Movie recommendation system using enhanced content-based filtering algorithm based on user demographic data

G.Sunandana[1], M.Reshma[2], Y.Pratyusha[3], Madhuri Kommineni[4], Subbarao gogulamudi[5]

[1, 2,3,4,5] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Greenfields, Vaddeswaram, Guntur-522502

[4]madhuri.cbit@gmail.com

Abstract-

Now-a-days people are consuming content in form of movies, series, etc. for entertainment. In this modern era, people always look up to entertainment and in that process, they waste their time in searching for movies. Everyone wants to watch good films that have great content. It takes lot of time to search for a movie they like. Recommendation system comes into play in such situations. It helps to people by recommending movies. This paper develops a Movie Recommendation System to recommend movies based on different parameters. The principal objective of the project is to construct a movie recommendation framework to prescribe pictures to users. There are many algorithms that help to build a recommendation system. Here, the Content-based algorithm has been employed to recommend movies based on the similarity with other films by analyzing the content of the movie. To find the similarity, the cosine similarity method has been used. Here, the cosine similarity has been computed by using linear kernel, where the parameters are taken by the result of TF-IDF vectorization. Then the most similar movies are recommended.

Keywords—Content based, Recommendation System

I. INTRODUCTION

Recommendation system helps in the predictions of the data based on previous data of user. These are used to save time by the user. Movie Recommendation system is a system which helps the users by recommending movies based on previous data of the user. It mainly relies on the browsing history and previously watched films and those rated by the user.

Movie Recommendation engines provides its users with films that are similar to those that they have seen and or rated. The recommendations are based with users that have similar interests. They help the users by mapping them with the content they like to watch. These days there is lot of content on internet, and we don't know what to watch. So, recommendation engines help us by recommending movies based on our own interests.

There are many algorithms out there to build recommendation engines. In this project we are using Content-based filtering algorithm to build our engine. There are many ways to approach using Content-based filtering algorithm. We are using tf-idf vectorization to compute cosine similarity which helps recommend movies.

Content Based Filtering: Content-based frameworks consider the portrayal and highlights of an item alongside the client's inclinations to give suggestions. **TF-IDF Vectorizer:** TF-IDF is shortened form for term Frequency Inverse document Frequency. Set of rules to change over text appropriate into a significant portrayal of numbers that's utilized to fit framework set of rules for suggestion.

Any recommendation engine can commonly deal with different types of users. Here we deal with active users. Active users are the ones that have seen and rated few films. Active users are recommended based on their watch history and similar user interests.

II. LITERATURE SURVEY

Coordinating literature survey going before start on position paper is fundamental in

understanding the AI based algorithms. A Proposal framework or suggestion engine is a model utilized for data sifting where it attempts to anticipate the inclinations of a client and give recommendations dependent on these inclinations. In any recommendation engine we commonly deal with two types of users. They are active users and cold-start users. Active users are the ones that have seen and rated few films.

[1] Registers semantic likeness between films utilizing relevant highlights and predicts evaluations of the unrated films. The Extent of this paper was Use of profound learning procedures, primarily word portrayal models, over text-based information to recognize context-oriented highlights appears to be one of the promising headings of examination for the advancement of setting mindful recommender frameworks.

[10] Proposed a deep neural network algorithm for generalized recommendation system.

[7] In this paper, it permits a user to pick his decisions from a given set of attributes then suggest him a pic list supported the additive weight of various attributes and victimization K-means algorithm.

[18] Observed that the specific domain seekers recommendation is done using Ontology based Recommendation System by finding the hidden relationships among entities.

[16] Observed that the collaborative filtering technique is applied in recommendation of business for users.

[6] It is proposed that the cluster based recommender system for social media data analysis.

[3] The subjective take a look at results show the employment of sentiment analysis to get rid of cold begin drawback. Therefore, the take a look at shows however songs are counseled to new users even once there's no info regarding the new users.

[14] Observed that the usage of K Nearest neighbor technique and map reduce technique for cuisine-based recommendations.

[19] A few suggestion frameworks are projected that region unit upheld agreeable separating, content based and hybrid proposal ways and afterward route the majority of them can tackle the issues though giving higher suggestions.

[20] Specified that the deep learning technique is deployed using hybrid recommender system in bid data.

[24] Observed that the social votes recommendations done based on Collaborative filtering technique.

[23] Our tests propose that item-based calculations offer drastically better than user-based calculations, while at consistent time giving higher quality than the most effective on the market user-based algorithms.

[7] In this paper they have executed a steady and exact recommender framework a hybrid based and also collaborative is utilized.

[21] The accuracy of the generated content-based predictions of the recommender that uses the optimizations was measured, therefore on verify the degree of improvement gained by the optimization ways.

[11] In this paper, a moving-picture show suggestion structure upheld by hybrid proposal and sentimental examination is projected to support the exactness of recommender frameworks.

[25] This paper is to show various difficulties identifying with the methods that square measure being utilized for creating proposals.

[4] In this paper it covers collaborative filtering approach on film trust movie data set to provide recommendation of movies for a particular user.

[8] It is observed that how collaborative filtering technique and regression technique are used to recommend the location travel

system due to covid19 pandemic situations. [13] In this paper we introduced and tested assessed another calculation for CF-based recommender frameworks. Our outcomes shown that item-based strategy hold the guarantee of permitting CF-based calculations to scale to huge informational collections simultaneously to deliver excellent proposals.

[2] Another estimation of utilizing both client's and item's ascribes and their connection for top-N suggestion.

[17] In this paper, an examination was performed on Movie Lens knowledge set, and subsequently the outcome shows procedure arranged during this paper performs higher than standard method in proposal exactness and quantifiability.

[12] Observed the analysis of risk factors for movie recommendations using machine learning technique.

[15] This paper estimated the exhibition of recommender framework utilizing a solitary model (for example k-NN) and utilizing various models (for example k-NN and Network Factorization). However utilizing different models for recommender framework, give better execution.

[22] It is observed that the recommendations done by user-based recommendation with a scalable machine learning tool.

[5] In this paper they have implemented personalized recommendation system for any user. The Future Scope of this paper was it can improve the recommendation system for cold start user.

III. STRUCTURED AND UNSTRUCTURED DATA

The information separated from clients and the hotspot for recommender scheme comprise of organized and unstructured information. Information stockroom doesn't deal with unstructured information, which is a significant hotspot for a recommender scheme.

This is wised up by utilizing enormous information.

IV. DATASET

We have taken different datasets for active users from Kaggle. For the active user, we have taken two datasets. First one is credits dataset which contains movieId, cast crew, title and second one is movies dataset which has budget, genre, id, popularity, vote count, vote average, keywords and the popularity given to the movie by the user.

V. METHODOLOGIES

The movie recommendation engine is used for recommending movies to users that has been widely used. We have different types of users. Here we deal with active user to whom we suggest/recommend movies.

For Active users- we use cosine similarity from content-based filtering. To evaluate the similarity between the user's scores, we used Cosine Similarity which is computed using linear kernel and tf-idf vectorization.
Step1: Data Preprocessing and applying IMDB formula.

return (v/(v+m)*R) + (m/(m+v)*C)

Step2: The TF-IDF algorithm is used to match device set of rules for prediction.

  Tfidf=TfidfVectorizer (stop_words='english')
Step3:
The Figure gives the numerical equation to the Pearson Correlation. The higher the coefficient is, the more associated the two clients' decisions are-

$$R= \frac{\sum(x-x')(y-y')}{\sqrt{\sum(x-x')^2\sum(y-y')^2}}$$

Start
FUNCTION get_recommendation
    Get the index of the movie that matches the title
Get the pairwise similarity scores of all the movies with that movie
Sort the movies based on similarity scores
Get the scores of the 10 most comparative films
Get the movie indices
Return the main 10 most comparable films
stop
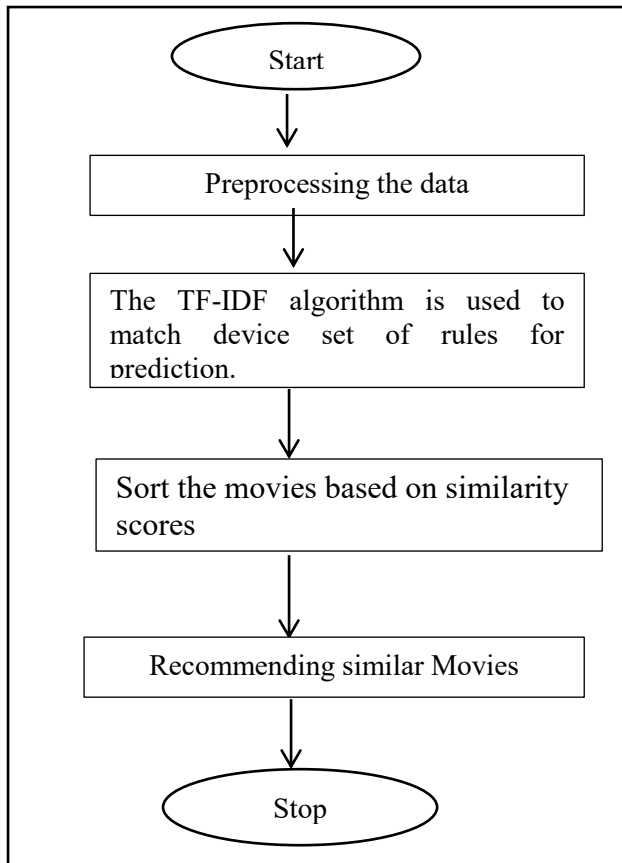
Figure 1: Flow of recommendation system

VI. RESULT

We used demographic features of users such as vote count and vote average to apply IMDB formulae for getting the weights of movies. Popularity of each movie can be calculated based on the outcomes of the IMDB values.

We used this popularity column to recommend movies based on the popularity of the movies to users as given in Figure. 2.
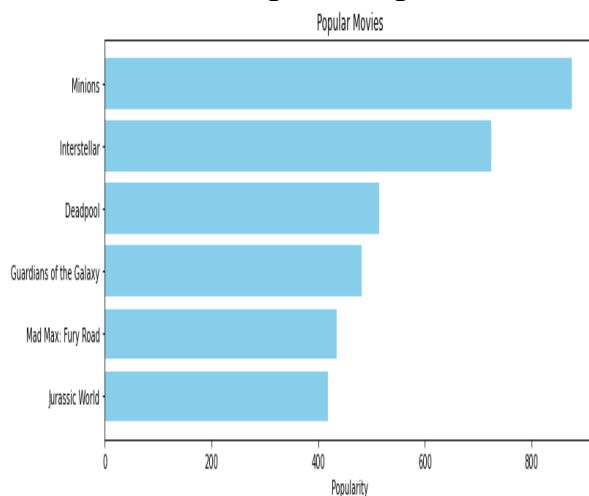
Figure 2:Comparision of top movies based on popularity

VII. CONCLUSION

This paper focussed on the recommendation of movies by using content based algorithm. This system is developed based on highlights like the class of the film, the chiefs, the entertainers, etc. which could be considered also to give ideas. For future work, the recommender plan could be made using hybrid isolating methodology as opposed to content based. Continuous assessment shows that cross variety plans are found to be really convincing and give more accurate ideas. Therefore, hybrid plan would be an improvement.

REFERENCES

[1] Vineet K, Muhammad Abulaisha "Context-Based Recommender System using Collaborative Filtering and LOD",2020

[2] Ming He, Bo Wang "Exploring Knowledge in Heterogeneous Information for Movie Recommendation",(2019)

[3] Supriya Singh, Abhisheks "Recommender System using Sentiment Analysis",(2018)

[4] Gourav Bathla ,Deepak Yadav "Movie Recommendation System Using Collaborative Filtering",(2018)

[5] Plan and Implementation of Movie Recommendation System Based on Knn Collaborative Filtering Algorithm by Authors: "Kevin Shah Jeet Gandhi".(2017)

[6] Kumar, K.K., Kiran, T.S.R., Kumar, M.R., Aruna Sri, P.S.G., "An Interactive Clustered base Recommender system for effective Social Data Analysis", IOP Conference Series: Materials Science and Engineering, 2020.

[7]MOVIEMENDER-A Movie Recommender System by Authors:"Rupali Hande, Ajinkya Gutti".(2016)

[8] Daya Sagar, K.V., Arunasri, P.S.G., Sakamuri, S., Kavitha, J., Kamesh, D.B.K, "Collaborative Filtering and Regression Techniques based location Travel Recommender System based on social media reviews data due to the COVID-19 Pandemic", IOP Conference Series: Materials Science and Engineering, 2020.

[9] Kumar, Manoj & Yadav, D.K. & Singh, Ankur & Kr, Vijay, "A Movie Recommender urnal of Computer Applications. 124. 7- 11. 1System: MOVREC", 2015 International Jo0.5120/ijca2015904111

[10] Anitha, R., Krishnaiah, K., Anusha, G., "A generic recommender system for continuous optimization based on deep neural network", Annals of the Romanian Society for Cell Biology, 2021.

[11]A Sentiment-Enhanced Hybrid Recommender

System for Movie Recommendation: A Big Data Analytics Framework: Yibo Wang,1 Mingming Wang.

[12] Gupta, S., Rao, K.V., Dubba, N., Subrahmanyam, K., "Risk Analysis in Movie Recommendation System Based on Collaborative Filtering", Advances in Intelligent Systems and Computing, 2021.

[13] B. Sarwar, G. Karypis, 1. Konstan, 1. Riedl, "Item-based collaborative filtering recommendation algorithms", Proceedings of the 10th international conference on World Wide Web, pp. 285-295, 2001.

[14] Kodali, S., Dabbiru, M., Rao, B.T., "A Cuisine Based Recommender System Using k-NN And Mapreduce Approach", International Journal of Innovative Technology and Exploring Engineering, 2019

[15] Jain, A., & Vishwakarma, S. K., "Collaborative Filtering for Movie Recommendation using Rapid Miner" International Journal of Computer Applications (0975 - 8887) Volume 169 - No.6, July 2017.

[16] Rokade, P.P., Aruna Kumari, D., "Business recommendation based on collaborative filtering and feature engineering - Aproposed approach", International Journal of Electrical and Computer Engineering, 2019

[17] Collaborative filtering idea estimation subject to blend customer model by Author: Qian Wang, Xianhu Yuan, Min Sun.

[18] Selvaraj, P., Burugari, V.K., Sumathi, D., Nayak, R.K., Tripathy, R., "Ontology based Recommendation System for Domain Specific Seekers", Proceedings of the 3rd International Conference on I-SMAC IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2019

[19] A Survey of Recommender Systems: Approaches and Limitations by Author: Meenakshi Sharma.

[20] Ramanujam, S.S., Sivaneshwar, P., Naren, J., Madhumitha, S., Vithya, G., "A study on hybrid recommender system with deep learning and deployment in big data", Test Engineering and Management, 2019

[21]A Content Based Movie Recommendation System by Author:- "ÖZNUR KIRMEMİŞ".

[22] Veena, C., Vijaya Babu, B., "A user-based recommendation with a scalable machine learning tool", International Journal of Electrical and Computer Engineering, 2015

[23] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms", In WWW '01: Proceedings of the tenth international conference on World Wide Web, New York, NY, USA, 2001.

[24] Ismail, M., Prakash, K.B., Rao, M.N., "Collaborative filtering-based recommendation of online social voting", International Journal of Engineering and Technology(UAE), 2018

[25] Prodan Andrei-Cristian, "Implementation of a Recommender System Using Collaborative Filtering", Studia univ. Babes_bolyai, Informatica, volume lv, number 4, 2010.