Movie Recommendation System Using Machine Learning

Dr. Prasant Dixit
Professor, Department of Computer Science and Engineering
Galgotias University
Greater Noida, India

Yash Raj
Student, Department of Computer Science and Engineering
Galgotias University
Greater Noida, India

Arun Kumar
Student, Department of Computer Science and Engineering
Galgotias University
Greater Noida, India

Abstract-- The recommendation system of today helps users locate the items which they need most. Movie The purpose of recommendation systems is to guide movie enthusiasts toward selecting their next movie view A large number of movies that might take considerable time to select can be bypassed through this system. The process requires substantial time along with causing confusion from several thousands or millions of options. This paper seeks to decrease the need for human involvement. A suggestion system cuts down human work by recommending movies that match user preferences. To handle such problems, we the proposed model united elements from both content-based and collaborative filtering systems. It will give This system delivers better outcome clarity when compared to content-based systems established previously. approach. Content-based recommendation systems function for people only and therefore they limit users to specific choices. The prescriptions from this system exist outside common recommendations which prevents users from expanding their discovery potential. Hence, we have focused Our system aims to address and solve the identified problems.

I. INTRODUCTION

Millions of new movies appear on different platforms and territories daily so users now face substantial challenges to discover appropriate films. Watching movies plays an essential role in entertainment as people use it to enjoy their free time or spend quality screen time with family members. Making a decision regarding a movie from among the numerous available choices causes significant difficulty for viewers. The job of recommender systems as intelligent platforms consists of scanning extensive movie databases to identify contents suitable for users based on their unique preferences [1–4]. Service providers gain advantages from recommender systems which complement the numerous benefits they provide to users. Such systems deliver precise user-tailored recommendations through which users discover items that match their proven preferences. Recommender systems stretch beyond serving movie recommendations and extend to suggest music selections along with advertisements and grocery merchandise and news items along with any multiple-item lists. Systems utilize different sets of techniques which work to evaluate and rate content through analysing user experiences and engagement data. User feedback and user behavioural data enables the platform to suggest content that matches individual user preferences [5–8]. The result? The system reduces user search activities and enhances buying opportunities and maintains customer loyalty through continued user delight. Research indicates that platforms using recommender systems gain more users who are willing to engage with their content primarily within premium and luxury service segments. These systems acquire learning capability through multiple varieties of data including: a) user actions and preferences, and b) customer reviews and feedback, The system determines highly personalized recommendations through its ability to learn from various types of data.

II. LITERATURE REVIEW

Throughout the movie oracle Nessel discusses how using examples serves as an essential element for human interaction. Human connection functions properly through this method together with the system that provides functionality. User behaviour serves as the engine that drives the movie recommendation feature of the system. More computing power is necessary to conduct these operations. System requirements increase proportionally since the analysed texts exceed in size but maintain the same algorithm structure. The algorithms function identically [3] according to this research. To function as a movie recommendation system the algorithm makes use of the system analyses several textual aspects from each movie

such as plot description and cast list and release information and genre information. The system uses year data together with production information to perform extensive analysis. The paper identifies the most suitable options from the pool of available options[2].

Analyses application similarity measure for recommendations forecasting in recommendations systems. The research presents evidence about the computational technique used for similarity assessment. Recommendations systems employ two main similarity methods that include cosine similarity measure and Pearson correlation coefficient [1]. As the User recommendation processing depends on multiple factors which the movie recommendation service accepts. Movie watching history plays an essential role in determining recommendations because content-based features were integrated into the system. recommendation approach. Typically, people have a tendency Many users have the belief that positive customer reviews result in positive effects. negative reviews have negative impact. Sentiment analysis the accuracy of recommendation results will benefit from this method. results. Our experimental results required the proper explanation of this methodology. The problem requires distributed system implementation to achieve its solution. scalability and timeliness of recommender system [5].

III. RELATED WORK

Recommendation systems emerged in the last decades. The various recommendation approaches used by these systems include collaborative approach, content-based approach, utility base approach, hybrid approach and others. Based on shopper purchase histories Lawrence et al. (2001) created a recommender system that provides new market options. The authors enhanced recommendation accuracy by using combination methods from collaborative and content-based filtering systems. Most present-day recommendation systems determine potential customers by evaluating previous-user ratings. Next the system uses these ratings to determine and suggest the preferred item to the user. Weng Lin and Chen completed an evaluation study in 2007 that established multidimensional analysis combined with additional customer profiles enhances recommendation quality. For this purpose, Weng implemented MD recommendation model (multidimensional recommendation model) according to Tuzhilin and Adomavicius's (2001) initial proposal.

IV. RESEARCH METHODOLOGY

1. The Basic K Mean Algorithm

K-means algorithm came to existence through MacQueen's early work the novel algorithm known as ISODATA was developed by Ball and Hall after the original K-means method proposed by MacQueen [6]. The K-means clustering method begins with selecting K preliminary cluster centres. The clustering technique requires K initial centroids to achieve the specified cluster number K The algorithm assigns the point to the cluster which has the nearest average. centroid of the cluster. We move to update the cluster area for each Cluster points belong to the cluster where they receive their assigned values. We Continue the process so the cluster centre remains unchanged. (centroid). The methodology completes its execution by seeking to reach the minimum level of It uses an objective function which includes a squared error function.

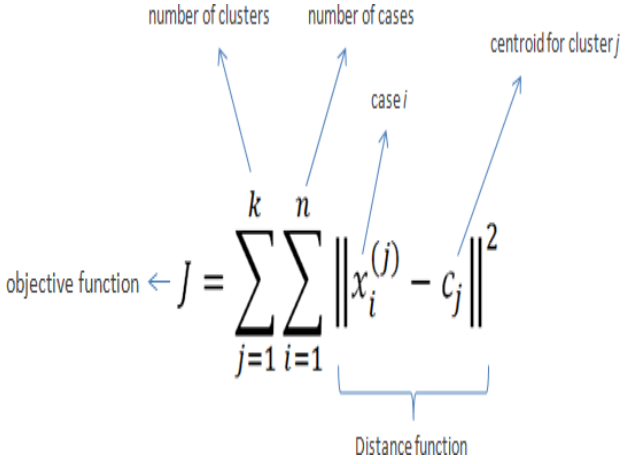

Figure 1: K Mean Algorithm Formula

The equation uses two parameters: k represents cluster number while n signifies case number. The equation uses two parameters: k represents cluster number while n signifies case number

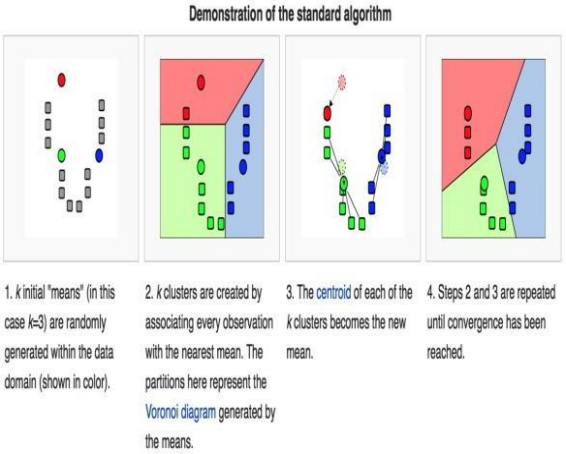

Figure 2: K Means Clustering

2. Data description

The proposed model needs a pre-filter to function before K-means application. At the stage of centroid point calculation, the attributes used for distance assessment include

1. Genre
2. Actor
3. Director
4. Year
5. Rating

Each feature receives a different weight value for calculation purposes. Previous user ratings for movies serve as the foundation for creating the best recommendations so our analysis gives special weight to the rating element above other attributes. The website www.imdb.com provides movie ratings because it holds an extensive movie collection and receives ratings from numerous worldwide users. The proposed model includes total votes received by each specific movie as one of its critical elements. The number of votes was separated into three vote divisions where one group contains fewer than 1000 votes and another hold votes between 1000 and 10000 while the last category has more than 10000. Research evidence shows that higher vote counts should correspond to increased rating weightage in direct proportion. The numbers of vote ratios we use in our research are 1:1, 1:2 and 1:3 according to the total votes a movie receives. Additionally, we have discovered rating lower than 5 indicates movies get least Favor and least suitable recommendations from users. The user desire to watch quality movies prompts us to select prediction sets containing popular films that receive high ratings from a wide range of users. Other attributes receive weighting values derived by calculating the average between their total linked movies and the total number of movies within the dataset.

3. Challenge faced

The major development obstacle for any system arises from the necessity to please end users who receive the system implementation. A number of obstacles confronted our development team while constructing our system. User friendly system design and easy understandability along with straightforward utilization stand as key system requirements. Additionally, we encountered two significant challenges which included developing the most suitable recommended movie list and making the system adaptable for multiple geographic user groups. We also needed to assign weights to system attributes.

V. CONCLUSION

The research presents MovieREC as a recommender system which enables users to select from predefined attributes followed by automated recommendation of movies. Users can choose their preferences from available attributes before receiving recommended movie lists through this system A set of movies receive recommendation to the user via a calculation that combines all selected attributes through the K-means algorithm. The evaluation process of such a system proves difficult since A recommendation without specific merit whatsoever exists between correct and incorrect suggestions since the effectiveness depends merely on user opinions. of opinions. Our system was evaluated through undisclosed assessments we conducted. We performed assessments with limited participants who reacted favourably. them. Our project needs additional data which will enhance its effectiveness. The application delivers more significant outputs when utilized with our system. Additionally, we Our system requires different machine learning and clustering algorithms while examining their relative outcomes. We will study different clustering algorithms for their performance results. We aim to develop a web interface for user interaction which contains a database feature to serve individual user needs. The system operates through an interface that both manages user-related information and implements the learning model tailored to each user.

VI. REFERENCE

[1] Mykhaylo Schwarz, Mykhaylo Lobur, Yuriy Stekh, Analysis of the Effectiveness of Similarity Measures for Recommender Systems, 978-1-5090-5045-1/17/$31.00 ©2017 IEEE M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Rujhan Singla, Saamarth Gupta, Anirudh Gupta, Dinesh Kumar Vishwakarma, FLEX: A Content Based Movie Recommender, 978-1-7281-6221-8/20/$31.00 ©2020 IEEE. Jochen Nessel, Barbara Cimpa, The MovieOracle - Content Based Movie Recommendations, 978-0-7695-4513-4/11 $26.00 © 2011 IEEE.

[3] Shreya Agrawal, Pooja Jain, An Improved Approach for Movie Recommendation System, 978-1-5090-3243- 3/17/$31.00 ©2017 IEEE

[4] Yibo Wang, Mingming Wang, and Wei Xu, A Sentiment-Enhanced Hybrid Recommender System for Movie Recommendation: A Big Data Analytics Framework, Hindawi Wireless Communications and Mobile Computing Volume 2018, Article ID 8263704

[5] F. Furtado, A, Singh, Movie Recommendation System Using Machine Learning, Int. J. Res. Ind. Eng. Vol. 9, No. 1

[6] MacQueen J.. Some methods for classification and analysis of multivariate observations. In Proc. Of the 5th Berkeley Symp. On Mathematical Statistics and Probability, pages 281-297. University of California Press, 1967.