International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 11 | Nov 2022 www.irjet.net p-ISSN: 2395-0072

A Review Study OF Movie Recommendation Using Machine Learning

Devani Radhika[1], Saket Swarndeep[2]

[1]Student of Masters of Engineering, Ahmedabad, Dept. of Computer Engineering, L.J. University, Gujarat, India
[2]Assistant Professor, Ahmedabad, Dept. of Computer Engineering & Technology, Gujarat, India

---

Abstract - *Recommendation systems forecast consumer preferences for yet-to-be-viewed goods like movies, songs, or books. The goal of movie recommendation systems is to anticipate users' interests and suggest products that are probably interesting to them. To assist users in finding the most pertinent films, movie recommendation systems employ a variety of filtration techniques and algorithms .From a commercial standpoint, the more relevant information or films a user discovers on a certain platform, the better their engagement and, thus, the higher revenue. Content-based filtering, collaborative filtering systems, and hybrid filtering are the most widely utilized subcategories of machine learning algorithms for movie suggestions. We can create a model that can propose movies based on historical data using machine learning. In this essay, we will examine various machine learning techniques that can be applied to movie recommendation.*

Key Words: Movie recommendation system, collaborative filtering, content-based filtering, hybrid system, k-means, k-nearest neighbour (KNN), clustering algorithm

1 INTRODUCTION

The concept of machine learning had emerged as early as the late 19th century [1]. One of the main reasons we need a recommender system is that in modern times, with the help of the internet, we have far too many options for everything! Because of the revolution in the entertainment industry, the source of entertainment has grown rapidly in today's world. A recommendation system or recommendation engine is a model used for information filtering where it tries to predict the preferences of user and provide suggests based on these preferences [2]. The recommendation system is mostly used in digital entertainment platforms like Netflix, Prime Video, and IMDB, as well as e-commerce platforms like Amazon, Flipkart, and eBay.

Movie suggestions for users depend on Web-based portals. Movies can be easily differentiated through their genres, such as comedy, thriller, animation, and action. Another possible way to categorize the movies based on its metadata, such as release year, language, director, or cast. Most online video-streaming services [3].The main motto of the current recommendation system is formulated the basic fact that "share" and "learn". The process of sharing and witnessing each other's opinions is considered as the heart of the recommending system [4]. Because of its ability to provide enhanced entertainment, a movie recommendation is becoming increasingly popular as a part of our social lives.

1.1 RECOMMENDAION SYSTEM

A Recommendation system is a type of information filtering system which is used to predict the "rating" or "preference" a user would give to an item [5]. The ability to predict user preferences and needs when analyzing user behavior or other user behavior to produce a personalized recommender is a very helpful feature in a recommendation system.

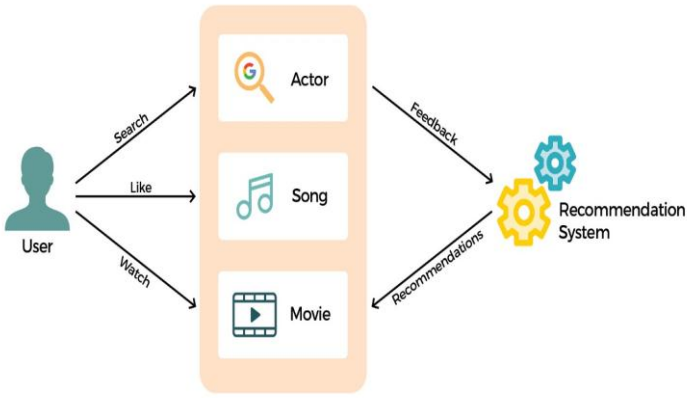

Fig - 1: Recommendation system [6]

Recommendation system are typically classified into the following filtering technique are:

- Collaborative filtering
- Content-based filtering
- Hybrid filtering.

1.1.1 Collaborative Filtering

The majority of recommendation systems use collaborative filtering to identify similar user patterns or information; this method can exclude items that users like based on the ratings or reactions of similar users.

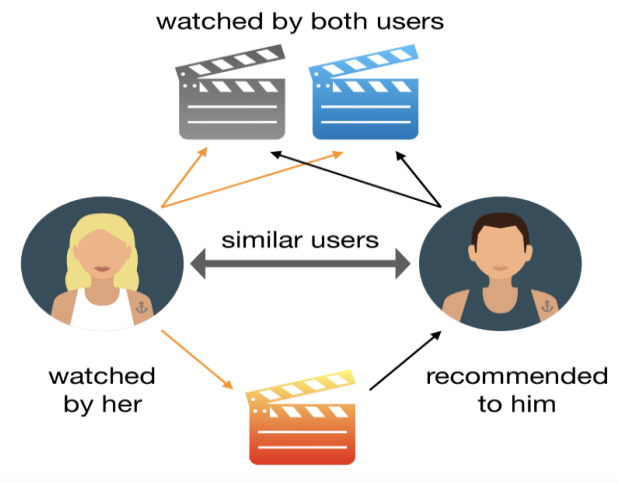

Fig – 2 Collaborative Filtering technique [7]

Cooperative filtering comes in two flavors:

A. **User-based**: measures the degree to which target users and other users are comparable.

B. **Item-Based**: This method assesses how similar two items are that target user's rate or interact with.

1.1.2 Content-Based Filtering

According to the user's past behavior or explicit feedback, content-based filtering uses item features to suggest additional items that are similar to what they already like.

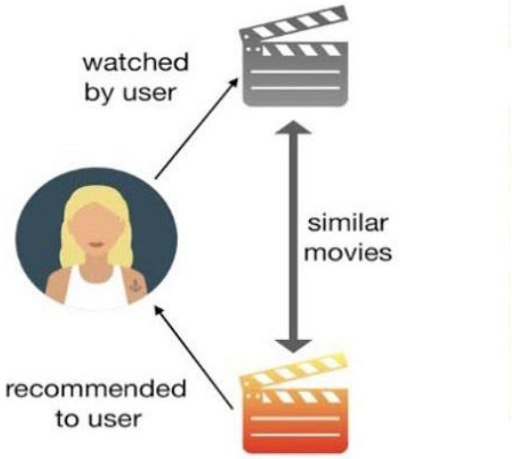

Fig – 3 Content-based Filtering technique [8]

The similarity between item vectors can be computed by three methods [9]:

A. Cosine similarity
B. Euclidian distance
C. Pearson's correlation

1.1.3 Hybrid Filtering

Which can be thought of as a hybrid of the content- based and collaborative filtering methods. It avoids the flaws of all recommender techniques.

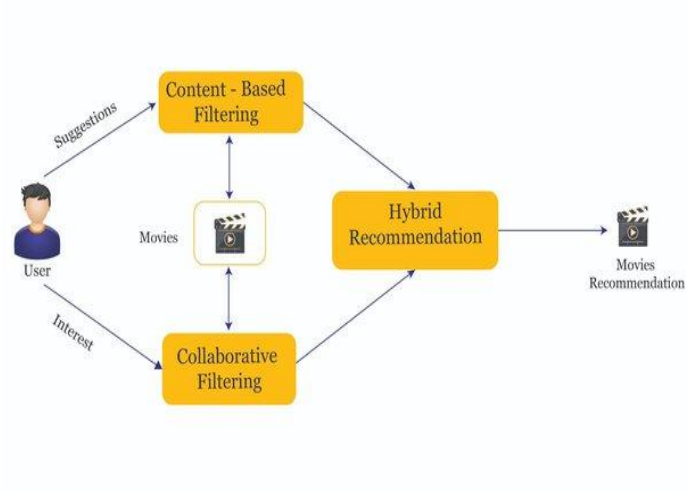

Fig – 4 Hybrid Filtering technique [10]

1.2 Problems Related to the Recommendation System

There are some problems related to the recommendation system are According to [9, 11, 12].

A. **Cold-start problem**: when a user registers for the first time, he/she not watched any movie. So, the recommendation system does not have any movie based on which it can give result [9]. This called cold-start problem.

B. **Data scarcity problem:** This problem occurs when the user has rated very few items based on which it is difficult for the recommendation system to accurate result [9].

C. **Scalability:** In this, the encoding goes linearly on items. The system works efficiently when the data set is of limited size [9].

1.3. Machine Learning Algorithms

Which focuses on using data and algorithms to mimic how humans learn, gradually improving its accuracy. There are three types of machine learning techniques: supervised, unsupervised, and Semi-supervised and reinforcement learning. The machine learns under supervision in Supervised Learning. It includes a model that can predict with the help of a labelled dataset. A labelled dataset is one in which the target answer is already known. The following supervised learning algorithms are widely used: decision tree, logistic regression, support vector machine, and k-nearest neighbors. Unsupervised Learning involves the machine learning on its own using unlabeled data. The machine attempts to find a pattern in the unlabeled data and

responds. K-means clustering, Hierarchical clustering, and the apriority algorithm are the most commonly used unsupervised learning algorithms.

A. K-Means Algorithm

Clustering[8] is a process to group a set of objects in such a way that objects in the same clusters are more similar to each other than to those in other clusters.

Unsupervised learning algorithm K-Means Clustering divides the unlabeled dataset into various clusters. Here, K specifies how many pre-defined clusters must be created as part of the process; for example, if K=2, there will be two clusters, if K=3, there will be three clusters, and so on.

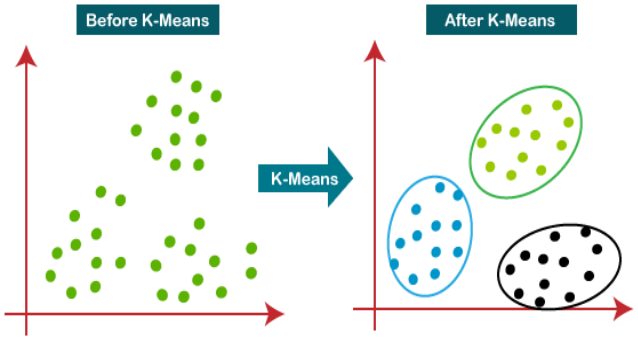

Fig – 5 K-means clustering [13]

Different types of distance measures, including the Euclidean distance measure, are supported by K-Means clustering.

- The Manhattan distance
- A Euclidean distances measurement of distance

Cosine distance calculation

1) Euclidean distances: The most common application is calculating the distance between two points. If we have two points P and Q, the Euclidean distance is a straight line. It is the distance in Euclidean space between two points. Formula for Euclidean distance is:

B. K-Nearest Neighbors

The k-nearest neighbor algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier that uses proximity to classify or predict the grouping of a single data point. While it can be used for either regression or classification problems, it is most commonly used as a classification algorithm, based on the assumption that similar points can be found nearby.

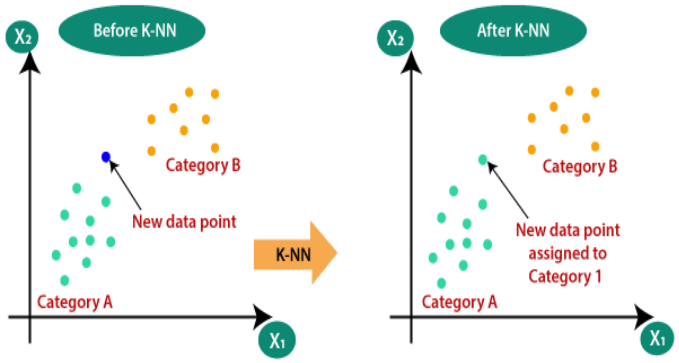

Fig – 6 K- Nearest Neighbors [14]

C. Support Vector Machine

A Support Vector Machine" (SVM) can be applied to classification or regression problems. However, classification issues are where it is most frequently used. When using the SVM algorithm, each data point is represented as a point in n-dimensional space, with each feature's value being the value of a specific coordinate. Next, we perform classification by identifying the hyper-plane that effectively distinguishes.

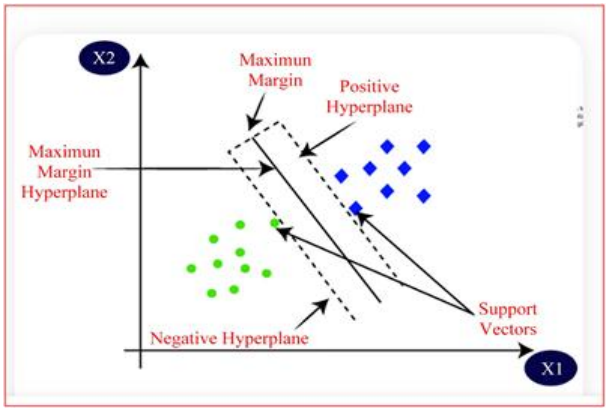

Fig – 7 Support Vector Machine [15]

2. LITERATURE REVIEW

Raghavendra C K and Srikantaiah K.C [16] are discussed about the various similarity metrics applied in recommendation tasks are compared by considering MovieLens dataset for developing a recommendation system for movies using user-based and item-based. that item-based CF gives better accuracy and faster than user based. Accuracy respectively is 0.84 and 0.76.

M Ashok, S Rajanna, PV Joshi [17] are discussed about Machine Learning and Sentiment Analysis based techniques are used for further optimizing search query results. This provides the user with quicker and more relevant data, thus

avoiding irrelevant data and providing much needed personalization. Based on classification technique are naïve bayes, svm, maxEnt, Random forest more accurate result provide is maximum Entropy respectively 77.17%

N Banik, MHH Rahman[18] discussed about polarity detection system on textual movie reviews in Bangla by using two popular machine learning algorithms named Na¨ıve Bayes (NB) and Support Vector Machines (SVM) and provided a comparative results where SVM performed slightly better than NB by considering stemmed unigram as feature with an excellent precision of 0.86.

P Vilakone at. [19]All discussed about , to evaluate the performance; collaborative filtering methods are monitored using the *k* nearest neighbors, the maximal clique methods, the *k*-clique methods, and the proposed methods are used to evaluate. experiment and the output showed was more effective results is maximal clique method for accuracy is 59.63%.

Choudhury, S. S., Mohanty, S. N., & Jagadev, A. K. [20] discussed about recommendation models such as Backpropagation (BPNN) model, SVD (Singular Value Decomposition) model, DNN (Deep Neural Network model) and DNN with Trust were compared to recommend the suitable movie to the user. Results imply that DNN with trust model proved to be the best model with high accuracy of 83% with 0.74 MSE value and can be used for best movie recommendation.

3. PROPOSED SYSTEM

Our proposed system employs a number of methods, including the collaborative filtering, SVM, KNN. For the purpose of compare SVM and KNN for better recommended the movie, our proposed model will focus on collaborative filtering and will briefly introduce two algorithms that fall into this category, K Nearest Neighbor (KNN) and Singular Value Decomposition (SVD).

Step 1: The first step is to use the kaggle movie lens dataset to predict recommended movies for each individual.

Step 2 The second step is carried out by employing two techniques to investigate the data at hand. To see a more comprehensive EDA guide.

Step 3: The third step would be to introduce two collaborative-based filtering algorithms for matrix decomposition, KNN and SVM.

Step 4: The fourth step is to evaluate the model and apply the egression evaluation metrics to our recommendation system.

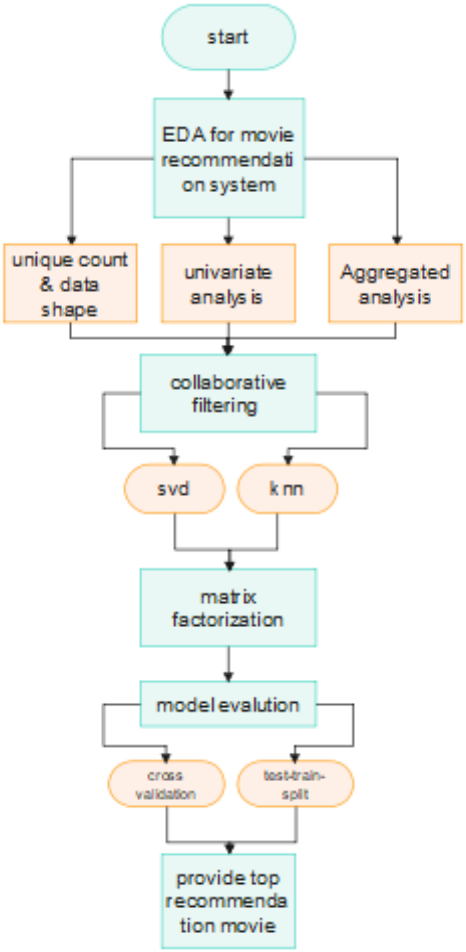

Fig – 8 Proposed Model

4. CONCLUSION

In this paper, we use collaborative filtering to implement a movie recommendation system. It is clear from the recommendation that the research project helps users exert less effort. In the proposed system, a system for suggesting movies to users is built using collaborative filtering to predict higher ratings.

REFERENCES

[1] Lin, C. H., & Chi, H. (2019, March). A novel movie recommendation system based on collaborative filtering and neural networks. In *International Conference on Advanced Information Networking and Applications* (pp. 895-903). Springer, Cham.

[2] Wu, C. S. M., Garg, D., & Bhandary, U. (2018, November). Movie recommendation system using collaborative filtering. In *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)* (pp. 11-15). IEEE.

[3] Kumar, S., De, K., & Roy, P. P. (2020). Movie recommendation system using sentiment analysis from microblogging data. *IEEE Transactions on Computational Social Systems*, *7*(4), 915-923.

[4] Subramaniam, R., Lee, R., & Matsuo, T. (2017, December). Movie Master: Hybrid Movie Recommendation. In *2017 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 334-339). IEEE.

[5] Ahuja, R., Solanki, A., & Nayyar, A. (2019, January). Movie recommender system using K-Means clustering and K-Nearest Neighbor. In *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 263-268). IEEE.

[6] S. (2022, August 12). *System Design Interview: Recommendation System Design (As Used By YouTube, Netflix etc.)*. Medium. Retrieved November 24, 2022, from https://medium.com/double-pointer/system-design-interview-recommendation-system-design-as-used-by-youtube-netflix-etc-c457aaec3ab

[7] Grimaldi, E. (2019, June 25). *How to build a content-based movie recommender system with Natural Language Processing*. Medium. https://towardsdatascience.com/how-to-build-from-scratch-a-content-based-movie-recommender-with-natural-language-processing-25ad400eb243

[8] Moshksar, R. (2021, March 13). *Recommender systems: context-based & hybrid recommender systems*. Reza Moshksar. https://omnidatatec.com/recommender-systems-context-based-hybrid-recommender-systems/

[9] Singh, R. H., Maurya, S., Tripathi, T., Narula, T., & Srivastav, G. (2020). Movie recommendation system using cosine similarity and KNN. *International Journal of Engineering and Advanced Technology*, *9*(5), 556-559.

[10] *What are the types of recommender systems?* (n.d.). Quora. Retrieved November 24, 2022, from https://www.quora.com/What-are-the-types-of-recommender-systems

[11] Mahata, A., Saini, N., Saharawat, S., & Tiwari, R. (2016, December). Intelligent movie recommender system using machine learning. In *International Conference on Intelligent Human Computer Interaction* (pp. 94-110). Springer, Cham.

[12] Ghazanfar, M., & Prugel-Bennett, A. (2010). Building switching hybrid recommender system using machine learning classifiers and collaborative filtering. *IAENG International Journal of Computer Science*, *37*(3).

[13] *K-Means Clustering Algorithm - Javatpoint*. (n.d.). www.javatpoint.com. Retrieved November 24, 2022, from https://www.javatpoint.com/k-means-clustering-algorithm-in-machine-learning

[14] *K-Nearest Neighbor (KNN) Algorithm for Machine Learning - Javatpoint*. (n.d.). www.javatpoint.com. Retrieved November 24, 2022, from https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

[15] Boateng, E. Y., Otoo, J., Abaye, D. A., & Publishing, S. R. (2020, September 29). Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review. Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review. Retrieved November 24, 2022, from https://www.scirp.org/journal/paperinformation.aspx?paperid=104256

[16] Raghavendra, C. K., & Srikantaiah, K. C. (2021, May). Similarity Based Collaborative Filtering Model for Movie Recommendation Systems. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1143-1147). IEEE.Meshram, N. G., & Bhagat, A. P. (2014, February). Connotative features based affective movie recommendation system. In *International Conference on Information Communication and Embedded Systems (ICICES2014)* (pp. 1-7). IEEE.

[17] Ashok, M., Rajanna, S., Joshi, P. V., & Kamath, S. (2016, March). A personalized recommender system using machine learning based sentiment analysis over social data. In *2016 IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (pp. 1-6). IEEE.

[18] Banik, N., & Rahman, M. H. H. (2018, September). Evaluation of naïve bayes and support vector machines on bangla textual movie reviews. In *2018 international conference on Bangla speech and language processing (ICBSLP)* (pp. 1-6). IEEE.

[19] Vilakone, P., Park, D. S., Xinchang, K., & Hao, F. (2018). An efficient movie recommendation algorithm based on improved k-clique. *Human-centric Computing and Information Sciences*, *8*(1), 1-15.

[20] Choudhury, S. S., Mohanty, S. N., & Jagadev, A. K. (2021). Multimodal trust based recommender system with machine learning approaches for movie recommendation. *International Journal of Information Technology*, *13*(2), 475-482.