JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Neural Collaborative Filtering For Movie Recommendation System

Ritika Varpe
CSE – AI & DS
MIT-WPU
PUNE, INDIA

Seema Patil
Department of CSE
MIT-WPU
Pune,, India

***Abstract*—This report conducts research on movie recommendation systems using deep learning concepts. Which movies are best depends upon how they have received the ratings. By classifying and predicting the ratings given by the user which movie should be recommended to the user is decided. It has classified the ratings according to the unique users and the unique movies. A traditional method for recommendation systems is used, matrix factorization. The summary has been driven by matrix factorization and a neural collaborative filtering model. It gives input, different layers, and output. Deep learning has great success in various areas such as segmentation, image classification, speech and text recognition, facial expression recognition, etc. A small comparison between matrix factorization and neural collaborative filtering is made which gives a brief idea about a recommendation system. To suggest two different methods to use of the neural network that is collaborative filtering and matrix factorization where the concatenation of user and item embedding is performed. The machine learning algorithm that is mean absolute error has been used for the prediction of the models. After this, the model uses the evaluation matrix for calculating the predictions received by the model. The recommendation system has its evaluation matrix such as Discounted Cumulative Gain(DCG) and Normalized Discounted Cumulative Gain(NDCG) which has helped to give accurate predictions. Additionally, the report explores the limitations and challenges in applying recommendation systems along with serval avenues for future research.**

Keywords*— *Recommendation System, Deep Neural Network, Matrix factorization, Neural collaborative filtering, Embedding.

I. Introduction

In the information explosion era, recommender systems play an important role in information overloading. It has been widely adopted by many online advertising, E-commerce, streaming platforms, Education, etc. The recommendation system primarily uses collaborative filtering where analysis is made on the user interaction patterns with items. The research is carried out on different concepts of recommendation systems. It consists of the recommendation model as well as its evaluation matrix. Based on individual preferences, and user feedback future research could focus on the personalizing recommendation system for advanced technology. To make personalized recommendations to the user this information is considered. The report also discusses neural collaborative filtering. The information goes from different layers which are present in the neural collaborative filtering and then the output is achieved. By adjusting specific parameters the output is received according to the desired. The same is performed for the matrix factorization.

Matrix factorization is considered a model for recommendation systems. It is a traditional method that is frequently used for recommendation systems. The research more focuses on how it improves the efficiency of the recommender system by using different techniques. Recent studies aim to enhance its efficiency through various techniques related to user-item interactions. The latest and old movies are uploaded and according to the ratings they are classified and arranged for the trending section or most viewed section. The whole purpose is the user should get relaxed and entertained for some time from their busy schedule. These types of models are generally used by famous companies such as Netflix, Amazon, Hotstar, etc. As the number of viewers increases accordingly the company gains its profit. A small comparison is made between the matrix factorization and the neural collaborative filtering so that it would be easy to under both concepts.

Netflix	3/5th of the movies watched are recommended
Amazon	nearly 75% of the user engagement has increased.

Table 1. Benefits of the Companies Using Recommendation System

Further explaining this model evaluation metrics are used for information retrieval and recommendation systems. The metrics such as Discounted cumulative gain and normalized discounted cumulative gain are used. Both metrics are considered for the ideal ranking, common usage in information retrieval, and interpretability. To demonstrate the effectiveness of NCF approaches and deep learning concepts the performance is made on the two real-world datasets.

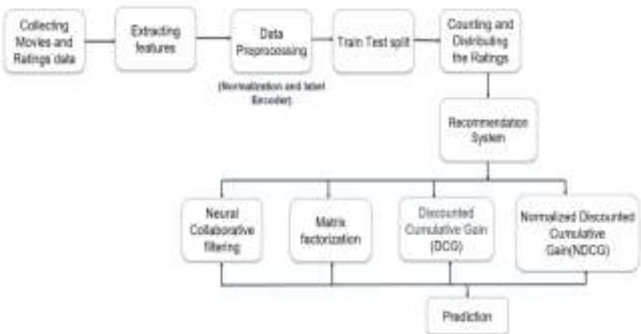
II. LITERATURE SURVEY

Neural collaborative filtering focuses on the recent advancement in deep learning–based recommendation models. There are different papers that discuss the role of collaborative filtering and its advantages in this field. To find the potential customers most of the recommendation systems today use ratings given by the previous users. Therefore an efficient data handling model is essential for improving efficiency and accuracy. Others have proposed novel metrics for measuring system efficiency using clustering techniques.

A paper in an international journal for computer applications has used approaches such as content-based, collaborative, utility, and hybrid. Additionally, it also explores research on data augmentation techniques, performance evaluation, and training strategies. It focuses on the evaluation of the metrics that are specified in the recommendation systems for movies. Other research has proposed a new way of finding solutions that learn from past experiences and recommend changes to the system. It mainly focuses on adapting to the new technique and making the model advanced. For example, Netflix asks users for reviews so that it can make its content favorable to the users.

III. METHODOLOGY

In these reports, a recommendation system is made that uses deep learning techniques. The dataset used is a movie and the rating files. For the movie recommendation system, the NCF and matrix factorization are used. The data is collected and



preprocessing is applied to it. Further, the train test split is used for dividing the data into the train and the test model. For clear measures, the mean absolute error is used which returns the mean of the absolute error. Matrix factorization of neural networks is used. It has used the activation function of "ReLU" and the optimizer as Adam. It has shown the different layers and how the data is processed from one layer to another. Similarly, neural collaborative filtering is also mentioned. For a better recommendation system, the DCG and NDCG are used. By using these metrics the prediction and the evaluation are made. The main problem is addressed which is working on large datasets and making the model efficient.

A. Architecture Diagram

The architecture diagram is shown in Fig 1.

Fig 1. System Architecture Diagram for Movie Recommendation System

B. Dataset

The datasets include the movie and rating data. Fig 2.1 shows the movie data which has 9742 rows and 3 columns. Fig 2.2 shows the rating data which has 100837 rows and 4 columns.

	MovieID	Title	Genres
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	2	Jumanji (1995)	Adventure|Children|Fantasy
2	3	Grumpier Old Men (1995)	Comedy|Romance
3	4	Waiting to Exhale (1995)	Comedy|Drama|Romance
4	5	Father of the Bride Part II (1995)	Comedy

Fig 2.1 Movies Dataset

	UserId	MovieId	Rating	Timestamp
0	1	1	4.0	964982703.0
1	1	3	4.0	964981247.0
2	1	6	4.0	964982224.0
3	1	47	5.0	964983815.0
4	1	50	5.0	964982931.0

Fig 2.2 Rating Dataset

C. Data Preprocessing

The preprocessing is performed on the data. First, the missing data values are handled then it is normalized for stability. The label encoder is used to convert the categorical data into the numerical on the rating dataset. It shows the unique number of users for the unique movies. The Fig 3 shows the rating diagram.

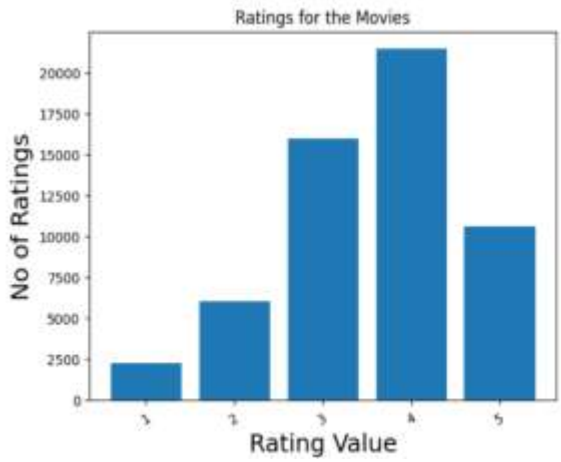
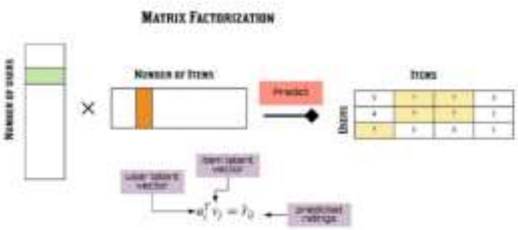


Fig 3. Shows the Distribution of Ratings

D. Model Training

Using the training data the model has been trained using various deep learning concepts. For model training, matrix factorization and neural collaborative filtering are used. Their accuracy is also mentioned.

a) **Matrix factorization:** To show the user-item interaction the matrix factorization technique is used. It has used the activation function as 'Relu'. It is used to analyze and predict user preferences



which are based on historical interactions with movies. To provide personalized recommendations to the users the system has to approximate the original matrix with the product of two lower-dimensional matrices. Through these, the system can estimate the missing entries in the matrix. From this model, the total parameters we have received is 384281 which is 1.47MB. For adaptive learning rates, the Adam optimizer is used. Then the model is compiled using metrics such as mean absolute error. The model is fitted with the epochs as 10 and the validation split as 0.1. The model evaluation is done using the 'userid', 'movie', and the 'ratings'. The rest of it is shown in a tabular manner. Fig 4.1 Shows the working of the matrix factorization for clear understanding.

Fig 4.1 Matrix Factorization

b) **Neural Collaborative filtering:** For learning of complex relationships between the user and the item the neural collaborative filtering approach is used. It directly learns from the patterns and maps them into the latent space. To capture the intricate patterns multi-layer perceptron is used. By training on the given data, the model can predict user preferences for unseen data as well. Accordingly, it gives high accuracy and provides more personalized and effective recommendations. The model has used the MLP which consists of layers like the input layer, the hidden layer, and the output layer. The activation function used for the hidden layer is 'Relu' and for the output, the layer is 'Sigmoid'. For compiling the model, the Adam optimizer is used. From the model summary, the total parameter we received is 686273 which is 2.62MB. Fig 4.2 shows the detailing of the multi-layer perceptron.

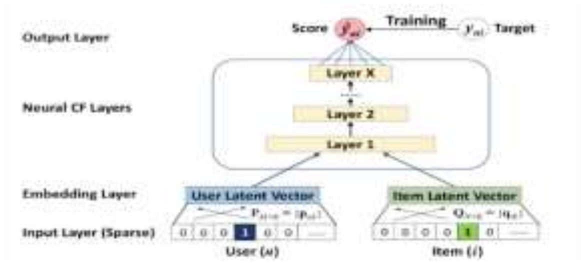


Fig 4.2 Neural Collaborative Filtering

E. Model Evaluation

The model evaluation is done by finding out the accuracy and the predictions. Here is a small idea about the models.

a) **Discounted Cumulative Gain:** It is considered a crucial metric for evaluating the recommendation system. The calculation of DCG involves the summing of relevance scores of movies at each position. Here higher the DCG value is considered as more relevant as it reflects the system's ability for user preferences and movie suggestions. The Fig 5.1 represents the formula for DCG.

$$\mathrm{DCG_p} = \sum_{i=1}^{p} \frac{2^{rel_i} - 1}{\log_2(i+1)}$$

Fig 5.1 DCG formula

b) **Normalized Discounted Cumulative Gain:** To assess the quality of recommendations that are provided by users these metrics are used. It normalizes the DCG score by further dividing it and this represents the maximum possible DCG for a given movie set of recommendations. It makes it easier for algorithms to have large datasets. The higher NDCG tells us that the model is more relevant. It also indicates the accurately ranked list of movie recommendations and compelling movie suggestions to the users. The Fig 5.2 shows the formula for NDCG.

$$NDCG = \frac{DCG}{IDCG}$$

Fig 5.2 shows the formula for NDCG.

c) **Accuracy:** An evaluation metric of several correct predictions made by the model. Fig 5.3 shows the formula

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

Fig 5.3 The Formula for Accuracy.

F. Result :

The matrix factorization gives an accuracy of 0.84. That is of 84 %. The Neural Collaborative filtering gives as accuracy as 0.87. The fig 6 shows the prediction of neural collaborative filtering. Higher accuracy values suggest that the neural collaborative model is more effective in predictiong user prefernces and providing relevant recommendations. The filtering approach is slightly having higher accuracy compared to matrix factorization, implying that it may be better at capturing complex patterns and relationships in user-item interactions.

Model	Accuracy
Matrix Factorization	0.84
Neural Collaborative Filtering	0.87

FIG 6 PREDICTION

IV ACKNOWLEDGMENT

In conclusion, neural collaborative filtering is considered a powerful and effective approach for movie recommendation systems. From the result, we came to know that neural collaborative has higher accuracy than matrix factorization. The NCF learns from the complex patterns and clearly understands the relation between the user and the item. As the model is flexible it allows for easy integration of additional features. In the future, there will be a user interface so that the user can interact with the model and get results accordingly. Also focusing on large and complex datasets so that the model would give more accurate and precise results. From this, the efficiency of the model also increases. Thus NCF has proven to give better results and more accuracy for recommendation systems. As it uses advanced techniques rather than traditional methods.

Overall NCF represents the best way for a recommendation system, providing the users with tailored and engaging movie suggestions while improving users' satisfaction and engagement. By using the evaluation metrics of deep learning techniques such as DCG and NDCG it is given the accurate accuracy. Hence this model has proven to be best for recommendation systems. The big companies are using such models for better results and user satisfaction.

REFERENCES

[1] K Revanth sai, E Shankar et al. "*Movie recommendation using sentiment analysis from micro-blogging data*" International Journal of creative research thoughts (IJCRT) 2023, Volume 11, Issue 3 March 2023 | ISSN : 2320-2882.

[2] Phonexay Vilakone, Doo-Soon Park, Khamphaphone Xinchang and Fei Hao et al. "*An Efficient movie recommendation algorithm based on improved k-clique*.", 2018 Journal of human-centric computing and information sciences.

[3] Manoj kumar, D.K. Yadav, Ankur Singh, Vijay Kr. Gupta et al. "*A Movie Recommender System : MOVREC*", International Journal of Computer Application(0975-8887) Volume 124 – No.3, August 2015.


```
631/631 [==============================] - 3s 4ms/step
0.8758992961371489
```


[4] Xiangnan He ,Lizi Liao, Hanwang Zhang, Liqiang Nie Xia Hu, Tat-Seng Chua," *Neural Collaborative Filtering*", 2017 International World Wide Web Conference Committee (IW3C2), published under Creative Commons CC BY 4.0 License. WWW 2017, April 3–7, 2017, Perth, Australia. ACM 978-1-4503-4913-0/17/04.

[5] . Redwane Nesmaoui, Mouad Louhichi, Mohamed Lazaar,"*A collaborative filtering movies recommendation system based on graph neural network*", Procedia Computer Science,Volume 220,2023, Pages 456-461, ISSN 1877-0509.

[6] Nisha Bhalse, Ramesh Thakur, "*Algorithm for movie recommendation system using collaborative filtering*", February 2021,Materials Today Proceedings DOI:10.1016/j.matpr.2021.01.235

[7] Luyao Li, Hong Huang, Qiangian Li, Junfeng Man," *Personalized movie recommendations based on deep representation learning*", 2023; 9:e1448.Published online 2023 Jul 12. doi: 10.7717/peerj-cs.1448,PMCID: PMC10403217 PMID: 37547384.

[8] Sonu Airen, Jitendra Agrawal, "*Movie Recommender System Using Parameter Tuning Of User and Movie Neighbourhood via Co-Clustering*" International Conference on Machine Learning and Data Engineering, Procedia Computer Science 218 (2023) 1176-1183.

[9] Jayalakshmi S, Ganesh N, Čep R, Senthil Murugan J.*" Movie Recommender Systems: Concepts, Methods, Challenges, and Future Directions"*. Sensors (Basel). 2022 Jun 29;22(13):4904. doi: 10.3390/s22134904. PMID: 35808398; PMCID: PMC9269752.