IOSR Journal of Engineering (IOSRJEN) www.iosrjen.org
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 11, Issue 3, March 2021, ||Series -II|| PP 43-46

Implementation of a Movie Recommendation System for Various Online Streaming Services: A Review

Mahesh Babasaheb Dahiwal[1], Prachi Avinash Joshi[2]
Department of Information Technology, Government Polytechnic, Aurangabad
Dr. JayshriDamodar Pagare[3]
Department of Computer Science and Engineering, Jawaharlal Nehru Engineering college, Aurangabad
Received 22 March 2021; Accepted 06 April 2021

Abstract
There are millions/billions of products on every entertainment website or online shop. It becomes difficult for the consumer to choose the best option. Recommender systems enter the picture at this point and assist the user in finding the right item by reducing the number of choices. Recommendation systems assist users in selecting the appropriate item by presenting a likely list of options; as a result, they have become an integral part of e-commerce, movie and music streaming sites, and so on. They are rapidly becoming one of the most commonly used applications of machine learning, which has increased in popularity in recent years. In this paper we are reviewing various recent implementations of Movie recommendation systems for various Online Streaming Services. These models have achieved more accuracy compared to traditional Collaborative filtering approach but Sentiment Analysis and user's feedback can be used to further improvement in accuracy.
Keywords: Collaborative filtering, Content-based filtering, Movie recommendation system, Neural Network, rating prediction

I. INTRODUCTION

When one wants to see a movie, he/she often gets confused about what movie to see. At this point, he/she may ask the people around him/her for advice. Currently, society produces huge amounts of information every day. Faced with so much information, many people do not know what information is needed (Shen et al., 2020).

With ever-increasing, a vast collection of movies, it would be tedious for customers to narrow down movies that would best satisfy their needs unless they know exactly which titles they would like to watch (Pongpaichet et al., 2020). The increasing interest and competition in the online streaming market have become apparent, urging media streaming companies to invent innovative strategies to impress customers and remain relevant in the market space. One of those strategies includes the content recommendation that automatically selects a subset (typically very small) of movies for a user. The purpose of such a functionality is to help the user to quickly and efficiently identify relevant items within the database that contains millions of movie titles (Pongpaichet et al., 2020).

The literature on recommendation systems is extensive. While there are a variety of recommendation algorithms that have been proposed in the past, only a subset was proven applicable in online movie settings.(Pongpaichet et al., 2020)A primary concern of the current recommendation system is how to provide personalized recommendation to users and improve the accuracy and user satisfaction. (Cheng et al., 2020)

This paper provides a literature review of recent implementations of Movie Recommendation Systems in online streaming services and explores various factors which needs to consider to further improvement in performance of recommendation systems.

II. TECHNICAL PART

The two most popular approaches for implementation of recommendation systems are:
(1) Content based filtering
(2) Collaborative filtering
(3) Hybrid based approach

A definition of the item and a profile of the user's interests are used to create content-based filtering methods. In contrast, Collaborative filtering is based on the premise that people who have agreed in the past will agree again in the future, and that they will like similar things.

The data generated about the objects is used to create a content-based filtration strategy. The algorithm suggests items that are close to those that the consumer has previously enjoyed. This similarity is calculated using information about the objects as well as the user's previous preferences.For example, if a consumer enjoys

movies like "1920," we can suggest films by "Rajneesh Duggal," films in the "Horror" genre, or even films directed by "Vikram Bhatt." So, the recommendation system looks at the user's previous tastes and finds the film "1920," then tries to find similar movies based on the details in the database, such as the lead actors, the director, the film's genre, the production house, and so on, and finds movies similar to "1920." However, by using content-based filtering, various items are not exposed to the user as much. The user base cannot be extended since the user does not experiment with various types of content. Collaborative filtering techniques can be further divided into two main categories:
(1) user-user based methods and
(2) item-item based methods.
User-user based methods build the similarity matrix between users. A recommendation can then be given to the target user by looking at the set of items that other "similar" users also like. On the other hand, item-item based methods process the collection of items to find the similarity among them. The system then uses this interitem similarity to infer the target user's preferences on an item based on his/her previous preference on other similar items (Pongpaichet et al., 2020).
The Collaborative Filtering approach uses the rating data consisting of two main phases:
(1) movie similarity and
(2) movie rating prediction.
Meanwhile, the Hybrid based approach adds the benefit of a Content-based to the Collaborative Filtering based approach. Thus, it uses both the rating and movie data and is consisting of four main phases:
(1) text pre-processing,
(2) term weighting,
(3) movie clustering, and
(4) Collaborative Filtering based approach.

Empirical results show that the recommendation performances of both approaches are linear to the size of the movie neighbourhood. However, the Hybrid based approach's required neighbourhood size is naturally a lot smaller than that of the Collaborative Filtering since the former employs a clustering technique. The performance comparisons show that the Collaborative Filtering based approach always outperforms the Hybrid based at any top-N position in Precision and NDCG metrics. These findings conjecture that the Hybrid approach does not always improve the Collaborative Filtering approach in movie recommendation (Ifada et al., 2020)

Traditional recommendation techniques, such as collaborative filtering algorithms, are often affected by the sparsity of user-item interactions and cold start issues (Cheng et al., 2020). Traditional collaborative filtering methods do not assume additional information available for each entity type that could enable content-based similarity to be quantitatively measured (Pongpaichet et al., 2020). Execution of collaborative filtering technique on high-dimensional dataset tends to capture high computation time (Sinha et al., 2020).

When a dataset moves from scarcity to abundance, there are two issues with recommendations: fluctuating user interest over time and long computation times.(Sinha et al., 2020).

Also as the dataset shifts from scarcity to abundance, the complexity of understanding user behaviour increases exponentially (Sinha et al., 2020).

For review purpose we have selected papers which have used MovieLens dataset for testing of proposed model.MovieLens is the public dataset provided by the GroupLens research team of Minnesota State University. MovieLens is a kind of movie rating dataset, which is widely used in the research of recommendation algorithm because of its rich dataset, clear and accurate data attributes(Shen et al., 2020).

Normalized Discounted Cumulative Gain or NDCG can be used to measure of ranking quality for recommendation systems(Y. Wang et al., 2013)

III. LITERATURE REVIEW

(Cheng et al., 2020)proposed a movie recommendation model based on Recurrent Neural Network (RNN) and KG-RKAN (Knowledge Graph-Recurrent Knowledge Attention Network), which uses the auxiliary information in the KG to look for the potential interests of users for personalized recommendations. They solve the problem of user's individual interests, by designing an attention module in RKAN, using different weights to converge user's interest. For testing purpose, they mapped data collected from the real movie data set Movielens and IMDB into a new data set for testing. Their model had significantly improved the recommendation accuracy.

(Pongpaichet et al., 2020) proposed rating prediction algorithm using singular value decomposition (SVD). They extended the singular value decomposition (SVD) based movie recommendation algorithm using Paralleled Stochastic Gradient Descent (PSGD) and improved its speed. They compared their proposed algorithm with the state-of-the-art rating prediction algorithm based on the traditional user-user collaborative filtering algorithm on MovieLens dataset and their proposed algorithm outperforms the baseline in terms of accuracy, in both the rating prediction and movie recommendation tasks.

(W. Wang et al., 2020) proposed a combined recommendation model of LSTM and CNN. Their model combines CNN to fully mine the local information of movie data, and uses LSTM to capture the context of user ratings. They used the MovieLens 1M data set. They compared with the traditional recommendation model and other recommendation models based on deep learning, the combined recommendation model of LSTM and CNN proposed in this paper have a MSE loss reduction of 4.4%~18.7% and a MAE loss reduction of 3.0%~52.2%.

TimeFly algorithm is a novel behavior-inspired recommendation algorithm that operates on the concept of changing the user's behaviour with respect to time. Their proposed model considers two recommendation problems (fluctuating user interest over time and high computation time when datasets go from scarcity to abundance) and shows a real-world implementation of the approach in the field of recommendation engines. On the MovieLens 1M dataset, they compared the results of the TimeFly algorithm with the results of other well-known algorithms. They discovered that using TimeFly results in more accurate predictions in less time. (Sinha et al., 2020)

(Shen et al., 2020) used collaborative filtering algorithm to implement the movie recommendation system. They used the MovieLens data set for experimentation. Their system achieved high efficiency and reliability in large datasets.

The system, which adopts the Hadoop technique, can meet the needs of the big data and the cloud computing environment (Shen et al., 2020).

The KG provides an effective way for the design of recommendation systems in a big data environment. As an emerging type of auxiliary data, it can effectively solve data sparsity and cold start problems, thus improving the accuracy, diversity, and interpretability of recommendation results (Cheng et al., 2020).

IV. CONCLUSION

In this paper we reviewed recent implementation of Movie Recommendation System for Online Streaming Services. In this conclusion let's discuss some findings.

Fluctuating user interest and high computation time are two recommendation problems. Also, when data is inadequate or there is plenty of data, complexity of understanding user behaviour is high.

With Content based filtering approach userbase cannot be expanded if user does not watch different types of movies. Collaborative filtering often affected when users watch time is less and when we cannot draw any inference from user's behaviour. Collaborative filtering does not assume additional information available for each entity type that could enable content-based similarity to be quantitatively measured.

Execution of collaborative filtering technique on high dimensional dataset tends to capture high computation time. Hybrid based approach adds benefit of a content based to the collaborative filtering-based approach.

Recommendation performances of content based filtering and collaborative filtering are linear to the size of the movie neighbourhood. With employment of Clustering technique neighbourhood size required for Hybrid based approach is lot smaller than that of the collaborative filtering. Collaborative Filtering based approach always outperforms the Hybrid based at any top-N position in Precision and NDCG metrics. Hybrid approach does not always improve the Collaborative Filtering approach in movie recommendation.

Movie recommendation model based on Recurrent Neural Network (RNN) and KG-RKAN (Knowledge Graph-Recurrent Knowledge Attention Network (RKAN) solve the problem of user's individual interests, by designing an attention module in RKAN which significantly improved the recommendation accuracy.

When extended the singular value decomposition (SVD) based movie recommendation algorithm using Paralleled Stochastic Gradient Descent (PSGD) outperforms traditional user-user collaborative filtering algorithm in terms of accuracy, in both the rating prediction and movie recommendation tasks.

Combined recommendation model of LSTM and CNN reduced MSE and MAE losses. TimeFly algorithm has high accurate predictions in less computation time and considers solving fluctuating user interest over time and high computation time when data is inadequate or there is plenty of data. With KG (Knowledge Graph) we can achieve better accuracy and can interpret recommendation results in better way.

For high dimension data Hadoop technique can be adopted for decreasing computation time.

Even though all the models reviewed has significant improvements compared to traditional Collaborative filtering approach, there need more work to make use of Sentiment Analysis over various sources for more accurate personalized recommendation. Also, user feedback can be used for reinforcement learning.

REFERENCES

[1]. Cheng, Y., Liu, N., Lu, Y., & Tang, X. (2020). Recurrent knowledge attention network for movie

recommendation. *Proceedings - 2020 3rd International Conference on Electron Device and Mechanical Engineering, ICEDME 2020*, 648–651. https://doi.org/10.1109/ICEDME50972.2020.00153

[2]. Pongpaichet, S., Unprasert, T., Tuarob, S., & Sajjacholapunt, P. (2020). SGD-Rec: A Matrix Decomposition Based Model for Personalized Movie Recommendation. *17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2020*, 588–591. https://doi.org/10.1109/ECTI-CON49241.2020.9158308

[3]. Shen, J., Zhou, T., & Chen, L. (2020). Collaborative filtering-based recommendation system for big data. *International Journal of Computational Science and Engineering*, *21*(2), 219–225. https://doi.org/10.1504/IJCSE.2020.105727

[4]. Sinha, B. B., Dhanalakshmi, R., & Regmi, R. (2020). TimeFly algorithm: a novel behavior-inspired movie recommendation paradigm. *Pattern Analysis and Applications*, *23*(4), 1727–1734. https://doi.org/10.1007/s10044-020-00883-8

[5]. Wang, W., Ye, C., Yang, P., & Miao, Z. (2020). Research on Movie Recommendation Model Based on LSTM and CNN. *Proceedings - 2020 5th International Conference on Computational Intelligence and Applications, ICCIA 2020*, 28–32. https://doi.org/10.1109/ICCIA49625.2020.00013

[6]. Wang, Y., Wang, L., Li, Y., He, D., Chen, W., & Liu, T. Y. (2013). A theoretical analysis of NDCG ranking measures. *Journal of Machine Learning Research*, *30*, 25–54.

Mahesh Babasaheb Dahiwal, et. al. "Implementation of a Movie Recommendation System for Various Online Streaming Services: A Review." *IOSR Journal of Engineering (IOSRJEN)*, 11(03), 2021, pp. 43-46.