MRML-Movie Recommendation Model with Machine Learning Techniques

Ravi Gorli[1], Bagusetty Ajay Ram[2]

[1]CSE, GST, Gitam University
Email: *ravigorlee[at]gmail.com*

[2]CSE, GVPCOE
ajay.ram05[at]gmail.com

Abstract: *In the traditional days people use to visit theaters to watch movies. Some of the reviews used to give some view to visit which movie in what kind of theater and so on. According to the reviews and rating people use to visit the theaters and enjoy the movie. Some movies used to run in theaters for many days like starting with 50, 100 to 365 days and so on. Depending on various suggestions the theater owners use to predict the best movie in several aspects like actors, directors, story, songs and many other factors. In the present days it is very hard to see that a single movie is not in any theater for at least 20 to 50 days. OTT has done a lot for replace the platform for movie lovers. Within 10 to 30 days every movie is entering into the OTT platform as a lot of users are increased for the OTT. Different OTT platforms such as Amazon prime, Netflix, Hot star, Zee cinema, have attracted a lot of movie lovers where they are providing databases from past old movies to the present new movies. New movie rights are brought by many of them in a competitive way. Along with movies they are casting several daily serials, web series and different live streaming also. All the databases are maintained by each of them with high performance servers and attracting the users with reasonable subscription charges.*

Keywords: Movie, Review, Automation, Subscription, OTT
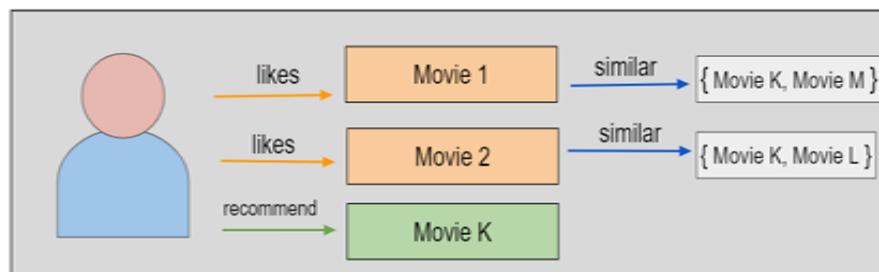
1. Introduction

Figure 1: Movie Recommended system

OTT platforms maintaining a large database is ok but they have to attract more users by providing an interactive platform. Like a movie should be suggested to the user depending on his mood interest. Review based and rating based suggestions for the new movies. In order to provide a more interactive platform to attract more users, a best search and suggestion-based methodology should be implemented. In this research a lot of literature survey is done as followed next, based on all algorithms more interactive platform will be proposed.

2. Literature Survey

F. Furtado & A. Singh, had proposed a model with its need to build up a system for the users to get suggestions based on the movie recommendation content-based system. Their system suggests movies which are critically acclaimed with maximum ratings by the end target of average viewers. They implemented a model that is going to decide with some metric scores for rating movies, calculating the score for each and every movie and then based on the scores and the results they are sorted. They evaluated data on IMDB's dataset. They got higher accuracy when compared to the previous models. Their main task is based on the rating and

preference by the value given by the user. Their system was based on a collaborative approach, by using different systems they explicit progress outcomes based on content-based approaches. Their system works on every user's individual rating. Their approach also has the connection between different clients and relies on their ratings, prescribing movies to others who have similar tastes. They developed a web application that allows users to rate movies along with recommendations depending on others ratings.

R. Ahuja, et. al, had explored by building a movie recommender system using clustering algorithms such as KNN-K-Nearest Neighbor algorithms, K-Means Clustering. They used a kaggle dataset for their findings. They implemented the technique using python scripts. The proposed work deals with the introduction of various concepts related to machine learning and recommendation systems with the usage of various tools and techniques for building their system. Various algorithms such as KNN-K-Nearest Neighbor algorithms, Content-Based Filtering-Means Clustering, Collaborative Filtering, have been studied in detail. After a lot of experiments using the above algorithms, they got a decision maker to apply which algorithm for which task they have to use in the areas of industries with the recommender systems, e-commerce, etc.

Various building blocks are proposed with an Architecture followed by Process Flow and Pseudo Code using all these Implementation and Working of the System developed. In their works different cluster values with different values of Root Mean Squared Error are also observed. Their approach gave the suggestion that the number of clusters decreases with the value of RMSE-Root Mean Squared Error also decreases. The best value of RMSE obtained is around 1 which is better compared to other methods.

Choudhury, et. al, given a study that addresses many problems such as recommendation of movies to watch with a decision taken in less time and also which movie to watch according to the user requirements. They proposed a trust matrix measure for combining similarities of users with propagation of weighted trust. According to Gorli, Ravi, A set of four models such as DNN-Deep Neural Network model, BPNN-Backpropagation Neural Network model, SVD-Singular Value Decomposition model and DNN-Deep Neural Network with Trust models are tested for recommendation and different observations which are measured are compared among all the algorithms concluding with the suggestion for recommending a movie for the user to view. Training and testing was done with different parameters and observations finally among all DNN-Deep Neural Network with Trust models gave better performance with a high accuracy of around 83% with 0.7 MSE-Mean Squared Error

Cintia Ganesha Putri, et al, developed a film recommender system for the users to recommend a movie in their recommendations. As there are a lot of movies available, choosing the best movie according to the user requirement is difficult as the amount of information about movies is increasing day by day. They proposed a model which will be useful for the customers to choose according to their preferred movies with required features compared with existing features. They implemented several algorithms in order to obtain different groupings, such as the Birch algorithm, mean-shift algorithm, *K*-Means algorithm, mini-batch *K*-Means algorithms, agglomerative clustering algorithm, spectral clustering algorithm and affinity propagation algorithm. They proposed methodology for optimizing *K* in order that each and every cluster may not able to significantly increase variance. Their findings discovered methods which are better in evaluating cluster algorithms. For finding the qualities of the recommender system, we adopted the MSE-Mean Square Error, such as the DN-Dunn Matrix and CVI-Cluster Validity Indices, and SNA-Social Network Analysis, such as DC-Degree Centrality, CC-Closeness Centrality, and BC-Betweenness Centrality. They also worked on Average Similarity, CT-Computational Time, and AR-Association Rule with Apriori analysis algorithm and evasion with cluster performance and compared recommended system performance using Silhouette Coefficient, Calinski-Harabasz Index, and Davies–Bouldin Index.

Kumaar, Et al, proposed a model with human diminish endeavor for recommended movies based models according to the interests of the users. Finding the correct movie to watch is always a difficult task. Sometimes users get frustrated, taking all these issues of the customers. Kumaar,

Et al, proposed a model for resolving troubles that occur, with building a model with the evaluation of content-based approach. Their main focus is to build a model which recommends a movie based on descriptions of movies and considering all other features. They have taken an example of the movie "GoldenEye", they evaluated and observed the result movie as "Skyfall" with a similarity score of near to 67% using the method of Count vectorizer, using the Jaccard Recommender they got near to 13% value, using TF-IDF Keywords they got near to 14. %, and using TF-IDF Plot Overview they got near to 10% using and using Google form they got near to 72%. Zhang, Jiang, et. al, [6], Constructed a real personalized web-based movie recommendation system and proposed a model for addressing two different issues. A High-efficient simple recommendation algorithm is proposed, which exploits attributes of users' profile for partitioning each of them into several mini clusters with a conception of virtual opinion leader, user-item matrix is reduced significantly. Gorli, Ravi, the recommendation system a Weighted Slope One VU method is designed and implemented. Compared to traditional clustering-based CF recommendation schemes, their method had significantly reduced the time complexity for recommendation of movies, while achieving comparable recommendation performance.

Li, Hui, et al, carried out a program with an intelligent program system for recommendation for different movie recommendation sites. They proposed a novelty model which is turning into the social networks and preferences of the users are mined with microblogs expressing information used for the evaluation of the similarities between episodes of TV and online movies. A gap is built between movie and TV watchers' domain with social networking activities. "Cold-start" is the first approach. Several data mining approaches and social computational models are adopted in their work. They developed the works micro bags for easy recommendation models.

Darban, et al, had proposed a system for recommendation of methods using graph-based model association along with the ratings of user's similarity and with the combination of user's location and demographic information. With consideration of advantages from a utilization as a feature extraction from the Auto-encoder. They extracted new features based on all the attributes with a combination of different methods. Recommendation accuracy is performed with their approach for a new set of clustered user's features. Their method had given significant results for a cold start problem. They performed the measures on the Movie Lens dataset and achieved better results.

Mukherjee, R, et al, proposed a model for building software agents for management tasks and information access for employing user model techniques. They developed a procedure with reasoning which had a systematic and meaningful trade-off between the preferences of the users. They have used the theory of voting for mechanism adaptation for a desirable guarantee for stored preferences with a recommendation. They developed a movie recommendation system for catering based on the user's interests. As per their experimental results they presented the issues and initial results for the employed voting theory for user's model, mainly focusing on important issues in the

context of user model. They proposed modalities of multiple queries through which the user can pose instance-based queries, unconstrained or constrained. They provided a robust and most effective recommendation model.

Bhalse, Nisha. Et, al, had proposed an algorithm with similarity with the combination of the rations of user values and the behavior of user value. Their findings showcased the values behavior of the users from the probability score in accessing the genre data. Therefore, their methods proposed a new algorithm for similarity called UPCS-User-Profile-Correlation-Similarity for examining the data-genre and the profiles data of users, namely gender, age, location and occupation. Finding the weights of all the user similarities of profile data are used to rate the value and behavior of user value. According to Gorli, Ravi, The correlation coefficients between the profile user date and rate of the behavior area calculated the similarity. The experimental results of algorithm UPCS reduced MAE-Mean Absolute Value by 1.64% and RMSE by 1.4%.

3. Proposed Model

In our proposed model, we want to develop a web platform which is going to provide a more interactive user interface, so that the user can view movies according to his/her interest, mood and with any other parameters based. In our interface we provide two methods, one method is to take input from the user with different parameters such as Language (English, Mandarin, French, Hindi, Telugu, Tamil, Spanish, Korean, German, Italian, Russian, etc. .), type of movie (drama, action, romance, horror, thriller, science fiction, animation, children, adventure, comedy, musical, documentary, crime, mystery, fantasy, The prestige), year of release, actor, director, etc. A large database is proposed to develop with movies from the above all parameters.

A Movie Recommendation algorithm is proposed for the calculation of rating which is taken into consideration for every movie with the consideration of various parameters and a dynamic model is also proposed for the rating, corresponding the database is updated from time to time. There are 2 types of rating, one is taken from the user's view who continuously watches the movies and another type is automatically calculated in background from time to time based on all criteria with several parameters and with immediate update. In the second method the movies are recommended automatically according to the rating, mood, previous history, situation based and so on. The web interface is developed in html, phpMySQL or python Django. MySQL database server is used for the storing of movies and its corresponding information like history of the user and several data which is used for the web interface and for the prediction methodology i. e, rating of movies.
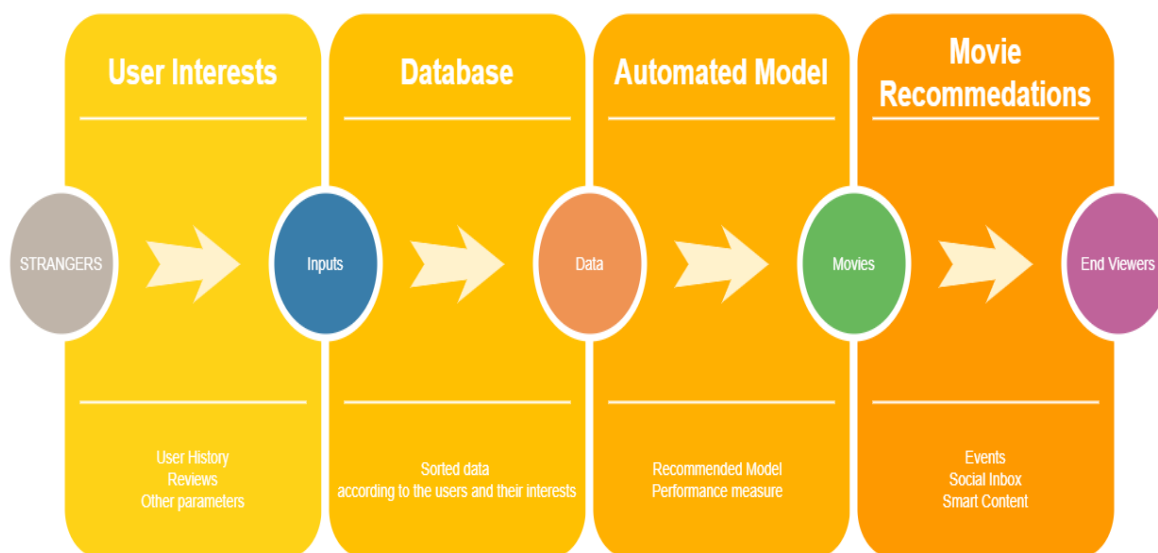

Figure 1: Movie Recommendation Model

The above Fig 1. Movie Recommendation model represents the approach of our model, In which input is taken from the users and the reviews such as user history, reviews and any other required parameters, all are saved in our database with a prior data cleaning methodology and the data is sorted and updated from time to time. From the updated database, the data is taken by our automated recommended model and several algorithms are applied with the calculation of the performance from here the user is recommended with several moves in a user perspective

In our recommendation model we have used 4 different algorithms such as Memory based Collaborative Filtering, Model based Collaborative Filtering, Content-Based Filtering, Deep Learning / Neural Networks.

The Memory based Collaborative filtering approach is divided into two main sections: one is item-item filtering and another user-item filtering. item-item filtering will take an item, find users who liked that item, and find other items that those users or similar users also liked. It takes items and outputs other items as recommendations. A user-item filtering takes a particular user, finds users that are similar to that user based on similarity of ratings, and recommends items that very similar users liked. Item-Item Collaborative Filtering: "Users who liked this item also like" for, User-

Item Collaborative Filtering: "Users who are similar to you are also liked. There are 3 distance similarity metrics that are usually used in collaborative filtering:

Jaccard Similarity: Similarity is based on the number of users which have rated item X and Y divided by the number of users who have ratings either X or Y. It is not a numerical rating but just a Boolean value like a product being bought or an ad being clicked.

Cosine Similarity: Similarity is the cosine of the angle between the two vectors of any item vectors of X and Y. Closer the vectors, smaller the angle and larger the cosine value.

Pearson Similarity: Similarity is the Pearson coefficient between the two vectors. It is used for the purpose of diversity.

Model-based collaborative filtering algorithms are used for describing movie recommendation, the first step of developing a model for calculation of user ratings. Probabilistic approaches and envisioning the process of collaborative filtering by computing the expected value of a user prediction, given his/her ratings on other items.

The Content Based Recommender model is the one which is based on similarities, such as if the user likes one item, he/she would like to the similar item, the same happens in movies if the user had watched/liked one kind movie, the movies which are similar to the previous movies are recommended to the user next time. On the methods of context or properties the recommendation is done. It works on the ratings, user profile. The concept of TF-IDF Term Frequency and Inverse Document Frequency is used for the information retrieval and the vector space model is used based on the user profile information. .

We have used nine models for calculating the performance, The models such as GB-Gradient Boosting, LGBM-Light Gradient Boosting Machine regressorregressor, RF-Random Forest regressor, CB-Cat Boost regressor, GBDT-Gradient Boosting Decision Tree, ET-Extra tree regressor, GOSS-Gradient-based One Side Sampling and EFB-Exclusive Feature Bundling are calculated and the performance scores are graded.

Dataset, we collect data from all the parameters like type of movie, year of release, actor, director, user review, no of views many other all these are saved in a backend database and the dataset is cleaned with the removal of unnecessary, unrelated or partial data.

Several models are developed using different machine learning algorithms such as random forest, regression, support vector machine, Naive bayes, Gradient Boosting and so on. A lot of research is done by taking into consideration parameters which are taken from the viewers and reviewers. In our model we have proposed a new recommendations model by recommendations of movie in demand and situation of the user. By considering different parameters-based approaches we are going to recommend the movies exactly the user wants to view. Along with all the manual

recommendations, an automated model is developed by training our system with different AI algorithms.

The advantage of our model is that a movie recommendation exactly as the user wants is going to predict and it will be given in prior notification to the user that the particular movie gives a good impression, our model is s going to be reviewed by auto med machine which will give more accurate results compared to a model of prediction based on the manual parameters given by the user. Compared to previous models our model gives more accurate and reliable movie recommendations, so that the user will benefit and the business model gets more popular with the reliable recommendations model.

4. Conclusion

Our model is going to be the best one for recommendations of the exact movie for the user with a clear description for each and every view, how the movie is going to be exactly the thinking of the user. The traditional models had developed their models with recommendations of movies using the manual parameters, In our model we have recommended the movie using reliable and user view based The movie recommendations are done with enough of algorithms and the performance is also calculated and compared with the previous models, and the model is going to give best and better results.

References

[1] F. Furtado; A. Singh. "Movie recommendation system using machine learning". *International Journal of Research in Industrial Engineering*, 9, 1, 2020, 84-98. doi: 10.22105/riej.2020.226178.1128

[2] R. Ahuja, A. Solanki and A. Nayyar, "Movie Recommender System Using K-Means Clustering AND K-Nearest Neighbor, " *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence),* 2019, pp.263-268, doi: 10.1109/CONFLUENCE.2019.8776969.

[3] Gorli, Ravi. "Interlinking OF IoT, big data, smart mobile app with smart garbage monitoring. " *International Journal of Computer Sciences and Engineering* 5.1 (2017): 70-74.

[4] Choudhury, SasmitaSubhadarsinee, Sachi Nandan Mohanty, and Alok Kumar Jagadev. "Multimodal trust based recommender system with machine learning approaches for movie recommendation. " *International Journal of Information Technology* 13.2 (2021): 475-482.

[5] CintiaGanesha Putri, Debby, Jenq-Shiou Leu, and Pavel Seda. "Design of an unsupervised machine learning-based movie recommender system. " *Symmetry* 12.2 (2020): 185.

[6] Gorli, Ravi. "Smart Jacket for Industrial Employee Health and Safety. " *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol* 2 (2017): 390-392.

[7] Kumaar, Hareesh, S. Srikumaran, and S. Veni. "Content-based Movie Recommender System Using Keywords and Plot Overview. " *2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET).* IEEE, 2022.

[8] Gorli, Ravi. "A New Approach for Employee Safety in Industries with IoT." *i-Manager's Journal on Information Technology* 7.2 (2018): 22.

[9] Zhang, Jiang, et al. "Personalized real-time movie recommendation system: Practical prototype and evaluation." *Tsinghua Science and Technology* 25.2 (2019): 180-191.

[10] Li, Hui, et al. "An intelligent movie recommendation system through group-level sentiment analysis in microblogs." *Neurocomputing* 210 (2016): 164-173.

[11] Gorli, Ravi. "VOICE OVER CONTROL TO INTERNET OF THINGS WITHDEEP LEARNING. " *Editorial Board* 9.2 (2020): 7.

[12] Darban, Zahra Zamanzadeh, and Mohammad Hadi Valipour. "GHRS: Graph-based hybrid recommendation system with application to movie recommendation. " Expert Systems with Applications 200 (2022): 116850.

[13] Mukherjee, R., Sajja, N. & Sen, S. A Movie Recommendation System – An Application of Voting Theory in User Modeling. User Model User-Adap Inter 13, 5–33 (2003). https://doi.org/10.1023/A: 1024022819690

[14] Bhalse, Nisha, and Ramesh Thakur. "Algorithm for movie recommendation system using collaborative filtering." Materials Today: Proceedings (2021)

[15] Gorli, Ravi, Ch Demudu Naidu, and G. Pandit Samuel. "Proposal of ASLR for Voice Disorders." Soft Computing for Intelligent Systems: Proceedings of ICSCIS 2020. Springer Singapore, 2021.