International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Prediction Model for Streaming Platform User Recommendation System Based on Collaborative Learning

[1]Nikita Jain Nahar, [2]Deepak Sukheja. [3]L. K Vishwamitra

Submitted: 26/01/2023 **Accepted:** 07/04/2023

Abstract: The rise of content overload on the internet has brought us a new issue for both consumers and service providers. To reduce the quantity of material shown, such as movies, music, or other items, Netflix and Amazon utilise recommender algorithms and user profiling, which try to direct the user through the available information. These technologies amass information about the user in order to provide customised experiences. The majority of current recommender systems use a content-focused approach, yet they often miss the nature of consumers' demands. This paper presents a hybrid approach to improving user profiling of content streaming platforms in order to improve user experience. In this paper, user experience was improved by hybridising the model with a similarity score and by hybridising the collaborative learning of the CNN model with latent based matrix factorization. The result was evaluated based on different latent sizes and also validated using 10-fold cross-validation. The result shows its superiority with respect to existing state-of-the-art models.

Index Terms: *Recommendation System, Collaborative Learning, Deep Learning, User Profile, Content Streaming.*

1. Introduction

Several countries adopted social distancing rules as a result of COVID-19, prompting theatres to limit or even close their doors, encouraging people to stay at home, and supporting the increase of OTT platform subscriptions. As a result, we came to the conclusion that it was time to investigate more OTT platforms. The Internet has more and more information and audiovisual content, which makes it harder to find what you're looking for. This is called cognitive load. Various techniques, such as online directories, online tools, and recommendati 2020 and Iftikhar Alam et.al 2021] on algorithms, have been developed to address this problem. [Bradley K et.al 2000, S. Amara et.al]. Today, users can access a v.ariety of services via a variety of electronic devices (such as smartphones and tablets) [C. I. Eke et.al 2019 and R. Kaur et. Al 2018]. User profiles are becoming increasingly crucial for service providers in this aggressive market to achieve effective service for the virtual representation of the data. User profiling (UP) is a critical and required step as a result. In customised services, users' needs, preferences, and aspirations are matched with service delivery [R. Sharma et.al 2017, M. E. Ibrahim et.al 2019, Shristi Shakya Khanal et.al 2020 and Sha Zhao et.al 2019]. The efficacy of these services is measured by how effectively the service provider comprehends and reflects the user's needs. User profiles, which are representations of persons, are created through the UP process. The user profiling process has two major problems. Creating a user's first profile is one of them, as is routinely updating it to reflect the user's evolving preferences, interests, and needs. Today, a lot of online businesses employ recommendation algorithms to pair customers' preferences with songs, movies, and other media. These tools are designed to help customers identify products that suit their preferences. [10]-[17]. This study looks at how a user profile and recommender system can be used to improve the streaming experience for users on multiple platforms.

A user profile is a set of information and interests about a specific person. It's a clear digital representation of a user's identity in working situations, like operating systems, software programmes, or websites [15, 18]. The user profile assists in identifying a user's characteristics as well as determining the viewer's interaction interests and preferences. The primary purpose of user profiling is to collect information about users and their preferences. Many studies on profiling in the context of recommender systems have been undertaken, and several profiling techniques have been developed over time. In general, data mining and machine learning methodologies have advanced user profiling. Everything started with the knowledge-data discovery paradigm [19]-[24].

The taxonomy of user profiling is shown in Fig. 1, and it contains categories like static and dynamic profiles as well as two different kinds of dynamic profiles: short-term and

[1]*Department of Computer Science and Engineering, Oriental University, Indore, India.*
nikitajain.0320@gmail.com
[2]*Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.*
Deepak_s@vnrvjiet.in
[3]*Department of Computer Science and Engineering, Oriental University, Indore, India.*
lkvishwamitra@gmail.com

long-term user profiles. Additionally mentioned are the stages of user modelling and user procedures.

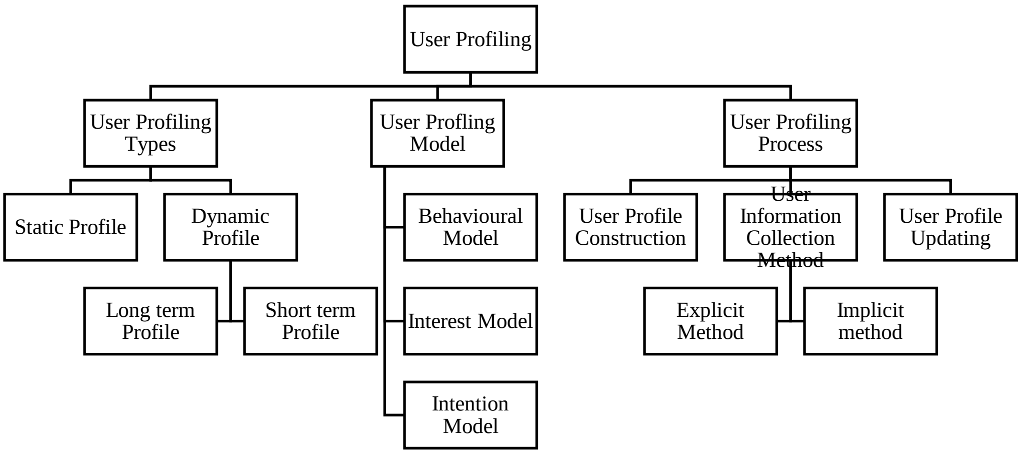


Fig. 1:User Profiling Taxonomy

The increase in social network usage has an impact on both user social interaction and entity profile sharing. Because of the volume of user data transferred across various OTT platforms, protecting individual users' privacy has become a key responsibility. Because of this, personal profiles are being used more and more to make profiles of people. This is a hard problem to solve because, most of the time, there is no good way to tell the difference between a fake profile and a real one [25–32]. An in-depth examination of fraudulent user ID detection methods and strategies will aid in the detection of bogus profile information while also enhancing trust in profile information. An extra security mechanism is required to reinforce the user profile and make it difficult for an intruder to misrepresent it [33],[34]. Basic diagram of a user profiling is shown in Fig 2. A recommender system is needed to forecast which movies, web series, and platforms a person will watch. A succession of recommendations may be made depending on the various user profiling techniques [12][13][41].

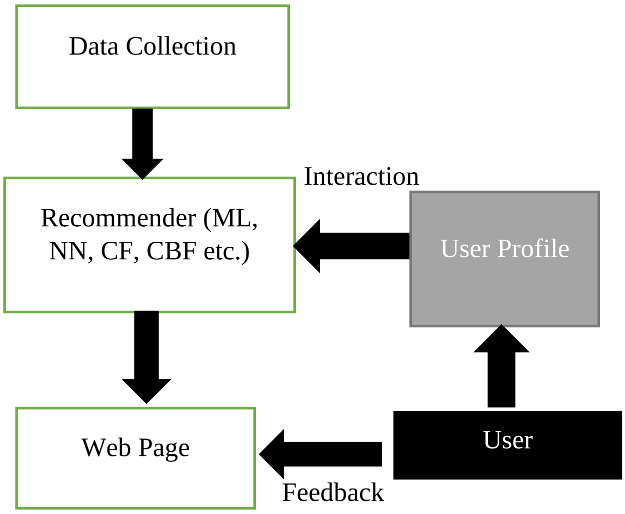


Fig 2: Basic Flowchart of a Recommender

2. Literature Review

Jing et al. [14] proposed an unique hybrid recommendation method is developed, which depends on the idea of combining movies type and viewers recommendation list to determine user similarity. The suggested approach alleviates the issue of data sparsity and changing user interests, and gives more accurate recommendation results than certain current algorithms, according to experimental findings. The end result is as follows: The suggested technique has a lower MAE than any of the previous methods, ranging from 1% to 12%. The training set accounts for 89.350 percent of the total information, whereas the test set accounts for 10.640%.

M. He et al. [15] discussed the goal of online news RS is to combat the data overload of news and provide consumers with individualized suggestions. However, most current approaches employ static representations of user profiles, and there has been limited study into effective user modeling that can dynamically capture user interests in news subjects. On real-world datasets, researchers ran numerous tests comparing the technique to state-of-the-art approaches, and the findings show that the approach greatly increases accuracy and efficacy in news recommendation.

B. Zhu et al. [16] allocated a dependability rating to every prediction and recommendation is crucial in todays modern recommendation system: users should be aware what suggestions are acceptable and which will be dangerous. Model-based work on collaborative filtering dependability have received little attention despite its growing importance. The reliability for every forecasting is calculated using a matrices factorization-dependent architecture and approach described in this work. The reliability values obtained have indeed been tested, and then when employed, they improve predictions and recommendation accuracy in a variety of recommendation system; they also provide a recognizable range of possible values for users. The best experiment results for Proposed technique using Netflix data are 0.229, which has the greatest (negative) correlation values.

Z. Zhu et al. [17] offered a user profile model in this work to explain viewers preference from many angles. Then, based on the user's reading behavior and the popularity of news, explore the degree of the user's preferences and offer technique for calculating the user preferences on various type of news. This strategy might help create more accurate user profiles. Furthermore, they provide a dynamic system for news suggestion that takes into account both type of user preferences. The results of the experiments show that our strategy may increase the recommendation effect greatly.

Sipra et al. [18] suggested a new customized online recommendation system. The web pages are initially transformed into a series of sequences. To determine the neighborhood of each user profile, an effective clustering technique known as clustering is used. The proposed method is then utilize to mine the frequently visited web sites using the association rule mining approach. Finally, the results of the experiments reveal that the suggested technique prioritizes web sites and also suggests similar web pages. Finally, the experimental results reveal that, when compared to current approaches, the suggested strategy enhanced web page suggestion accuracy by 50-60%.

Ouaftouh et al. [19] suggested a model based on collaborative filtering. This method is based on extrapolating a portion of a user's interests from the preferences of different viewers with similar user-profiles. The clustering method is used to provide collaborative filtering among the many ways. Based on a case study, researchers provide a comparison of hierarchical and flat user profile clustering in this paper. The suggested method is based on a collection of user profiles from an e-commerce setting.

Chaurasiya et al. [20] proposed collaborative filtering to create suggestions, a review-based recommendation technique is suggested in this study. Using the benchmark item-based collaborative filtering algorithm, the proposed technique uses review text as user input to create predictions. The raw review dataset, which contains a thorough user response to the product, is pre-processed before sentiment intensity ratings are calculated. When compared rating-based approach, the suggested recommendation model shows an increase in recall and a drop in RMSE. Amazon Videos has a Recall Score of 0.7369, while Automotives has an RMSE Score of 0.8072.

Bandana et al. [21] proposed NB and SVM based Recommendation system. The system model is built using heterogeneous characteristics such as ML features, as well as techniques Naive Bayes and SVM supervised learning method. Based on implementation and observation, the suggested heterogeneous features and hybrid technique may provide a more accurate sentiment analysis system than existing baseline systems. Researchers can utilize these heterogeneous traits to develop advanced and more accurate models utilizing Deep Learning (DL) methods in the future for huge data. The results suggest that using the Naive Bayes algorithm with a data set of 250 training and 100 testing, 89 percent accuracy and precision can be reached.

Luong Vuong et al. [22] presented a novel content-dependent CF method rely on item similarities in this article. Presented a novel method for determining movie similarity based on extracted characteristics such as username , and narrative material. To implement the measurement, researcher split two distinct processes. The first step is to convert all of the movie's retrieved features to vectors. researcher used the Word2Vec embedding approach in particular. In the second stage, researcher used Vector Soft Cosine-based Similarity to measure each feature, and then researcher got a set of movie similarity. The results of the experiments showed that the suggested techniques can successfully manage sparse datasets. The experiment was carried out using actual data from system. The testing findings revealed that the suggested approach consistently outperformed the competition in terms of Accuracy, Precision, Recall, and F1 metrics. The suggested similarity measure will be applied to other kinds of recommendation systems in the future, and the similarity will be measured using other metrics. 82.6360 is the Accuracy metric. 83.44 is the precision measure. 77.300 is the recall measure. 81.027 is the F1 metrics value.

Yi S et al. [23] used machine learning algorithms to learn, analyze, and categorize product and store information based on user experience. Machine learning algorithms have been shown to outperform other methods based on the findings and comparisons. When compared to other current methods, the suggested HRS system has higher MAPE of 96 percent and accuracies of almost 98 percent. The suggested HRS system's mean absolute error is approximately 0.6, indicating that the system's performance is very effective. This method may be expanded in the future to collect consumer interest for various goods in different geographical areas. Some of their contributions are presented in table 1.

Table 1: Major Contributions of Researchers

Ref	Method	Dataset	Discussion
[14]	Hybrid recommendation method	Movie review set	Achieved MAE ranging from 1 % to 12%.
[18]	Fuzzy C-Means	MSNBC dataset	Achieved accuracy of approx. 50-60%. Enhanced web page suggestion
[20]	Product Recommendations Using User Review	Amazon Videos, uses review text	Achieved recall score of 0.7369 and RMSE of 0.8072.
[21]	Naive Bayes (NB) and Linear Support Vector Machine (LSVM).	Product Recommendations Using User Review	Achieved 89% of accuracy.
[22]	Content-based CF approach	Movie's data extracted on different basis such as title, genre etc.	Achieved 63.6% of accuracy and precision of 83.44.
[23]	Machine Learning algorithms	-	Achieved MAPE of 96%, accuracy of 98% and MAE of 0.6. Machine learning algorithms outperform than other approaches.

3. Methodology

User profiling's primary job is to collect information about users and their interests. In the subject of recommender systems, much investigation has been performed on profiling, and numerous profiling approaches have been developed throughout time [43]-[45]. In general, user profiling has progressed as a result of the data mining and machine learning methodology. In our study, we propose a CNN based recommender system using TD-IDF and cosine similarity scores. The broad steps of working are shown in fig. 3.

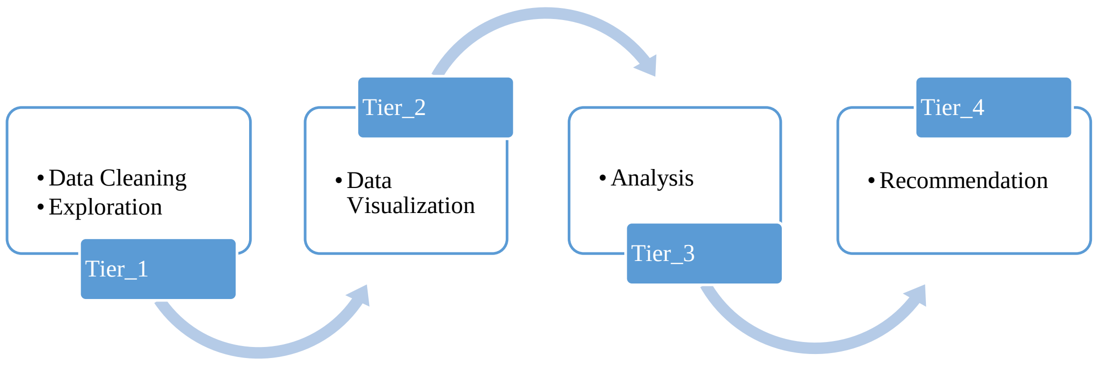


Fig. 3: Steps of Working

Fig 4 shows the more detailed working flowchart which is divided into three main parts data cleaning, data visualization using TD-IDF, Cosine similarity and matrix factorization, further data analysis is performed using collaborative learning to calculate the recommendation output. Each step are discussed in below sections:

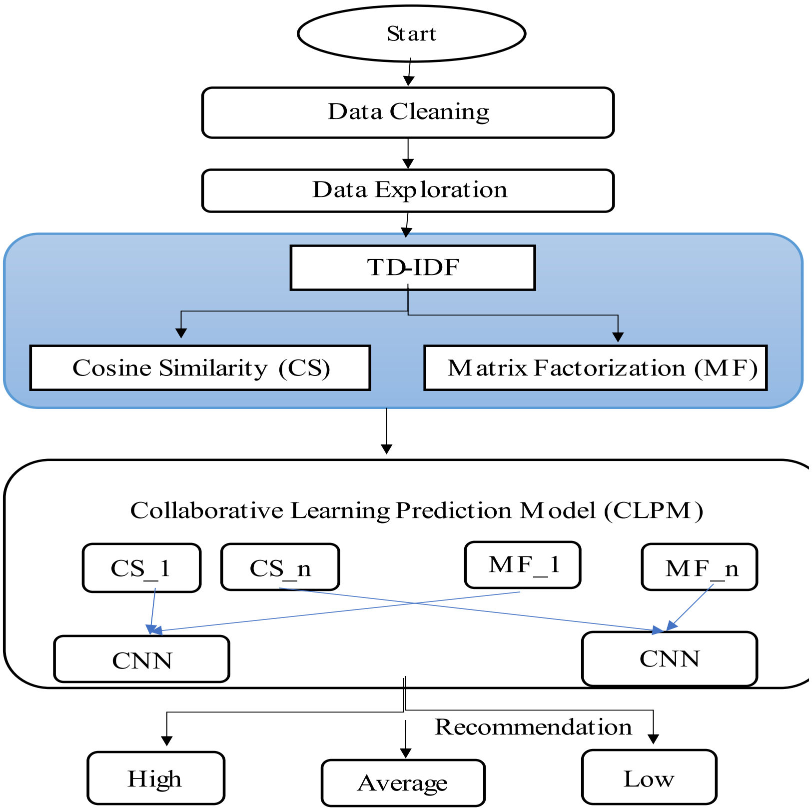


Fig. 4:Flowchart of Methodology

3.1 Tier-1: Data Cleaning

Movie lens [24] dataset was used in this work. For this paper, movielens25M dataset was taken. The methodology will use this information to do exploratory data analysis (EDA) in order to find solutions to some intriguing. For data cleaning, recognizing erroneous, missing, faulty, inappropriate, or missing bits of information and then altering, changing, or removing them as required is part of the data cleaning procedure.

3.2 Tier-2: Data Visualization and Analysis

Fig 5 shows the data visualization steps. It is done in two steps first the term frequency-inverse document frequency (TD-IDF) and then cosine similarity (CS) as well as matrix factorization (MF) is evaluated using three type of scores given as follows:

1. On the basis of user
2. On the basis of movie
3. On the basis of rating

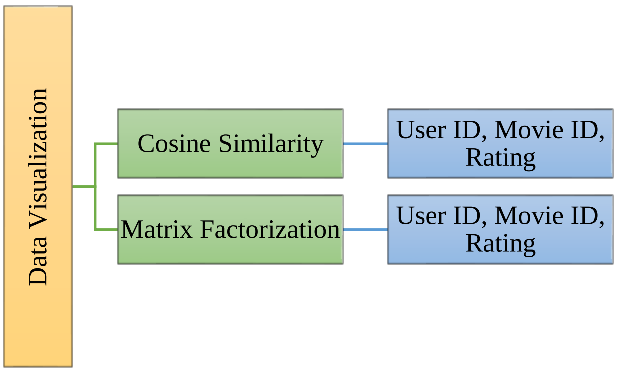


Fig 5: Steps for Data Visualization

3.2.1 Term Frequency- Inverse Document Frequency

TF-IDF is commonly used to determine the statistical probability of word frequency in a data set in a variety of activities, such as textual analysis, data mining, and online databases. It assesses the significance of a given word in a given text. Equation 1 shows how to calculate a TF-IDF score using two words: Term Frequency (TF) and Inverse Document Frequency (IDF).

$$\mathrm{TFIDF}(t, d, D) = \mathrm{TF}(t, d).\mathrm{IDF}(t.D) \quad (1)$$

Where t depicts the word, d depicts each document and D depicts the number of documents.

TF (t, d) is Term Frequency that calculate the number of times a term appears in a dataset . Since each document's duration changes, lengthier texts tend to have higher word frequency. As a result, TF is computed by dividing the frequency of every word in a text by the total words within this dataset. The TF score is determined using eqn (2).

$$\mathrm{TF(t,d)}= \frac{f(t,d)}{|D|} \quad (2)$$

Where, f(t,d) is total number of times t appears in document d and $|d|$ is the total number of times in d "document". IDF (T, D)is termed as Inverse Document Frequency that is used for determining the significance of a word. Numerous terms that occur multiple times in texts, such as stopping words, have little relevance or use. Eqn (3) is used to calculate each corresponding IDF score.

$$\mathrm{IDF(t,d)}= \log \frac{|D|}{|\{D\ I\ t\epsilon d\}|} \quad (3)$$

Where $|D|$ are total number of documents and $\{D\ I\ t\epsilon d\}$ is total number of documents with term t in it. The flowchart of the TF-IDF for a new feature set is shown in Fig 6.

The TF-IDF score of a feature indicates how essential it is for a certain text dataset.

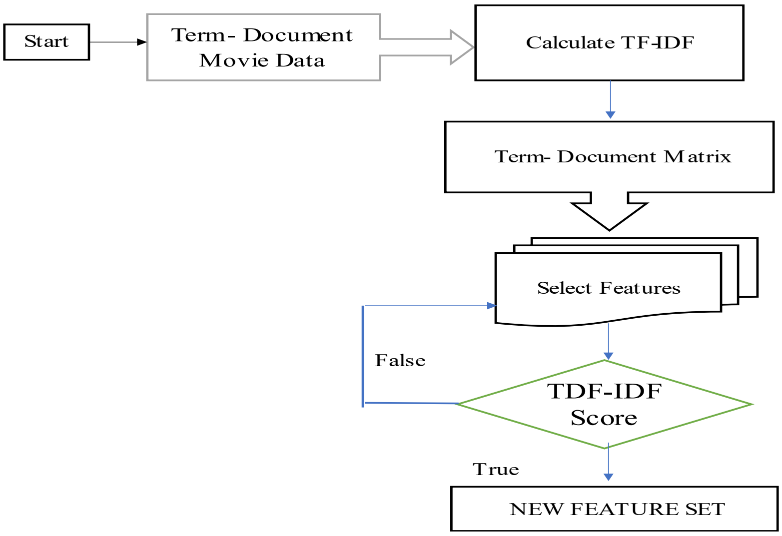


Fig 6: TD-IDF Flow Diagram

3.2.2 Cosine Similarity

Cosine similarity is a prominent Natural Language Processing approach that uses assessments of correlation between two vector (or document) and is frequently used in conjunction with TF-IDF to quantify document uniqueness. The cos of an angle between the two documents is calculated, yielding a value range of -1 to 1, with -1 indicating complete dissimilarity and 1 indicating absolute similarity between two documents. Yet, since the angle between the two documents could not be larger than 90^o as given in eqn (4), cosine similarity will vary from 0 to 1. The below formula is used to calculate cosine similarity:

$$\text{Cosine Similarity}= \frac{A.B}{|A||B|} = \frac{\sum_{I=1}^{N} AiBi}{\sqrt{\sum_{i=1}^{n} Ai^2}\sqrt{\sum_{i=1}^{n} Bi^2}} \quad (4)$$

Where A_i= Component vector of A,

Bi= Component vector of B

3.2.3 Matrix Factorization

Neighborhood modeling, which suggest things based on the similarity measurements, and latent factor modeling techniques, which recommend components depending on the latent factors revealed by the algorithm, are the two most common types of collaborative filtering approaches. In comparison to the preceding class, latent factor approaches have more reliability whenever it comes to managing recommendations, according to numerous studies in literature.

All items or attributes are characterized in matrix factorization as a latent vector. The inner product of x_u and y_i yields an estimating score of interactivities utilizing matrix factorization, presuming x_u and y_i relate to the latent vectors for user u and item I correspondingly:

$$\hat{P}_{ui} = f(u, iI\ xu, yi) = x_u^T yi = \sum_{d=1}^{D} x_{ud}\, y_{id} \quad (5)$$

Where, D is the latent space's dimension. After matrix factorization, each element or feature is considered as an element in latent vector and are independent of other features in latent vector.

In lower dimension of latent vector, it is quite difficult to identify the proper features interactions. Whereas in higher dimension of latent vector, overfitting issue arises. Therefore, in this work, to reduce the overfitting issue latent vectors are fed into CNN model for further accurate learning and for better recommendation.

3.3 Tier-3: Analysis

Last step is recommendation using a recommender. Recommender systems have gained popularity in recent years as a possible answer to the issue of information overload. Recommender systems are a kind of customized information filtering technology that helps users find things of interest by making suggestions. They've been used to enhance the quality of web services in a variety of applications with great success. Recommender systems employ a variety of profiling approaches to gather data about user interactions, which they then incorporate into user profiles. In this research work, we used a Matrix Factorization with CNN based Recommender system. The description of CNN model is presented as below:

CNNs was initially designed for computer vision but has since found widespread Uses in voice and text analysis. In this paper, the latent vector is collaborated with CNN model and termed as collaborative learning prediction model (CLPM). Collaborated data is fed into the Convolution Neural Network to understand the movie interaction with user with a high degree of flexibility and non-linearity. Different layers of CNN are as follows:

Input Layer: Two feature vectors defining user u and movie item I were utilized as input in the first layer, which is the input layer. These are used for collaborative learning.

Collaborative Learning Layer: The CS matrix, x_u and MF, y_i matrix are fed as input in this layer. The interaction vector, z^{cnn}, was created by concatenating the two vectors. Mathematically it is represented as:

$$z^{cnn} = \varphi 1(xu, yi) = \begin{bmatrix} xu \\ yi \end{bmatrix} \tag{6}$$

Convolution Layer: Contextual information were extracted using the convolution layer. By analyzing the interaction vector, z^{cnn}, the convolution structure was used to assess the contextual aspects among the input vectors.

$$c_n^m = f(w_c^m * z^{cnn}_{(:n:(n+w-1))} + b_c^m) \tag{7}$$

Where, f is an activation function, * is a convolution operator. Leaky Rectifier Linear Unit (LReLU) is adopted as an activation function as it avoids the vanishing gradient issue. The feature latent vector might be generated as follows:

$$c^m = [c_1^m, c_2^m, c_3^m \dots \dots \dots c_{1n}^m, c_{1k-w+1}^m] \tag{8}$$

Pooling Layer: This layer pulls relevant features from the convolutional layer, in which an interaction is described by j contextual FV. There may be many contextual elements in Eq. 9, and the majority of them may be useless. As a result, max-pooling was used to extract just the most important contextual feature from each contextual FV.

$$z_f^{cnn} = [\max(c^1), \max(c^2) \dots \dots \max(c^j)] \tag{9}$$

The contextual feature vector is retrieved by the m_{th} shared weight. For additional layers, the vector is created as in eqn. (10) that is evaluated as:

$$z_1^{cnn} = \varphi 1(xu, yi) \tag{10}$$

$$z_2^{cnn} = \varphi 2(z_f^{cnn})$$

$$z_f^{cnn} = \varphi f(z_f^{cnn}) = cnn\left(w, \begin{bmatrix} xu \\ yi \end{bmatrix}\right),$$

To avoid clutter, W signifies all weight and bias variables, while the matrix formed by x_u and y_i represents the original interaction vector.

Output Layer: This layer estimates the probability score of final prediction after analyzing the features from the preceding layer. It is mathematically represented as:

$$\hat{P}_{ui} = \sigma[h^T z_f^{cnn}] \tag{11}$$

In which σ is the sigmoid function given by $\sigma(a) = \frac{1}{1+e^{-a}}$ is the weights of the output layer.

3.4 Tier-4: Recommendation

Traditional collaborative filtering provides suggestions based on the idea of individuals who are similar to one another, whereas content-based collaborative filtering generates recommendations by using the browsing histories of users. These two strategies are plagued by a number of drawbacks, which discourage new users from following the recommendation. The hybrid strategy is used in this study to attain improved performance by combining one or more recommendation strategies with other procedures. In light of this, the article suggested a hybrid approach as a solution to the problem of new users or new things. The hybrid filtering strategy can be broken down into three distinct phases as follows: In the first stage, it examines the profiles of other users in order to discover other individuals who are comparable to the active user. The ratings provided by users are what the cosine similarity algorithm uses to determine which people are most like one another. Following the completion of this calculation, the top-n most comparable users, also known as neighbours or friends, are found. In the second phase, it determines which item should be considered a contender for each neighbour by first getting vectors V_c and V_m that correspond to the user profile and the item contents, respectively. We are able to calculate the distance between these vectors by making use of the cosine similarity equation. After determining the prediction value for each item, the system moves on to the third phase, which is where it makes its recommendations to the user who is being targeted. In this part of the article, the recommendation is broken down into three levels: low, average, and high.

4. Results and Discussions

4.1 Data set Description

To verify the performance of the designed model, real MovieLens datasets: MovieLens 2M; are collected as part of the GroupLens Research Project of the University of Minnesota (and are available online at https://grouplens.org/datasets/movielens/). Rating data sets from MovieLens [24] have been collected and made available by GroupLens Research. Depending on the size of the data set, different periods were used to collect it. MovieLens has a total of 25 million movie reviews. Benchmark dataset that is stable. 162,000 users applied 25 million ratings and one million tag applications to 62,000 movies, resulting in a total of 25 million ratings and one million tag applications. Tag genome data with 15 million relevance scores spread across 1,129 tags is included. Fig 7 represents the rating probability distribution of the dataset. The probability distribution is maximum at 3 ratings.

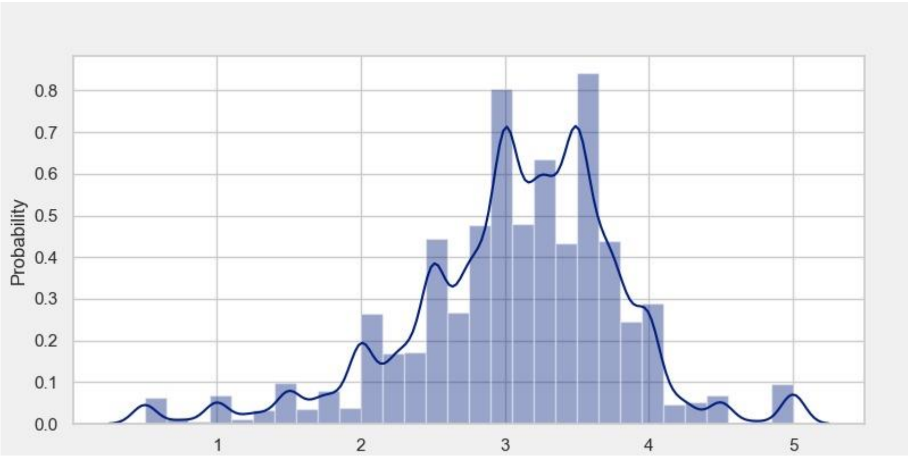


Fig. 7: Probability vs Ratings

4.2 Performance Parameters

For performance evaluation two parameters MSE and RMSE is calculated given below:

MSE: MSE is mean square error, it is the average of the squared difference between actual and estimated value. MSE is calculated using the eqn (12):

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{12}$$

Where, MSE= mean square error, n = the number of data points, Y_i = observed values

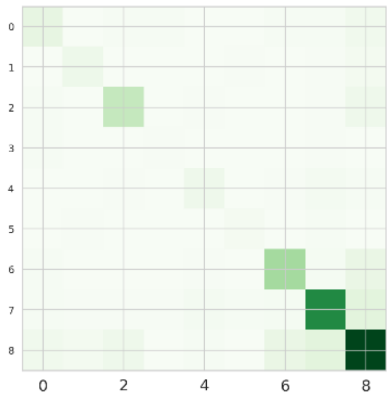
RMSE: RMSE is root mean square error. It is the square root of the average difference between the actual and observed values.

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)}}{\sqrt{N}} \tag{13}$$

Where, RMSE = the root mean square error, N = the number of values, Yi = the actual values and the other is the calculated values

4.3 Performance Evaluation

In this paper, we have simulated the model on python platform and evaluated the output. The Simulation was performed on system having configuration of i5 processor with 16GB RAM and 1TB SSD. The training of the model was performed with variable number of latent size and 10-fold validation. We have created three CNN models with latent size of 25 for CNN layer, latent size of 50 for CNN layer, latent size of 75 for CNN layer and latent size of 100 for CNN layer. Before, training in CNN model, we have generated the cosine similarity and weight factorization graph. In Fig 8 (a), we have represented the cosine similarity scores for movies. Fig 8(b) is used to represent the recommendation result using cosine similarity score in three categories “LOW”, “AVERAGE” and “HIGH”.



(a) Cosine Similarity

movieId	Recommendation	score	Rank	user_id
50	High Recommended	397	1.0	628
181	High Recommended	376	2.0	628
100	Average Recommended	349	4.0	628
286	Average Recommended	344	5.0	628
294	Low Recommended	324	7.0	628

(b) Recommendation Result

Fig. 8: Sample of Cosine Similarity Score and Recommendation

The learning progress is presented in fig 9. The graph presented in fig 9 shows the training and validation loss graph. Fig 10 shows the MSE comparison of different latent size on CNN model names as CLPM25, CLPM50, CLPM100 and CLPM500. For CLPM100 MSE has highest error with the value of 0.898 and minimum for CNN500 of value 0.888. Similarly, the RMSE value is represented in Fig 11. For CLPM100 RMSE has highest error with the value of 0.947 and minimum for CNN500 of value 0.942. Further, the analysis was performed on 5-fold validation, that was performed on CLPM500 and achieved MSE of 0.886 and RMSE of 0.9412.

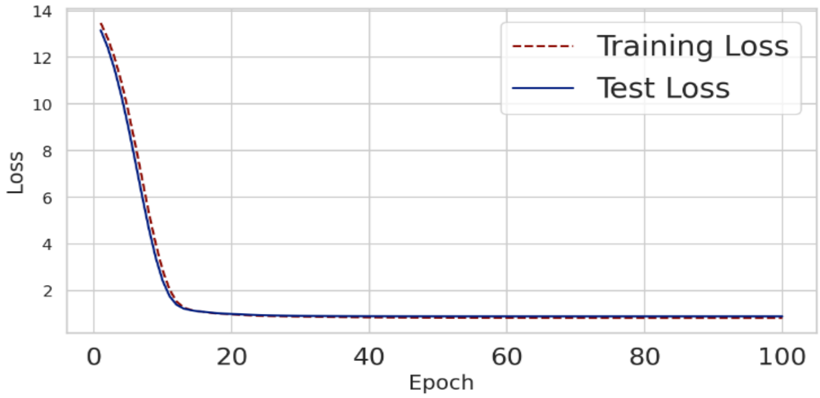


Fig. 9: Training and Validation Loss

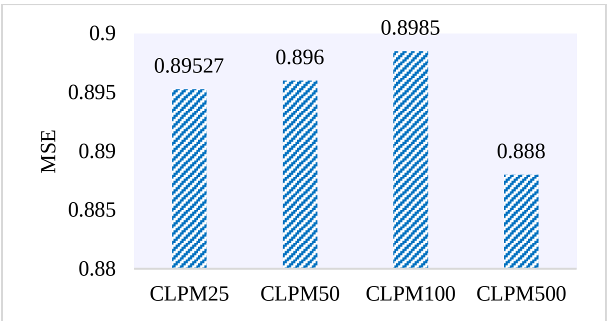


Fig. 10. MSE Evaluation on Variable Latent Size

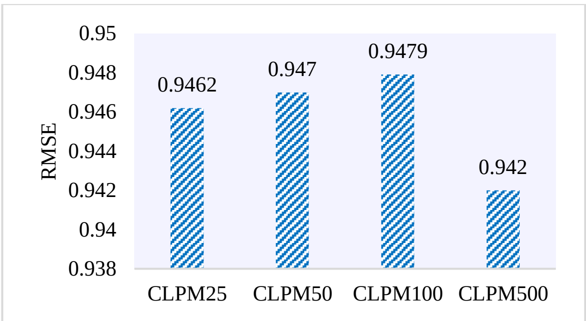


Fig. 11. RMSE Evaluation on Variable Latent Size

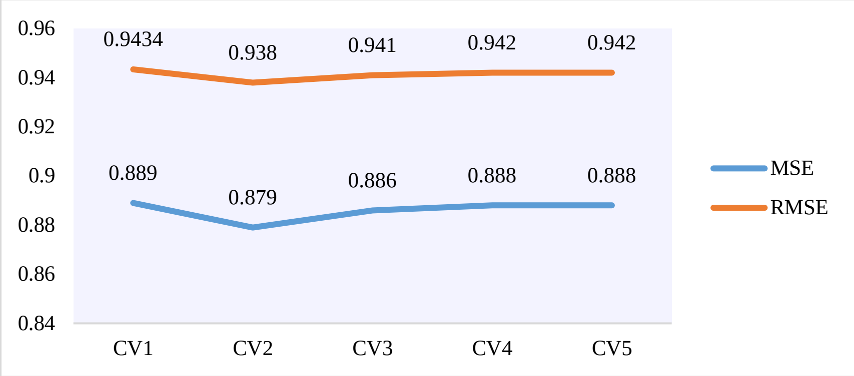


Fig. 11: Cross Validation MSE and RMSE Evaluation

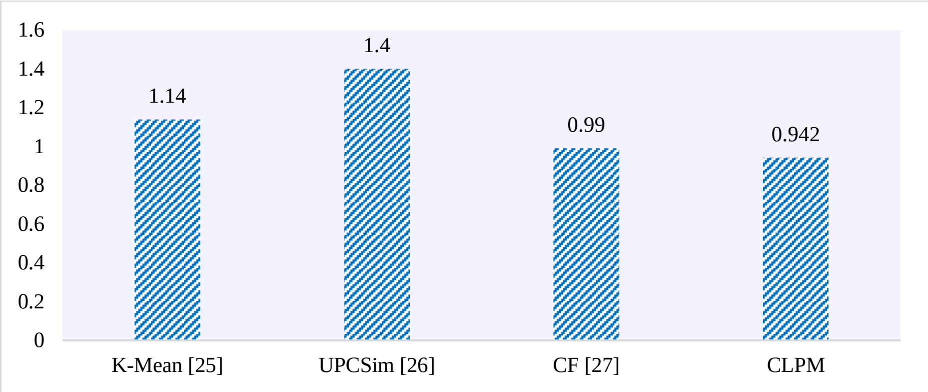


Fig. 12: Comparative Performance Evaluation

Fig 12, comparative state-of-art performance is presented in terms of RMSE. In [25], k-mean clustering algorithm is proposed and its RMSE is about 1.14. in [26], UPCSim is presented that is based on user profile clustering algorithm and achieved approx. 1.4. In [27], CF is presented that have 0.99 RMSE. Whereas in proposed approach, CLPM, the least RMSE is achieved i.e., 0.942.

5. Conclusion

There are many OTT services available now for watching TV shows and movies online, like Netflix, Prime Video, and Disney Plus. Because there is too much information available and there are many parameters to consider when comparing different OTT platforms, users are finding it more difficult to find the perfect match for their interests. Since artificial intelligence is a rapidly developing field, it has a variety of uses, including games, learning settings, and disease diagnostics. The goal of developing AI techniques is to automate intelligent behaviour. This study suggests a hybrid form of collaborative learning that makes use of CNN. The 10-fold cross validation on various CNN layers showed the analysis's efficacy. As a result, this paper can give users information about each platform, propose their favourite content across platforms, and improve the effectiveness of the user's recommendation system.

Refreneces

[1] Bradley K., Rafter R., Smyth B. (2000) Case-Based User Profiling for Content Personalisation. In: Brusilovsky P., Stock O., Strapparava C. (eds) Adaptive Hypermedia and Adaptive Web-Based Systems. AH 2000. Lecture Notes in Computer Science, vol 1892. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-44595-1_7

[2] S. Amara and R. R. Subramanian, "Collaborating personalized recommender system and content-based recommender system using TextCorpus," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 105-109, doi: 10.1109/ICACCS48705.2020.9074360.

[3] Iftikhar Alam, Shah Khusro, Mumtaz Khan, Personalized content recommendations on smart TV: Challenges, opportunities, and future research directions, Entertainment Computing, Volume 38, 2021, 100418, ISSN 1875-9521, https://doi.org/10.1016/j.entcom.2021.100418.

[4] C. I. Eke, A. A. Norman, L. Shuib and H. F. Nweke, "A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions," in IEEE Access, vol. 7, pp. 144907-144924, 2019, doi: 10.1109/ACCESS.2019.2944243.

[5] R. Kaur, S. Singh, and H. Kumar, ‘‘AuthCom: Authorship verification and compromised account

detection in online social networks using AHP-TOPSIS embedded profiling based technique,'' Expert Syst. Appl., vol. 113, pp. 397–414, Dec. 2018.
[6] R. Sharma, D. Gopalani and Y. Meena, "Collaborative filtering-based recommender system: Approaches and research challenges," 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT), 2017, pp. 1-6, doi: 10.1109/CIACT.2017.7977363.
[7] M. E. Ibrahim, Y. Yang, D. L. Ndzi, G. Yang and M. Al-Maliki, "Ontology-Based Personalized Course Recommendation Framework," in IEEE Access, vol. 7, pp. 5180-5199, 2019, doi: 10.1109/ACCESS.2018.2889635.
[8] Khanal, S.S., Prasad, P., Alsadoon, A. *et al.* A systematic review: machine learning based recommendation systems for e-learning. *Educ Inf Technol* 25, 2635–2664 (2020). https://doi.org/10.1007/s10639-019-10063-9
[9] Sha Zhao, Shijian Li, Julian Ramos, Zhiling Luo, Ziwen Jiang, Anind K. Dey, Gang Pan, User profiling from their use of smartphone applications: A survey, Pervasive and Mobile Computing, Volume 59, 2019, 101052, ISSN 1574-1192, https://doi.org/10.1016/j.pmcj.2019.101052
[10] Soares M., Viana P. (2017) The Semantics of Movie Metadata: Enhancing User Profiling for Hybrid Recommendation. In: Rocha Á., Correia A., Adeli H., Reis L., Costanzo S. (eds) Recent Advances in Information Systems and Technologies. WorldCIST 2017. Advances in Intelligent Systems and Computing, vol 569. Springer, Cham. https://doi.org/10.1007/978-3-319-56535-4_33
[11] Boulaalam O., Aghoutane B., El Ouadghiri D., Moumen A., Malinine M.L.C. (2020) Design of a Tourism Recommendation System Based on User's Profile. In: Ezziyyani M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). AI2SD 2019. Advances in Intelligent Systems and Computing, vol 1104. Springer, Cham. https://doi.org/10.1007/978-3-030-36671-1_19
[12] Farid, Marina and Elgohary, Rania and Moawad, Ibrahim and Roushdy, Mohamed, User Profiling Approaches, Modeling, and Personalization (October 27, 2018). Proceedings of the 11th International Conference on Informatics & Systems (INFOS 2018), Available at SSRN: https://ssrn.com/abstract=3389811 or http://dx.doi.org/10.2139/ssrn.3389811
[13] H. Liang, Y. Xu, Y. Li and R. Nayak, "Collaborative Filtering Recommender Systems Using Tag Information," 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008, pp. 59-62, doi: 10.1109/WIIAT.2008.97.
[14] M. Nair A., Preethi N. (2022) A Pragmatic Study on Movie Recommender Systems Using Hybrid Collaborative Filtering. In: Nayak P., Pal S., Peng SL. (eds) IoT and Analytics for Sensor Networks. Lecture Notes in Networks and Systems, vol 244. Springer, Singapore. https://doi.org/10.1007/978-981-16-2919-8_44
[15] Jing Li, Wentao Xu, Wenbo Wan, Jiande Sun, Movie recommendation based on bridging movie feature and user interest, Journal of Computational Science, Volume 26, 2018, Pages 128-134, ISSN 1877-7503, https://doi.org/10.1016/j.jocs.2018.03.009.
[16] M. He, X. Wu, J. Zhang and R. Dong, "UP-TreeRec: Building dynamic user profiles tree for news recommendation," in China Communications, vol. 16, no. 4, pp. 219-233, April 2019, doi: 10.12676/j.cc.2019.04.017.
[17] B. Zhu, F. Ortega, J. Bobadilla, A. Gutiérrez, Assigning reliability values to recommendations using matrix factorization, Journal of Computational Science, Volume 26, 2018, Pages 165-177, ISSN 1877-7503, https://doi.org/10.1016/j.jocs.2018.04.009.
[18] Z. Zhu, D. Li, J. Liang, G. Liu and H. Yu, "A Dynamic Personalized News Recommendation System Based on BAP User Profiling Method," in IEEE Access, vol. 6, pp. 41068-41078, 2018, doi: 10.1109/ACCESS.2018.2858564.
[19] Sipra Sahoo, Bikram Kesari Ratha, Rapid Frequent Pattern Growth and Possibilistic Fuzzy C-means Algorithms for Improving the User Profiling Personalized Web Page Recommendation System, International Journal of Intelligent Engineering and Systems, Vol.11, No.2, 2018 DOI: 10.22266/ijies2018.0430.26
[20] R. K. Chaurasiya and U. Sahu, "Improving Performance of Product Recommendations Using User Reviews," 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), 2018, pp. 1-4, doi: 10.1109/ICRAIE.2018.8710414.
[21] R. Bandana, "Sentiment Analysis of Movie Reviews Using Heterogeneous Features," 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), 2018, pp. 1-4, doi: 10.1109/IEMENTECH.2018.8465346.
[22] Luong Vuong Nguyen et al. "Content-Based Collaborative Filtering using Word Embedding: A Case Study on Movie Recommendation", RACS '20, October 13–16, 2020, Pages 96–100, https://doi.org/10.1145/3400286.3418253

[23] Yi, S., Liu, X. Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review. *Complex Intell. Syst.* **6,** 621–634 (2020). https://doi.org/10.1007/s40747-020-00155-2
[24] https://movielens.org
[25] Rahul, Mayur, et al. "Impact of similarity measures in K-means clustering method used in movie recommender systems." IOP Conference Series: Materials Science and Engineering. Vol. 1022. No. 1. IOP Publishing, 2021.
[26] Widiyaningtyas, Triyanna, Indriana Hidayah, and Teguh B. Adji. "User profile correlation-based similarity (UPCSim) algorithm in movie recommendation system." Journal of Big Data 8.1 (2021): 1-21.
[27] Salloum, Salam, and Dananjaya Rajamanthri. "Implementation and Evaluation of Movie Recommender Systems Using Collaborative Filtering." Journal of Advances in Information Technology Vol 12.3 (2021).
[28] M. Tkalcic, A. Kosir, and J. Tasic, "Affective recommender systems: the role of emotions in recommender systems," in Proc. The RecSys 2011 Workshop on Human Decision Making in Recommender Systems, 2011, pp. 9-13
[29] S. Kanoje, S. Girase, and D. Mukhopadhyay, ''User profiling trends, techniques and applications,'' Mar. 2015, arXiv:1503.07474. [Online]. Available: https://arxiv.org/abs/1503.07474
[30] C. Jang, H. Chang, H. Ahn, Y. Kang, and E. Choi, ''Profile for effective service management on mobile cloud computing,'' in Advanced Communication and Networking. Berlin, Germany: Springer, 2011, pp. 139–145.
[31] Y. Yang, ''Web user behavioral profiling for user identification,'' Decis. Support Syst., vol. 49, no. 3, pp. 261–271, Jun. 2010
[32] A. Farseev, M. Akbari, I. Samborskii, and T.-S. Chua, '''360◦ user profiling: Past, future, and applications' by Aleksandr Farseev, Mohammad Akbari, Ivan Samborskii and Tat-Seng Chua with Martin Vesely as coordinator,'' ACM SIGWEB Newsletter, no. 4, 2016.
[33] D. Poo, B. Chng, and J.-M. Goh, ''A hybrid approach for user profiling,'' in Proc. 36th Annu. Hawaii Int. Conf. Syst. Sci., Jan. 2003, p. 9.
[34] S. Liang, X. Zhang, Z. Ren, and E. Kanoulas, ''Dynamic embeddings for user profiling in Twitter,'' in Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Aug. 2018, pp. 1764–1773
[35] Qian Gao, Su Mei Xi and Young Im Cho, "A multi-agent personalized ontology profile based user preference profile construction method," IEEE ISR 2013, 2013, pp. 1-4, doi: 10.1109/ISR.2013.6695695.
[36] S. Alaoui, Y. EL Bouzekri EL Idrissi, and R. Ajhoun, ''Building rich user profile based on intentional perspective,'' Procedia Comput. Sci., vol. 73, pp. 342–349, Jan. 2015.
[37] Wei, Jian & He, Jianhua & Chen, Kai & Zhou, Yi & Tang, Zuoyin. (2016). Collaborative Filtering and Deep Learning Based Recommendation System For Cold Start Items. Expert Systems with Applications. 69. 10.1016/j.eswa.2016.09.040.
[38] Zhou, Xiaojun & Xu, Zhen & Wang, Liming & Chen, Kai & Chen, Cong & Zhang, Wei. (2018). Inside the Closed World: User and Device Profile Analytics for SCADA Security. MATEC Web of Conferences. 173. 03039. 10.1051/matecconf/201817303039.
[39] Pradeep Kumar Singh, Esam Othman, Rafeeq Ahmed, Awais Mahmood, Habib Dhahri, Prasenjit Choudhury, Optimized recommendations by user profiling using apriori algorithm, Applied Soft Computing, Volume 106, 2021, 107272, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.107272
[40] Hasan, Ali, Sana Moin, Ahmad Karim, and Shahaboddin Shamshirband. 2018. "Machine Learning-Based Sentiment Analysis for Twitter Accounts" *Mathematical and Computational Applications* 23, no. 1: 11. https://doi.org/10.3390/mca23010011
[41] Najafabadi, M.K., Mohamed, A.H. & Mahrin, M.N. A survey on data mining techniques in recommender systems. Soft Comput 23, 627–654 (2019). https://doi.org/10.1007/s00500-017-2918-7
[42] H. Liang, "DRprofiling: Deep Reinforcement User Profiling for Recommendations in Heterogenous Information Networks," in IEEE Transactions on Knowledge and Data Engineering, doi: 10.1109/TKDE.2020.2998695.
[43] Aljunid M.F., Manjaiah D.H. (2019) Movie Recommender System Based on Collaborative Filtering Using Apache Spark. In: Balas V., Sharma N., Chakrabarti A. (eds) Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing, vol 839. Springer, Singapore. https://doi.org/10.1007/978-981-13-1274-8_22
[44] Jian Peng, Kim-Kwang Raymond Choo, Helen Ashman, User profiling in intrusion detection: A review, Journal of Network and Computer Applications, Volume 72, 2016, Pages 14-27, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2016.06.012.

[45]M. Heidari, J. H. Jones and O. Uzuner, "Deep Contextualized Word Embedding for Text-based Online User Profiling to Detect Social Bots on Twitter," 2020 International Conference on Data Mining Workshops (ICDMW), 2020, pp. 480-487, doi: 10.1109/ICDMW51313.2020.00071.