



Movie Recommendation System

Ashish Dhyani | Narinder Kaur

Department Information Technology, Maharaja Agrasen Institute of Technology, New Delhi, India.

To Cite this Article

Ashish Dhyani | Narinder Kaur. Movie Recommendation System. *International Journal for Modern Trends in Science and Technology* 2021, 7 pp. 176-181. <https://doi.org/10.46501/IJMTST0712034>

Article Info

Received: 12 November 2021; Accepted: 06 December 2021; Published: 13 December 2021

ABSTRACT

With the rise of Youtube, Amazon, Netflix, and many other such web services in the last few decades, recommender systems have taken on an ever increasing role in our lives. Whether it's e-commerce or online advertising, recommender systems are an integral part of our daily online lives. In general, recommender systems are algorithms that recommend relevant items to users. In this paper I will give you a brief description of Content based, collaborative filtering and weighted hybrid technique. As a result of the proposed system, I got recommended movies based on user behavior. The datasets is retrieved from kaggle and Movie lens.

KEYWORDS: Content-Based Filtering, Collaborative Filtering, Weighted Hybrid Technique.

INTRODUCTION

There is no doubt that the world is growing faster than ever and that each and every sector is advancing. A good example of this is online business, as the market is growing exponentially, and owners of websites are offering services to attract users to their sites. Recommended Engines is one of the facilities given to users. It is easier for users to find things through recommendation systems rather than having to browse through a large number of pages to find the thing they are looking for. The data is taken from the previous behavior of the user and from other users of the same type. In the following step, different algorithms are used to forecast the items they may be interested in.

The three main approaches to building a Movie recommendation system are content-based recommendation systems, in which the filtering is based on the movie overview or genre. According to their genre or overview, the system finds similar movie. By looking at the user's previous history, similar movie are found that he or she may like.

The other method is the collaborative filtering recommendation system in which we group users of the same type and recommend movie based on their history and the last method is hybrid recommendation system which offer a broader range of products to customers, hybrid recommendation systems use both content-based and collaborative filtering simultaneously. This recommendation system provides more accurate recommendations than other recommender systems.

PROPOSED SYSTEM

A movie recommendation system is proposed in this paper the objective of recommender systems is to provide recommendations based on recorded information on the users' preferences. [8] These systems use information filtering techniques which are weighted hybrid technique, Content based filtering and Collaborative based filtering to process information and provide the user with potentially more relevant movies.

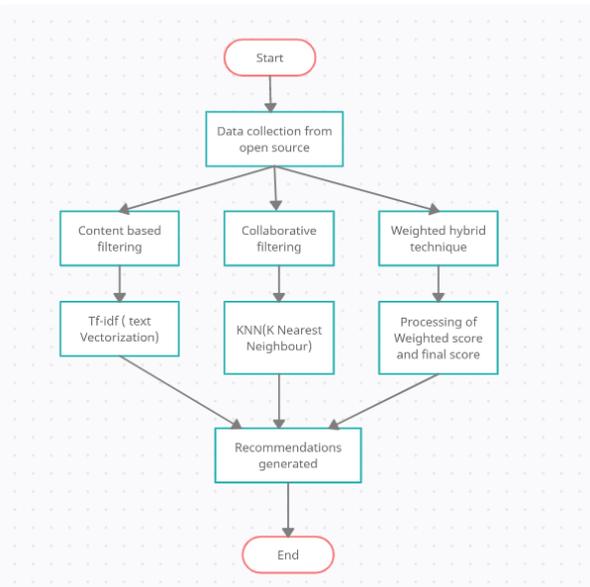


Fig. 1: Proposed Movie Recommendation System.

A. Datasets used

There are two different databases to be used to build the movie recommender system the first one is tmdb 5000 movie dataset which is retrieved from kaggle and other is movielens dataset.

IMPLEMENTATION OF CONTENT BASED FILTERING

Content based filtering is based on the description or some data provided for that product [5]. Comparing products based on their context or description allows the system to determine their similarity. In order to find similar products, the user's past history is taken into account.

The system uses the content of the movie (plot description which is in "overview" feature of dataset) to determine its similarity to other movies. The most likely similar movies will then be recommended. I calculated a pairwise similarity score based on plot descriptions of all movies and the system recommends movies based on this similarity score the movie with the highest similarity score will be recommended first.

In order to calculate a pairwise similarity score, I first created a vector-matrix using TF-IDF and then used cosine similarity to calculate a score that denotes the similarities between two movies.

Our dataset provides a description of the plot in the "overview" feature.

IMPLEMENTATION OF COLLABORATIVE FILTERING

Content-based filtering (CBF) is a popular and widely researched method of recommendation [1]. It is a technique where individual user profile are taken into account. It enhances the user's interest and predicts whether the user would be interested in any particular movie.

This recommendation system uses ratings and similarity between users to recommend movie[4].

The KNN(k nearest neighbor) algorithm is used in this technique. The KNN algorithm employs the same principle. Its goal is to find all the nearest neighbors around a new, unknown data point in order to determine what class it belongs to[6]. The dataset which is used in this recommendation system is retrieved by movielens.

One of the limitations is that for new users there is no appropriate recommendation this is known as cold start problem [4].

IMPLEMENTATION OF WEIGHTED HYBRID RECOMMENDATION SYSTEM

As several recommender approaches have developed, hybrid recommender systems [3] have also emerged in which multiple algorithms are combined. The hybrid recommendation system is a special type of recommendation system that combines content and collaborative filtering. Using collaborative and content-based filtering together may help to overcome the shortcoming of using them separately, while also being more effective in some cases.

The system calculates weighted rating on the basis of the formula mentioned below:-

$$W = \frac{R \cdot v + C \cdot m}{v + m}$$

Where

- W = Weighted rating
- V = number of votes for the movie = (votes)
- M = minimum votes required to be listed in the Top 250 (currently 25,000)
- C = the mean vote across the whole report

Next the minmax scaler is used the main task of this scaler is to normalize each column independently and

making each and every value in the data set between 0-1. It will remove the difference in large and small values and different data unit values.

Then the system recommends movie based on the final score.

RESULT

A. Content-based recommendation system.

```
In [27]: content('Avatar')
Out[27]: 1341      Obitaemyy Ostrov
         634       The Matrix
         3604      Apollo 18
         2130      The American
         775       Supernova
         529      Tears of the Sun
         151       Beowulf
         311      The Adventures of Pluto Nash
         847       Semi-Pro
         942      The Book of Life
         Name: original_title, dtype: object
```

B. Collaborative filtering based recommendation system.

```
for i in range(0,len(distances.flatten())):
    if i == 0:
        print('Recommendations for {0}:\n'.format(features.index[query_index]))
    else:
        print('{0}: {1}, with distance of {2}.'.format(i, features.index[indices.f
```

Recommendations for Toy Story (1995):

- 1: Toy Story 2 (1999), with distance of 0.4273987396802844:
- 2: Jurassic Park (1993), with distance of 0.4343631959138433:
- 3: Independence Day (a.k.a. ID4) (1996), with distance of 0.43573830647233425:
- 4: Star Wars: Episode IV - A New Hope (1977), with distance of 0.4426118294200634:
- 5: Forrest Gump (1994), with distance of 0.4529040920598262:

C. Weighted hybrid recommendation system

```
In [43]: final_system.head(20)
```

Out[43]:

	id	title	normalized_score	normalized_popularity	final_score
22	57158	The Hobbit: The Desolation of Smaug	0.932606	0.827162	0.879884
1897	10167	Half Past Dead	0.645794	1.000000	0.822897
60	17979	A Christmas Carol	0.898847	0.549462	0.724155
294	116711	Epic	0.814527	0.587690	0.701108
516	2976	Hairspray	0.780770	0.495989	0.638379
5	559	Spider-Man 3	0.965987	0.219887	0.592937
8	767	Harry Potter and the Half-Blood Prince	0.949496	0.213941	0.581718
0	19995	Avatar	1.000000	0.158179	0.578090
1	285	Pirates of the Caribbean: At World's End	0.983010	0.164074	0.573542
235	2395	Asterix at the Olympic Games	0.831313	0.310620	0.570966
1666	6016	The Good Thief	0.662696	0.478206	0.570451
2	206647	Spectre	0.966308	0.167611	0.566960
21	20862	Robin Hood	0.932650	0.191397	0.562024
94	118340	Guardians of the Galaxy	0.881863	0.232685	0.557274
9	209112	Batman v Superman: Dawn of Justice	0.949394	0.157762	0.553578

CONCLUSION

So we know that we are in 21st century and the availability of internet allowed us to reach every market and to access everything in a single click . Hence we have a vast number of choices and to choose from these choices is a difficult task. Hence Recommendation System is a very important aspect of current world.

Whether to choose any product or to choose correct book to read or watch any movie that you like everywhere we have recommender system and different systems serve different purposes.

- Hybrid Type Recommendation System are best among all the recommendation system as they take the functionality of both the collaborative and content based recommendation system.

- Collaborative Filtering Based recommendation system have some problem associated with them – 1) They require more computational power. 2) For new users there is no appropriate recommendation this is known as cold start problem.

- Content Based recommendation System doesn't need any data about other users, since the recommendations are specific to the user. This makes it easier to scale to a large number of users.

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere gratitude to My mentor Narinder Kaur for the continuous support of during the project and research paper, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped in all the time of research and writing of this paper. I could not have imagined having a better advisor and mentor for our research-based project.

REFERENCES

- [1] Manoj Kumar, D.KYadav, Ankur Singh, Vijay Kr. Gupta, " A Movie Recommender System: MOVREC" International Journal of Computer Applications (0975 – 8887) Volume 124 – No.3, August 2015.
- [2] Luis M. de Campos, Juan M. Fernández-Luna *, Juan F. Huete, Miguel A. Rueda-Morales; "Combining content-based and collaborative recommendations: A hybrid approach based on Bayesian networks", International Journal of Approximate Reasoning, revised 2010.
- [3] Harpreet Kaur Virk, Er. Maninder Singh, " Analysis and Design of Hybrid Online Movie Recommender System "International Journal of Innovations in Engineering and Technology (IJJET)Volume 5 Issue 2, April 2015.

- [4] Sharma, Poonam and Yadav, Lokesh Yadav, "Movie Recommendation System Using Item Based Collaborative Filtering (2020)". International Journal of Innovative Research in Computer Science & Technology (IJIRCST), ISSN: 2347-5552, Volume-8, Issue-4, July 2020, Available at SSRN:<https://ssrn.com/abstract=3672412>
- [5] Simon Philip1, P.B. Shola; "Application of Content-Based Approach in Research Paper Recommendation System for a Digital Library", in October 2014 International Journal of Advanced Computer Science and Applications.
- [6] Meenu Gupta, Aditya Thakkar ; "Movie Recommender System Using Collaborative Filtering" Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020) IEEE Xplore Part Number: CFP20V66-ART; ISBN: 978-1-7281-4108-4
- [7] Hongli Lin, Xuedong Yang, Weisheng Wang A Content-Boosted Collaborative Filtering Algorithm for Personalized Training in Interpretation of Radiological Imaging J Digit Imaging. 2014 Aug; 27(4): 449-456.
- [8] <https://en.wikipedia.org/wiki/Recommendersystem>.

