USER-RATED MOVIE RECOMMENDATION SYSTEM USING KNN ALGORITHM

R. Jeeva, N. Gomathi, C. Rajeshwari, M. Sakthi Manisha, P. Tamilselvi

R. Jeeva, Research Scholar, AP / CSE Department, Thamirabharani Engineering College, Tirunelveli, Tamilnadu, India, jeeva3710@gmail.com
Final Year Student, Thamirabharani Engineering College, Tirunelveli,
gomsnehru2001@gmail.com, santhiyaraji95@gmail.com, manishasakthi5@gmail.com, tamilpackiya29@gmail.com

ABSTRACT:

The recommendation system analyzes the past preferences of the user concerned and then it uses this information to try to find similar movies. This information is available in the database (e.g., lead actors, director, genre, etc.). After that, the system provides movie recommendations for the user. It helps users to find the movies of their choices based on the movie experience of other users in efficient and effective manner without wasting much time in useless browsing. The recommended movie list is sorted according to the ratings given to these movies by previous users and it uses KNN (K-Nearest Neighbor) Algorithm, where K represents the number of nearest neighbors.

Index Terms— Recommendation system, Accuracy, Machine learning, collaborative filtering, k-nearest neighbors.

INTRODUCTION:

A Recommendation system is a smart system that can suggest items or products based on the previous information or data of the individual user. Currently, the recommendation system is used in all the major search engine websites or any e-commerce giants. The user almost gets overrun with product or item recommendation in these kinds of businesses. Majority of them however use the individuals past/historic data or preliminary matches to recommend. There are multiple studies that have shown that recommendation system proves to give colossal value to the businesses. Although recommendation systems are the background source of these businesses. It takes a few machine learning techniques and algorithms to build one. ”A recommendation system can be largely described into three categories: collaborative filtering systems, content-based systems and hybrid systems”. [1] Collaborative filtering is dependent on the relation between users and its items. However, bare bone collaborative filtering will not be a suitable choice for recommendations of movies. Rating a movie or a choice of music is a very profound personal experience of an individual, it cannot be generic like a grocery product. Their mood, experience, stage of life can all be an important factor or attribute for any kind of recommendation. Hence, the logic of suggesting an individual with a movie based on his/her previous choices can be a challenging effort. Recommending a movie to a person whom you know personally can be comparatively a very easy task as a few attributes such as character, age, personality, trait and movie taste is already known. Dismally, the movie recommendation system in unaware of such anomalies. Hence the personalised decision making can be quite a difficult task. To produce accurate results of recommendation data from multiple users have to be considered and used. Especially focusing on how to utilize the user’s particular movie rating data to predict the users next selection more effectively. To achieve this, the Under the supervision of Dr. Shagufta Henna, Letterkenny Institute of Technology, Letterkenny, CO. Donegal recommendation system that has been implemented here uses collaborative filtering which uses rating data in the form of matrix where the movies are represented in columns and users in rows and implementing K-nearest neighbours’ model to measure the similarity of the user using distance metrics to fill in the blanks for ungiven rating and then come up with a recommendation. [1]

LITERATURE SURVEY:

A. Content-based Filtering:

The recommendation system relies on the

likeness of the products. The overall idea is if you like a particular product then it is highly considered or accepted that you will also like a similar product. It works well when the property or the characteristics of a product can be determined easily. In terms of Movie recommendation system, A content-based recommendation system can recommend a movie based on user data provided by them explicitly, following which a user profile is generated. This profile is further used to make suggestions which becomes more accurate overtime. In content-based system the ”concepts of Term Frequency and Inverse document frequency are used for filtering systems and information retrieval”. [2] The primary use of these terms is to derive the importance of any movie. Term frequency can be described as the frequency or the number of times the word in a document. Whereas, inverse document frequency is the whole collection of documents. For Instance, suppose in a search engine we type ‘The result of English Premier League’. It is quite assertive that occurrence of the word ‘The’ will be more frequent than ‘English premier league’ for any other search queries. Therefore, the importance of ‘English premier league’ is very high with regards to ‘The’. In such cases term frequency and inverse document frequency weighting is used to determine the importance. [2]

B. Memory-based Collaborative Filtering:

The memory based collaborative filtering is totally based on previous behaviour and not on the context. It is generally considering the similarity choices and preferences of two users. It gives recommendations to a user based on the preferences of a similar user. For example, if an individual likes a movie A, X and Z and another individual likes a movie X, Z and U. Then it is assumed that the first individual is more likely to watch the movie U. This is also one of the most often used algorithm because it does not have to depend on any extra information. It is a unique property of the algorithm that allows it to feature learn on its own. [2] The two major types of memory-based collaborative filtering are 1) User-User collaborative filtering: The major criteria here is to find look alike or similar users based on their past selections. The algorithm is time and resource consuming but very effective too. Therefore, it requires and very strong parallel computing system because the algorithm is hard to implement with huge base platform. 2) Item-item collaborative filtering: This type of filtering mechanism is quite like the previous algorithm. But, instead of focusing on user similarity this targets the movie. It is used to recommend similar movies to user who have rated he movies from the dataset. This algorithm as compared to the previous one barely uses any resource and time and a correlated movie matrix is framed over the period. However, in either scenario there is a similarity matrix that is getting built which consists of a few distance metrics that is used for measuring the similarity. Firstly, Jaccard similarity in which the similarity is based on the amount of users who have rated the movies. It is applied to the scenarios when there is no available numeric rating but rather just available Boolean value. Secondly, we have cosine similarity which considers the closest vectors with smaller angles and larger cosines. Lastly, there is Pearson similarity which is considered the coefficient among two vectors. [2]

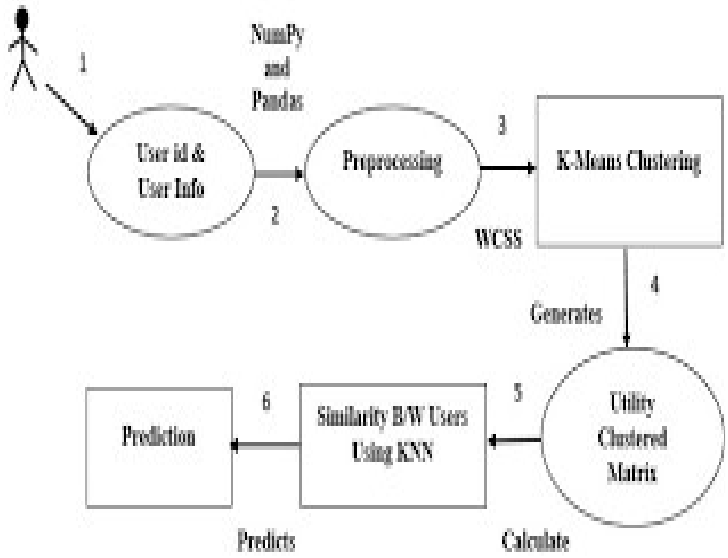
C. Model-based collaborative filtering:

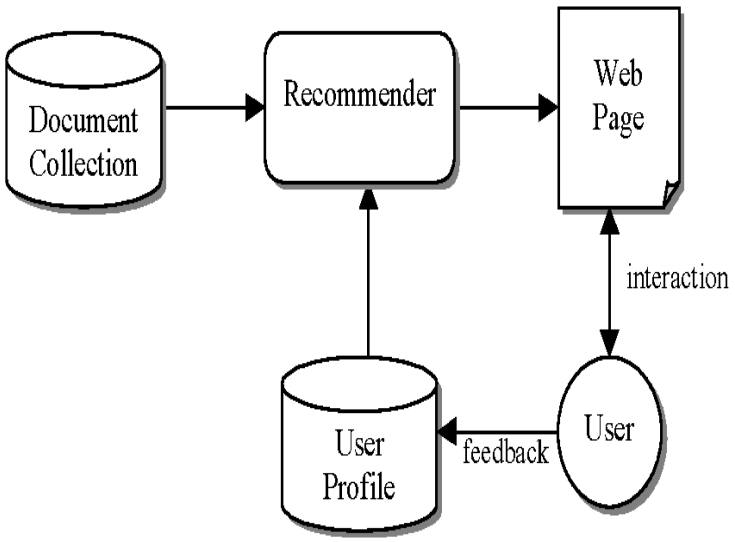
As we saw in the previous model that how users rating data can be used to recommend movies. However, using that can give high root mean squared which later on gives bad recommendations when the data set becomes enormous. The model-based collaborative filtering is based on a unsupervised learning method where it is exposed to Dimensionality reduction and latent variable decomposition. This model is used for the recommendation systems where it can deal with sparsity and scalability. The major target here is to learn the characteristics of the ratings and features such as user preference and item attributes. To predict the blank ratings with the help of dot product of its items and users. When given a sparse matrix with multiple dimensions the model can be restructured with the help of matrix factorization. Well, on a higher level it uses singular value decomposition in order to decompose into diagonal matrix and double unitary matrix. [2]

PROPOSED SYSTEM:

- It collects information have intensively search free online movie database and extract the information's.
- Predicting accurate recommend movies user have applied K-means clustering algorithm along with a pre filter.
- The usage of attributes that give priority for recommendation over movie database.

SYSTEM ARCHITECTURE:





CONCLUSION:

This recommendation system with the combination of many evaluation metrics makes itself relevant and useful. This paper focused on the idea behind matrix factorisation and implementation on K nearest neighbor and finding the "N" recommendations. This is a more personalised approach done via model based collaborative filtering The performance of the model

however has the scope of improvement. Additional factors from the dataset which may provide more insights can be considered in the future. Nevertheless, identifying correlations between customer or a user satisfaction to give recommendations with various different factors can also be a way in the future.

REFERENCES:

1. VIVAT THONGCHOTCHAT, YASUO KUDO, YOSHIFUMI OKADA, AND KAZUHIKO SATO, "Educational Recommendation System Utilizing Learning Styles: A Systematic Literature Review", 31 January 2023.
2. ABEER ALJOHANI, MOHAMED ALI RAKROUKI NAWAF ALHARBE AND REYADH ALLUHAIBI, "A Self-Attention Mask Learning-Based Recommendation System", 9 September 2022.
3. BO JIANG, JUNCHEN YANG, YANBIN QIN, TIAN WANG, MUCHOU WANG, AND WEIFENG PAN, "A Service Recommendation Algorithm Based on Knowledge Graph and Collaborative Filtering", 8 April 2021.