MOVIE RECOMMENDATION SYSTEM

Charulata Misal
Computer Science and Engineering
GH Raisoni, College OF Engineering
Nagpur, India
misal_charulata.cs@ghrce.raisoni.net

Ishita Borghate
Computer Science and Engineering
GH Raisoni, College OF Engineering
Nagpur, India
borghate_ishita.cs@ghrce.raisoni.net

Priyanka Singh
Computer Science and Engineering
GH Raisoni, College OF Engineering
Nagpur, India
singh_priyanka.cs@ghrce.raisoni.net

Muskan Gupta
Computer Science and Engineering
GH Raisoni, College OF Engineering
Nagpur, India
gupta_muskan.cs@ghrce.raisoni.net

Prof. Prashant Khobragade
Computer Science and Engineering
GH Raisoni, College OF Engineering
Nagpur, India
prashant.khobragade@raisoni.net

Abstract—**Many prominent applications employ the recommendation system, which plays an important part in the present period. The suggested method summarizes the approaches and techniques developed by the Collaborative Filtering-based recommendation system. Collaborative Filtering, Content-based, and Hybrid-based techniques were developed from the recommendation system. Collaborative filtering is characterized using different philosophies, for example, grid factorization, client based proposal, and thing based proposal in the recommended path. This software extracts aspect-based ratings from reviews and suggests reviews to users based on user similarities and rating trends. Finally, the suggested movie recommendation system was validated using multiple evaluation criteria, and the proposed system outperformed existing systems.**

Keywords— Recommendation System, Matrix factorization, Collaborative Filtering, Content based

I. INTRODUCTION

Suggestion frameworks are turning out to be more pervasive, and thus, many specialist organizations creation costs are rising. Because of today's overcrowding, recommendations are essential when promoting products or services. Recommender frameworks, then again, diminish exchange costs while expanding client experience and independent direction. The plan of a film suggestion framework presents a critical test, in light of the fact that other proposal systems rely on service providers and product distributors to provide quick calculation and processing services. Users' ratings are first gathered to suggest movies, and then the top list of items is suggested to the target user. Prior to watching a film, clients may likewise peruse surveys left by different clients. Several recommendation algorithms, such as collaborative filtering, content-based recommender systems, and hybrid recommender systems, have been described.

1. Collaborative filtering is a well known strategy for making recommender frameworks. The method creates recommendations exclusively based on information from rating profiles for various persons or goods. They generate suggestions by searching for peer users/items in this area who have a rating history like the present user or item.

2. A content-based recommender uses information provided by the user, either directly or implicitly. A client profile is made in view of this data, and giving proposals to the user is then used. The engine grows increasingly accurate as the user offers more inputs or acts on the recommendations.

.

II. LITERATURE REVIEW

The purpose of this project is to create a decent recommendation system based on movies in a user-friendly and simple manner, where one may have numerous choices depending on the interest formed by viewing the current movie in that context. Given below is the table based on observations made in previous research papers.

In these research papers, the authors have presented a wide approach to the recommendation system using various machine learning techniques. The main goal of this work is to provide an easy movie recommendation system in context with the real world.

Recommender Systems: An overview of different approaches to recommendations:

In this paper, the authors have, discussed that popular approaches to recommendations are content-based, collaborative and Hybrid. In the collaborative filtering technique, suggestions for every client are created by making correlations with the preference for one option over one more of different clients who have qualified the item comparatively to the dynamic client. The content based channel breaks down set of descriptors feed to it as contribution for a specific thing which is recently evaluated by the client. Content and Collaborative techniques are mixed in crossover recommenders to further develop execution. An outline of this large number of approaches shows that they enjoy both benefits and detriments.

Recommendation Systems with Machine Learning:

In this paper, the authors discussed that the given a system that has a huge number of users and a similar amount of content to present for them, the filtering process becomes crucial. A collaborative filtering-based recommendation model was created, using a new and open-source machine learning framework, ML.NET. The best recommendation method can be chosen by carefully observing what type of data is available and what problem should the recommendation address.

Content Based Movie Recommendation System:

In this paper, the authors discussed that the recommendation system aims to predict or take users' interests and recommend related items that quite likely are interesting for them. Recommender frameworks are utilized for suggesting items, producing playlists, matchmaking, furthermore significantly more. Recommender frameworks work with trademark data and client thing connections.

III. METHODOLOGY

This project is driven by Machine Learning algorithms, and it will help the users to find a better recommendation system. Following are the processes for this movie recommendation system.

A. Collaborative Filtering

Collaborative filtering is a popular method for creating recommender systems. Collaborative filtering is founded on the notion that individuals who have previously agreed would agree again and adore comparable things in the future. The method creates recommendations exclusively based on information from rating profiles for various persons or goods. They generate suggestions by searching for peer users/items in this area who have a rating history similar to the present user or item. Memory-based and model-based collaborative filtering systems are the two categories.

B. Data Filtering

The process of selecting a smaller subset of your data set for viewing or analysis is known as data filtering. Filtering is ordinarily (however not generally) impermanent; the whole informational collection is saved, yet a part of it is utilized in the computation.

Filtering can be used to

• Check out at results from a particular time span

• Work out results for explicit gatherings

• Eliminate any wrong or "poor" perceptions from a review.

• Create and test factual models.

C. K-Nearest Neighbor (KNN) Algorithm.

1) The K-NN method assumes that the new case/data and existing cases are similar and places the new case in the class that is generally like the to the existing categories.

2) The K-NN algorithm saves all existing data and classifies fresh data points according to their similarity. This means that new data can be quickly sorted into a well-defined category using the K-NN method.

3) The K-NN approach can be used for both regression and classification, but more commonly utilized for classification tasks.

4) The K-NN algorithm is a non-parametric which implies it makes no assumptions about the data. It is likewise called a sluggish student calculation since it doesn't gain from the preparation set quickly rather it stores the dataset and at the hour of characterization, it plays out an activity on the dataset.

5) KNN calculation at the preparation stage simply stores the dataset and when it gets new information, then it orders that information into a classification that is much like the new data.

D. Count Vectorizer

Count Vectorizer is a great tool provided by the scikit-learn library in Python. It is used to transform a given text into a vector based on the frequency (count) of each word that occurs in the entire text. Count Vectorizer creates a matrix in which each unique word is represented by a column of the matrix, and each text sample from the document is a row in the matrix. The worth of every cell is only the include of the word in that specific message test.

E. Cosine Similarity

To work out the distance between the objective film and the motion pictures in the dataset, cosine similitude is utilized. It measures the similarity between two documents irrespective of how different they are in size and calculates the cosine angle between two vectors in multidimensional space

F. CSR Matrix

A sparse matrix or sparse array is a matrix with the majority of its elements being zero in numerical analysis and scientific computing. There is no clear specification of how many zero-value elements a matrix must have to be considered sparse, but one frequent requirement is that the number of non-zero elements must be roughly equal to the number of rows or columns. When the majority of the components are non-zero, the matrix is said to be dense. The sparsity of a lattice is characterized as the quantity of zero-esteemed components partitioned by the complete number of components.

IV.   Design

Proposed collaborative filtering method. the target here is to recommend movies using the item-based technique. First, the extraction of the dataset to collect information about the target movie and therefore the user's rating.
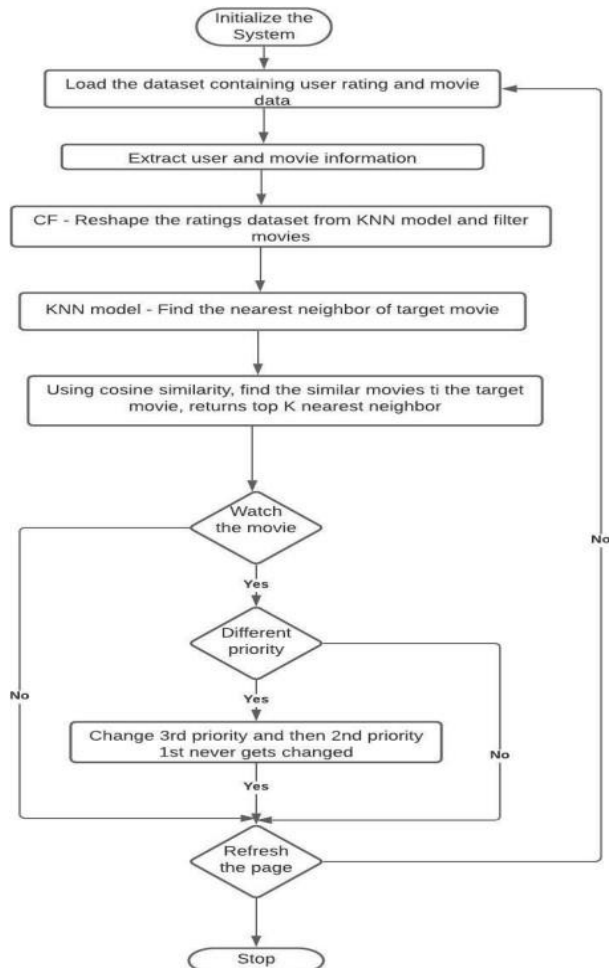
Second, the collaborative filtering begins with the formatting of the rating dataset in order that it is often consumed by the KNN model, to get rid of the large dataset handling problems. The dataset is reduced consistent with the recognition removing the noisy error pattern to urge the sparse matrix. pattern to urge the sparse matrix. Cosine similarity is employed to seek out the space between the target movie and other movies, which provides us the highest k nearest neighbor. and eventually displaying the specified recommended list of flicks with descending order of distance.

K-Nearest Neighbor (KNN) Algorithm

In KNN algorithm, all the data is converted into vector points and so a scattered plot is obtained. All the dataset have a label assigned to them which are known. A value for K is choosen which is the nearest data point. For solving the predictive problem KNN uses feature similarity and checks how closely the data matches the target data. With help of Euclidean distance formula the distance between the target point and

potential data is measured. Based on distance value data is sorted accordingly. Expecting that comparative things exists in closeness KNN calculation suggests the closest information as the outcome. In the KNN algorithm, if the worth of K=1, then the case is assigned to its nearest neighbor of that class. A case in KNN is assessed by the foremost majority vote of its neighbors, where the case is being allocated to the category commonest among its nearest neighbors measured by a distance function.
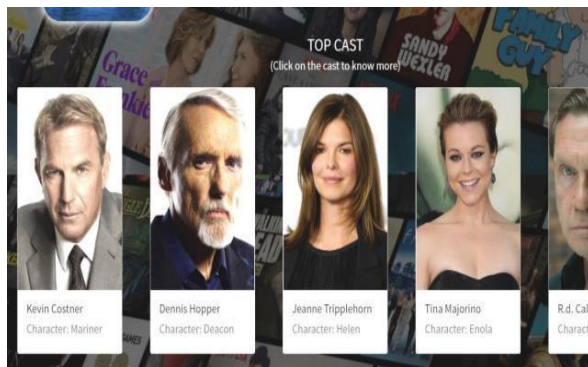
IV.   RESULTS

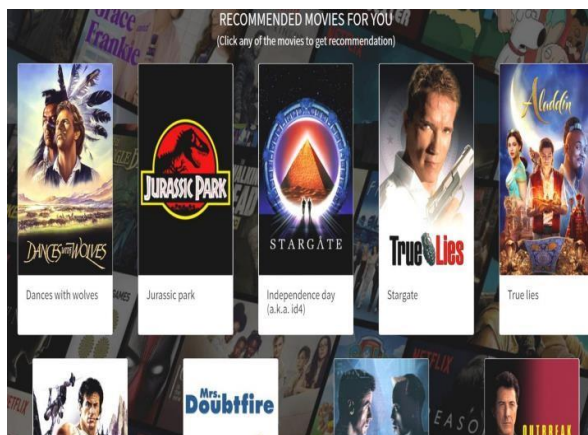Here is output of how the Movie Recommendation System website looks
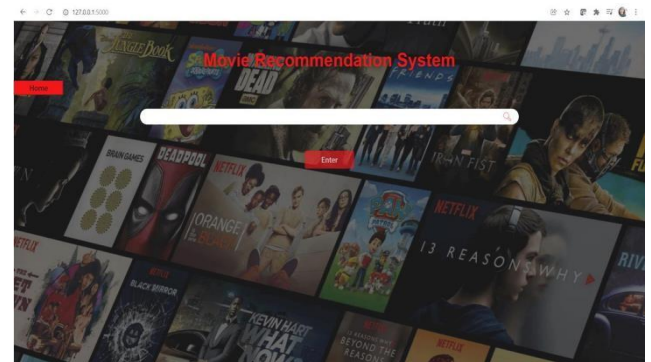
In this case the user has searched for 'Waterworld' movie so the details of the movie searched have been displayed.
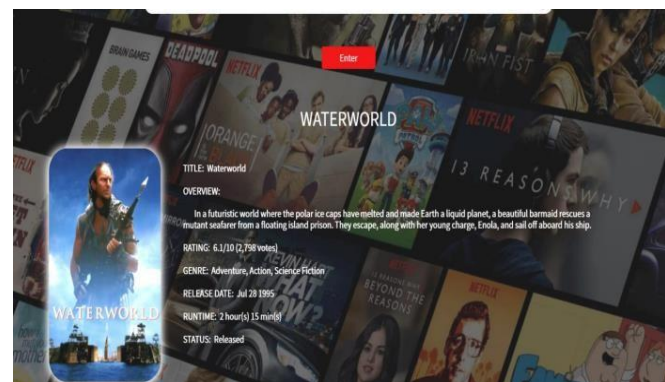
Movies having feature similarity i.e being closely related to the searched movie have been recommended as output based on cosine similarity. Thus the highest rated movies like Jurassic park are suggested to the current user with help of collaborative filtering approach.

V. CONCLUSIONS

We introduce a movie recommendation system based on a mix of opinion mining and user likeness analysis in this study. It allows a user to choose from a variety of attributes and then recommends a movie list to him based on the cumulative weight of the various attributes and the K-means algorithm. Finally, top-k movie is recommended for the targeted users. Our proposed system suggests leading and blockbuster movies, and the system is beneficial to millions and billions of consumers throughout the world. The accuracy of classification is increased here by employing the NbSVM classifier, which also meets the consumers' needs. Our suggested system has been tested on the Movie Lens Dataset, and our results demonstrate that it outperforms ALS and SEHRS.

REFERENCES

[1] K. Shah, A. Salunke, S. Dongare and K. Antala, "Recommender systems: An overview of different approaches to recommendations," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017, pp. 1-4, doi: 10.1109/ICIIECS.2017.8276172.

[2] A. Fanca, A. Puscasiu, D. -I. Gota and H. Valean, "Recommendation Systems with Machine Learning," 2020 21th International Carpathian Control Conference (ICCC), 2020, pp. 1-6, doi: 10.1109/ICCC49264.2020.9257290.

[3] Ali S.M., Nayak G.K., Lenka R.K., Barik R.K. (2018) Movie Recommendation System Using Genome Tags and Content-Based Filtering. In: Kolhe M., Trivedi M., Tiwari S., Singh V. (eds) Advances in Data and Information Sciences. Lecture Notes in Networks and Systems, vol 38. Springer, Singapore.

[4] A. Pujahari and V. Padmanabhan, "An Approach to Content Based Recommender Systems Using Decision List Based Classification with k-DNF Rule Set," 2014 International Conference on Information Technology, 2014, pp. 260-263, doi: 10.1109/ICIT.2014.13.

[5] Pradeep, N., Rao Mangalore, K. K., Rajpal, B., Prasad, N., & Shastri, R. (2020). Content based movie recommendation system. International journal of research in industrial engineering, 9(4), 337-348.

[6] Da'u, A., Salim, N. Recommendation system based on deep learning methods: a systematic review and new directions. Artif Intell Rev 53, 2709–2748 (2020). https://doi.org/10.1007/s10462-019-09744-1

[7] Furtado, A, Singh "Movie recommendation system using machine learning" Department of Master of Application, Jain University, Knowledge Campus, Bangalore, Karnataka, India. Int. J. Res. Ind. Eng. Vol. 9, No. 1 (2020) 84–98

[8] M. Gupta, A. Thakkar, Aashish, V. Gupta and D. P. S. Rathore, "Movie Recommender System Using Collaborative Filtering," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 415-420, doi: 10.1109/ICESC48915.2020.9155879.

[9] [1]. Zhang, J.; Wang, Y.; Yuan, Z.; Jin, Q.; "Personalized Real-Time Movie Recommendation System: Practical Prototype and Evaluation", Tsinghua Science And Technology, vol: 25, 2020, pp: 180-191

[10] [2]. Zhang, R.; Mao, Y.; "Movie Recommendation via Markovian Factorization of Matrix Processes", IEEE, vol: 7, 2019, pp: 13189-13199

[11] [3]. Haruna, K.; Ismail, M.A.; Suyanto, M; Gabralla, L.A.; Bichi, A.B.; Danjuma, S.; Kakudi, H.A.; Haruna, M.S.; Zerdoumi, S.; Abawajy, J.H.; Herawan, T.; "A Soft Set Approach for Handling Conflict Situation on Movie Selection", IEEE, vol: 7, 2019, pp: 116179-116194

[12] [4]. Kalra, N.; Yadav, D.; Bathla, G.; "SynRec: A Prediction Technique using Collaborative Filtering and Synergy Score", International Journal of Engineering and Advanced Technology, vol: 8, 2019, pp: 457-463

[13] [5]. Pavithra, M.; Sowmiya, S.; Tamilmalar, A.; Raguvaran, S.; "Searching an Optimal Algorithm for Movie Recommendation System", International Research Journal of Engineering and Technology, vol: 6, 2019, pp: 216-221