm in the application of a tailored movie recommendation system. Here movies are recommended based on ratings explicitly provided by the user and according to the ratings and reviews of movies provided by other users as well. Here the propose is to generate polarity ratings to Characteristics of a movie instead of generating a wholesome rating to an available text based review to gain better insights about preferences of users, thus refining Movie recommendation systems further.

General Terms

Data mining, movie recommendation engine, content-based filtering, collaborative filtering, text mining

Keywords

SVM classification, user based collaborative filtering, Content based filtering

1. INTRODUCTION

A recommendation system uses intelligent algorithms, which can provide recommendations to users according to their needs. Recommendation systems can be implemented in any domain from E-commerce to network security in the form of personalized services. They provide benefits to both the user and the service provider, by suggesting items to users on the basis of their known preferences [1]

Every recommender system consists of two entities, one is user and other is an item. A user can be any consumer or a customer of any items or products or services, who has an opinion of that item. The input to recommendation algorithm can be a database of user and items and output will be the recommendations. As in this case, inputs consist of a dataset of customer and database of movies and output denotes the movie recommendations.

Algorithms used in this paper for movie recommendations are content based algorithm [2], collaborative filtering algorithm [2] and review based text mining algorithm [3]. The content-based algorithm consists of user's information such as their interest, favorites, priorities etc. This type of information cannot already be available for a particular user; thus it is explicitly filled by the user. Content-based algorithms provide recommendations or suggestions based on this type of inputs from the user.

Collaborative filtering is a technique that has been used in recommendation systems which can predict and advise items that the user might like based on his or her known preferences. CF-based movie recommendations predict a list of top recommended movies for a given user based on ratings (numerical scores) from many users & the available ratings that are explicitly given by the user by rating a movie or movies [6]

User ratings partially represent user preferences. Thus, it is not easy to accurately identify similar users based on the available ratings. On the other hand, online movie reviews have become a common source for users to share and collect data about movies, but there have been very few studies that accurately integrate text mining techniques with traditional CF approaches to improve the suggestions based on user's preferences. Thus this paper proposes a combined approach of content based algorithm, collaborative filtering and review based algorithm that integrates user text reviews [3] and user numeric ratings [2] in order to model user's preferences better and in turn improve the performance of a movie recommendation systems.

2. RELATED WORK

2.1 Content-based Filtering

The content based filtering approach suggests items to users based on the comparison between the contents of the items or entities and the contents provided by users on their profiles [1]. For example, a movie profile could include various elements such as genre, the participating actors, its box office popularity, and so on. The profiles which are generated will allow the system to link users with matching products. Obviously, content-based strategies require gathering personal information from users that might not be available or easy to collect [2]
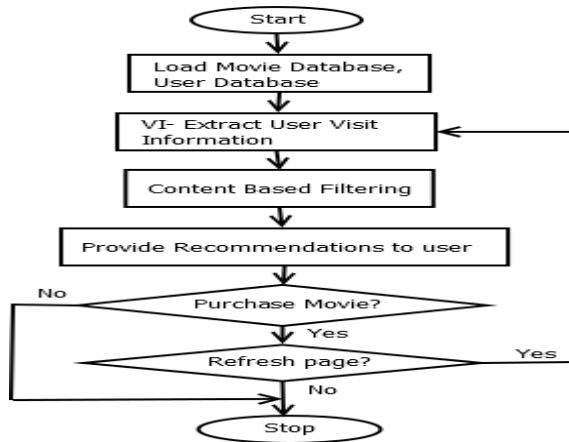
Figure 1: Flowchart of Content-Based Algorithm [2]

2.2 User-based Collaborative Filtering

Collaborative filtering evaluates different relationships between users and interdependencies among products to identify new user-item associations [1]. Collaborative filtering can be exercised either as a User-based Collaborative filtering or an Item Based-Collaborative filtering; in this case items refer to Movies. In a User Based-Collaborative recommendation engine, recommendations are generated on the basis of ratings given by groups of people or user clusters [1]. It locates different users with similar viewing preferences to the current user and creates recommendations for the user based on the ratings provided by those users [2]
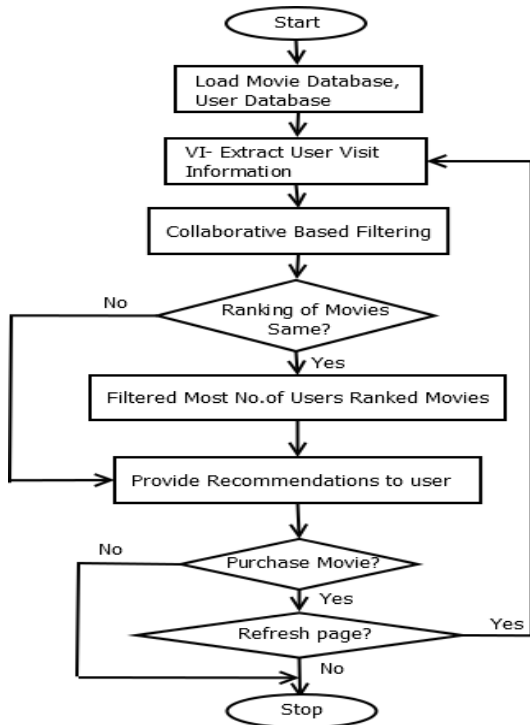

Figure 2: Flowchart of Collaborative-Based Filtering Algorithm [2]

2.3 Review based Filtering for Polarity Score generation

Collaborative filtering based recommendation systems experience a major problem as it suggests movies considering only one criterion i.e. the ratings which are provided by people. However, user ratings alone may not be able to give a Holistic insight of user's actual preferences. With the increasing popularity of Web, users have become more and more comfortable with expressing themselves and providing their opinions on the Internet using text. Such user reviews have potential to provide a system with more detailed and consistent user preference information. In other words, user text reviews can be used for generating a rating on Characteristics of a movie such as Directing, Story/Plot, Cinematography, Editing, Acting, Production Design, Sound and many other features, in conjunction with numerical ratings, to generate a more efficient recommendation process [3]
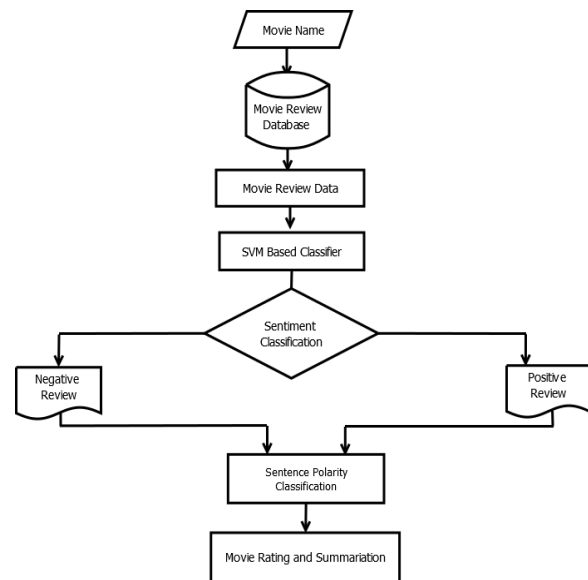

Figure 3: Movie Review Flowchart [3]

3. PROPOSED FRAMEWORK
3.1 Summary of Propositioned Routine

The objective of the proposed work is to generate a movie recommendation system by combining content-based algorithm [2], collaborative filtering [2] and preference scores generated on the basis of text mining [3] performed on the reviews & ratings provided by users. The Intersection stage will compare the movie recommendations from the combination of the Collaborative Filtering-Content based techniques & Polarity score generation using Text mining. This will provide the users with a refined and ranked List of Recommendations.
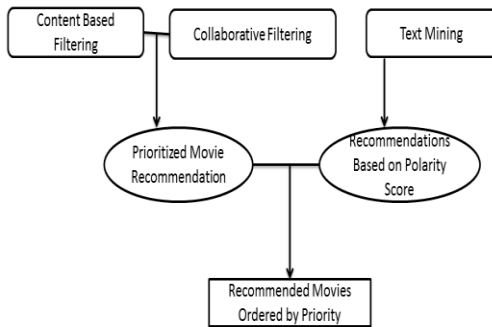
3.2 Ranking of Suggestions

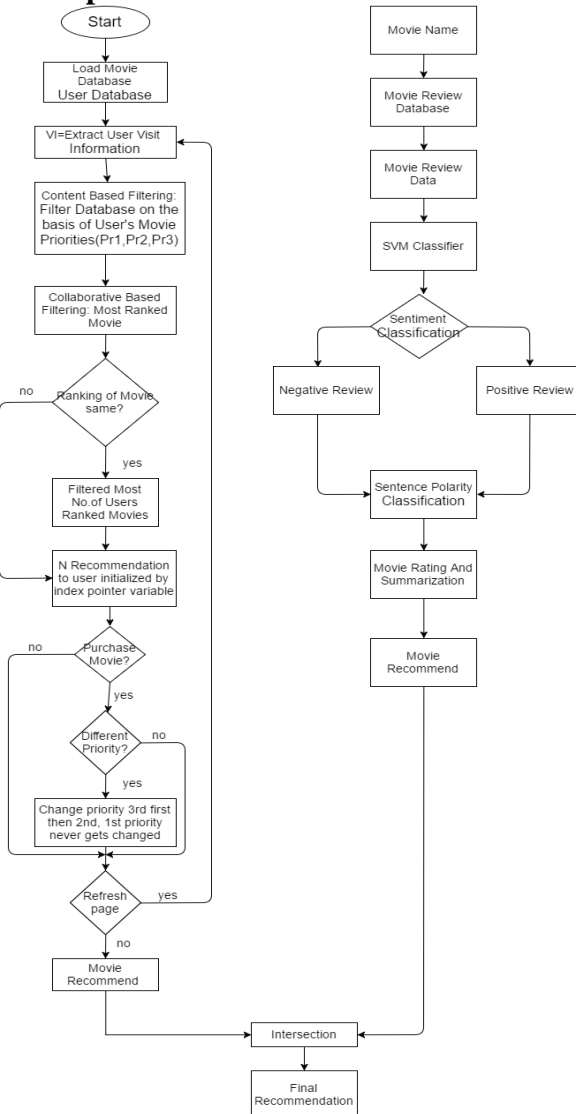**Figure 4: Flowchart of Intersection phase**

3.3 Proposed Flow

Figure 5: Proposed Three Way Hybrid Movie Recommendation System Flowchart

In the proposed work, implementation of content based algorithm and collaborative filtering together and text mining will be implemented independently. The next step is to combine the outputs obtained from both the approach to recommend the movies derived from both the implementations. SVM classification is used in text mining to gain positive and negative polarity scores from the raw unprocessed text data & segregating it on the basis of Characteristics of a movie, For the purpose of the study, System shall be limiting to only 3 characteristics which are vital when it comes to Reviewing a Movie, they are;

    a.   Movie Direction,

    b.   Storyline/Plot &

    c.   The acting of Individuals/Actors in the movie.

3.4 Advantages Inference from Comparative Study

i.   A Hybrid Recommendation engine based on Collaborative filtering, Content & Context based Algorithm [2], can provide, in most cases accurate recommendations but it is limited because this technique in a way disregards if a movie has been received well by the masses it also does not consider text-based reviews which in most cases give a higher perspective about the movie.

ii.   While Text mining Techniques [3] only consider the Rating & Reviews to summarize if a Movie has been received Positively or Negatively by engendering a Polarity Score. These techniques are effective for Summarizing a movie but have not been efficiently integrated on a Movie recommendation system.

iii.   A Recommendation system that Holistically takes into account Entities such as Information about the movie, Rating received by the movie & Reviews gathered would be a much more efficient resource for building a Movie Recommendation system.

4. CONCLUSION

Present Movie recommendation systems lack efficiency because Movies are recommended considering only factors such as Movie rated & reviewed by the User and Users having similar viewing preferences, by using text mining and suggesting movies on the basis of intersection of the three algorithms i.e. User Based Collaborative filtering, Content-based algorithm & text mining due to which the user will not only be suggested movies but this system will also provide the user with more refined recommendations as movies with a low rating score in any of the Movie characteristics generated on the basis of Text Mining will be filtered out during the priority allocation stage of the proposed Three Way Hybrid Movie Recommendation System.

5. FURTHER SCOPE

This paper proposes a system for incorporating Text based movie reviews in traditional recommendation systems, a similar system can be incorporated for Music or Book recommendation systems, it is vital that precise recommendations are provided to the users. Thus,

Accuracy of the recommendation plays a vital role in further refining of the recommendation system/engine. In the future, this recommendation system will provide more immersive and accurate results to the users. The approach in this paper can also be used for other types of recommendation systems/engines.

6. REFERENCES

[1] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. Computer 42, 8 (August 2009), 30-37. DOI=10.1109/MC.2009.263 http://dx.doi.org/10.1109/MC.2009.263

[2] D. Pathak, S. Matharia and C. N. S. Murthy, "ORBIT: Hybrid movie recommendation engine," Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN), 2013 International Conference on, Tirunelveli, 2013, pp. 19-24.doi: 10.1109/ICE-CCN.2013.6528589

[3] C. L. Liu, W. H. Hsaio, C. H. Lee, G. C. Lu and E. Jou, "Movie Rating and Review Summarization in Mobile Environment," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 3, pp. 397-407, May 2012.doi: 10.1109/TSMCC.2011.2136334

[4] S. Vinodhini, V. Rajalakshmi, B. Govindarajulu, "Building Personalized Recommendation System With Big Data and Hadoop MapReduce" ,IJERTV3IS042291 IJERT, April 2014.

[5] Aiming Diao, Minghui Qiu, Chao -Yuan Wu, Alexander J.Smola, Jing Jiang, and ChongWang. 2014.Jointly modeling aspects, ratings and sentiments for movie recommendation (JMARS). In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '14).

[6] Hwang, T., Park, C., Hong, J. et al. Multimed Tools Appl (2016) 75: 12843. doi:10.1007/s11042-016-3526-8

[7] Kyung-Yong Jung, Dong-Hyun Park and Jung-Hyun Lee, "Hybrid Collaborative Filtering and Content-Based Filteringfor Improved Recommender System"in Lecture Notes in Computer Science, 2004, Volume 3036/2004, Springer, pp. 295-302

[8] George Lekakos and Petros Caravelas, "A Hybrid Approach for Movie Recommendation" in Multimedia Tools and Applications, Volume 36, Numbers 1-2, January 2008, Springer, pp. 55-70

[9] Robin Burke, "Hybrid Recommender Systems: Survey and Experiments" in User Modeling and User-Adapted Interaction, Volume 12 Issue 4, November 2002, pp. 331-370

[10] Perm Melville, Raymond J. Mooney, Ramadass Nagarajan, "Content-Boosted Collaborative Filtering for Improved Recommendations" in the Proceedings of the 2002 American Association for Artificial Intelligence, pp. 187 – 192

[11] Y. Wang, Y. Liu and X. Yu, "Collaborative Filtering with Aspect-Based Opinion Mining: A Tensor Factorization Approach," 2012 IEEE 12th International Conference on Data Mining, Brussels, 2012, pp. 1152-1157. doi: 10.1109/ICDM.2012.76

[12] Zhenxue Zhang. 2013. Urcf: An Approach to Integrating User Reviews into Memory-Based Collaborative Filtering. Ph.D. Dissertation. University of Maryland at Baltimore County, Catonsville, MD, USA. Advisor(s) Dongsong Zhang. AAI356338.

ACM, New York, NY, USA, 193-202. DOI: http://dx.doi.org/10.1145/2623330.2623758