A Deep Learning Based Collaborative Neural Network Framework for Recommender System

1st Maram Almaghrabi
Faculty of Computer Science.
Umm-AlQura University
Makkah, Saudi Arabia
mammaghrabi@uqu.edu.sa

2nd Girija Chetty
Faculty of Science & Technology
University of Canberra
Canberra, Australia
Girija.Chetty@canberra.edu.au

Abstract*—Deep learning based neural networks have been attracting significant interest lately, due to their success in complex automatic recognition tasks in many artificial intelligence areas such as language recognition ,computer vision and expert systems. However, in recommendation systems, they have not been exploited fully, and most of the systems rely on traditional collaborative filtering with matrix factorization approaches. In this paper we suggest a novel approach based on deep learning-based augmentation of the collaborative filtering approach for predicting user ratings for different types of media collections in online databases and libraries, including movies, music and book collections. Though there have been few approaches proposed based on deep learning for recommender systems, they were mainly used for modelling additional complementary information available in terms of images, text or acoustic information from different items in the collections. The main component, the interaction between the item and the user was still modelled with traditional matrix factorization approaches, which uses the dot product to extract the latent features between users and items. By augmenting the latent features extracted with deep learning based neural network models, it is possible to capture the hidden interactions and can lead to better recommender system performance even with sparse and unbalanced datasets. The experimental evaluation of the proposed approach based on four different publicly available datasets involving movies, music and book collections, show promising results, in terms of different performance metrics including accuracy and mean absolute error. Experiments on four datasets show the effectiveness of our proposed frameworks.

***Keywords*— Deep learning, Recommender system, Collaborative Filtering, Neural Network.**

I. Introduction

Recommender systems are considered one of the most common ways of personalization and online search popular in recent years. The traditional recommender system has different criteria for users such as user's preferences or user's profile. Deep learning techniques with the recommender system can have good results in many areas such as computer vision, voice recognizer, and language processing. The basis of deep learning is imperative in the classification of complex data to improve the understanding that would come from them. Deep learning is known as deep structured learning. It is a new field of machine learning research and helps modelling data on multiple levels and can discover interactions and correlation between different types of structured and unstructured data, including, image, sound, and text data at a deeper level Building automatic computer-based recommendation systems based on capturing interaction and collaboration between users and items with deep learning approaches can predict and provide better recommendations to users. There are many companies such as Amazon, eBay and Netflix use recommender system to meet users' needs and help them in finding what they need [19]. The traditional recommender is a system based on collaborative filtering techniques, for capturing the user's preferences from online search behaviour. The collaborative filtering technique uses matrix factorization approach for modelling the interaction between the users and items, and merely captures collaboration between the users and the products at a shallow level. The usage of deep learning-based approach can capture the interactions at a deeper level, and model latent hidden features at multiple levels depending on the deep learning architecture used. This can result in better performance of recommender systems and provide more personalized recommendations. Deep learning-based approaches have shown significant promise in several fields of computer vision, voice recognition, Multilanguage processing, and other similar fields. Most of the recommender systems using collaborative filtering techniques, for filtering resources based on the users' ratings, and involves matrix factorization technique where an inner product between the items and users is used to capture the interactions. Moreover, two other issues arise in traditional collaborative filtering techniques, based on matrix factorization approaches. This is in relation to the correlation between several users who rate the items and the number of items, where the number of users who rate the items is much larger than the number of users who expressed the similar interest of items and will continue to maintain similar interests over the time. The second problem is cold–start of new items or old items that received fewer ratings and will be unlikely recommended in a collaborative filtering-based recommender system [23]. Many previous studies such as Sequential deep learning for human action recognition [6] have emphasized that deep learning can capture this interaction at a deeper level, as it involves a process of capturing deeply hidden interactions and learning nonlinear features and functions from complex data. They have proved to be the good learning models for speech recognition, image recognition, face detection, and other similar complex domains [8], [2],[22]. In this paper, we suggest a novel approach based on augmenting the traditional collaborative filtering technique involving matrix factorization of interaction between users and ratings, to extract latent joint features at a deeper level. We used two different deep learning architectures, one

978-1-7281-0404-1/19/$31.00 ©2019 IEEE
DOI 10.1109/iCMLDE.2018.00031

where the latent features from user and items are learnt separately first before matrix factorization, and the second architecture, where multiple stages of feed forward deep network are used for learning joint latent representations in items and users' joint space after matrix factorization. We found that joint learning of representations with multi-stage deep neural networks, results in significantly better performance as compared to individual learning of user and item features. To validate the proposed neural models, we evaluate recommender system performance with several publicly available datasets, and experimental results show that the proposed deep learning based neural models can improve recommender system performance; as compared to traditional matrix factorization based collaborative filtering schemes. We organized the paper as follows. Next Section discusses related works in this area, and Section 3 presents the proposed scheme for deep learning-based recommender system. The experimental evaluation of the proposed scheme and results are discussed in Section 4 and 5, and paper concludes with conclusions and plans for further work in Section 6.

II. Related Works

The recommendation/recommender systems have two ways to access data about the domain, namely the users and the items and the interaction between users and items in terms of reviews and ratings. In many recommender systems, such as Netflix site, the users can browse and watch different content, and they can rate the items they have purchased. The availability of such on-line content search platforms allows a relationship creation between users and items by linking their ratings with different items. This can be presented in a matrix [12]. Each cell within this matrix represents the user's rating for different items. Such as matrix can be useful for predicting ratings for the items that user has been not rated yet or newly added items. The collaborative filter matrix constructed by matrix factorization not just recommends the item to use based on the rating scales, but also provides top–n recommendations based on significant features, for recommending the item to users based on the user's interest [6], and hence, provide to users, several interesting items, such as which book to buy, which video to watch next and which music to listen and buy next. The collaborative filter matrix tends to capture the metadata about users and items. [8]. Several methods have been proposed in improving the recommender systems performance, as recommender systems play an important role in reducing the information overload in online search for resources in collections, be it movies, music, books or research articles. One of the most significant aspects of personalized recommender systems is modelling users' behavior and their preference for items based on their past behavior, interaction or engagement with a resource or item, captured by ratings, clicks and reviews. This interaction aspect has been addressed mostly by collaborative filtering [18, 19]. Among the several collaborative filtering techniques, matrix factorization (MF) [27, 11] is considered one of the most common technique, which extracts users and items interaction into a shared latent space, by projecting a vector of latent features to represent a user or an item. A user-item interaction is demonstrated as an inner product of their latent vectors. There have been also some efforts on combining matrix factorization (MF) based techniques with neighbor-based models [11] or combining it with item content using topic models [10], and use of factorization machines [15] for a generic modelling of features. The dot product or inner product in matrix factorization approach applied for collaborative filtering technique is realized by simply combining the multiplication of latent features linearly and may not be appropriate to capture the complex structure of user interaction data. This insufficient modelling leads to poor performance particularly for imbalanced and sparse data sets, normally in datasets related to user reviews and rating context, where there are large ratings or reviews of one or other category item, but hardly any rating of other categories of items in the dataset. This leads to singularities and intractability of matrix inner product, and poor recommender system performance. This paper explores the augmenting of the collaborative filtering approach with deep neural networks for discovering the complex and deep interactions in the shared space between users, and ratings/reviews, and provide significant improvement as compared to feature engineering and handcrafting techniques used in most of the previous approaches proposed [24, 10].

III. PROPOSED DEEP RECOMMENDER ARCHITECTURE

Our proposed framework consists of two models: Collaborative Filtering Deep Recommender architecture (CFDR) and Collaborative Filtering based Multistage Deep Neural Network architecture (CFMDNN) CFDR and CFMDNN learn to model complex structures of user and item interactions using matrix factorization for extracting latent feature representations either separately (CFDR) or jointly (CFMDNN), with joint learning architecture (CFMDNN) with a capability to learn the joint space at a deeper level, through several stages of learning and discovery, without a need for feature engineering or manual handcrafting needed in extracting and combining the latent feature sets. However, they differ in joint modelling capability of interaction between users and ratings, and in turn in their ability for deep discovery in joint latent feature space. All the models were built using Keras Deep learning tools [22], and for building deep learning based neural models for both CFDR and CFMDNN architectures, we used K-fold cross-validation for training and evaluation and used rating prediction accuracy and the Mean Absolute Error (MAE) as the performance metrics for the evaluation. Also, we have extensively used dropout as a regularization technique to reduce the complexity of the model with the goal to prevent overfitting. Then we applied densely, fully linked feed forward neural network, and Adam optimization algorithm, which is an extension to stochastic gradient descent that has recently seen broader adoption for several deep learning applications, including computer vision and NLP. The detail of the datasets used for model building and evaluation is described in the next Section.

IV. EXPERIMENTAL DETAILS

In this Section, we represent the details of experiments performed for evaluation of the suggested deep recommender models by using publicly available datasets. We first describe the datasets before providing the details of the experiments.

A. ***Datasets:***

1.MovieLens20M Dataset:

MovieLens 20M Dataset is a movie recommender system project dataset and details of this dataset is provided in [19]. Also, it contains 20 million user ratings and 465,000 movie tag which applied to 27,000 movies by 138,000 users and includes 1,100 tags [19]. Table 1 shows the data structure for the rating table, and rating distribution statistics.

Table 1. MovieLens 20M Dataset table structure and ratings distribution

	userid	itemid	ratings	timestamp
635836	4251	1	4.0	973568502
982700	6627	2	3.0	842879122
480565	3259	4	5.0	1230931053
97855	684	5	2.0	997585406
917364	6123	3	4.5	1076944877

2.Amazon Digital Music Dataset:

Amazon Digital Music dataset is from the Amazon product dataset which contains product reviews (reviewer, reviewer name, ratings, text, helpfulness votes, review time) specifically for Digital Music purchase [26]. The ratings distribution and data structure for this dataset is as shown in Table 2.

Table 2. Amazon Digital Music Dataset table structure and ratings distribution

	userid	itemid	ratings	Time
0	A2EFCYXHNK06IS	5555991584	5.0	978480000
1	A1WR23ER5HMAA9	5555991584	5.0	953424000
2	A2IR4Q0GPAFJKW	5555991584	4.0	1393545600
3	A2V0KUVAB9HSYO	5555991584	4.0	966124800
4	A1J0GL9HCA7ELW	5555991584	5.0	1007683200

3. Book crossing Dataset:

The book-crossing dataset is collected by Cai-Nicolas Ziegler, and is described in detail in [7], and [5].

Table 3. Book- Crossing Dataset structure

	userID	ISBN	bookRating
0	276726	0155061224	5
2	276729	052165615X	3
3	276729	0521795028	6
5	276736	3257224281	8
6	276737	0600570967	6

4.Amazon Books dataset:

Amazon books dataset is from the Amazon product dataset which contains only user's ratings (user, item, rating, timestamp) it includes more than 150,000 user's ratings [17].

Table 4. Amazon Books Dataset structure

	userId	itemId	rating	Time
0	AH2L9G3DQHHAJ	116	4	1019865600
1	A2IIIDRK3PRRZY	116	1	1395619200
2	A1TADCM7YWPQ8M	868	4	1031702400
3	AWGH7V0BDOJKB	13714	4	1383177600
4	A3UTQPQPM4TQO0	13714	5	1374883200

B. ***Experimental Setup***

In this section, we describe the details of the experimental setup to evaluate the effectiveness of our proposed CFDR and CFMDNN frameworks and its components by comparing with Collaborative Filtering (CF) without deep learning. In particular, we show, how can enhance the effectiveness of CFMDNN, a joint learning based recommender system model with accuracy and MAE to measure the performance. In addition, the ratings are often specified on a scale that shows the specific level of like or dislike of the item at hand [13]. All models were built using Keras deep learning tools and Python platform. [10]. We have used four datasets, and each dataset has the different rating system. Movielenes 20M dataset used 15 rating-scale, Amazon digital Music dataset used 1-5 rating-scales, Book-Crossing dataset used 0-10 Rating-scale and Amazon books dataset used 1-5 rating scales. When the system has learned the user's previous ratings of items it will give recommendation higher for those items in the future for the user. Besides, the system needs to know the relationship between users and item to make predictions more accurate. We can determine that by applying the Utility Matrix which the data itself is described as a utility matrix, giving for each user-item pair, a value that represents what is acknowledged about the degree of preference of that user for that item. The latent features extracted from the collaborative filtering stage with matrix factorization is processed through a different deep learning algorithmic pipeline for CFMDNN and CFDR architectures shown in Figures 1 and 2 below:

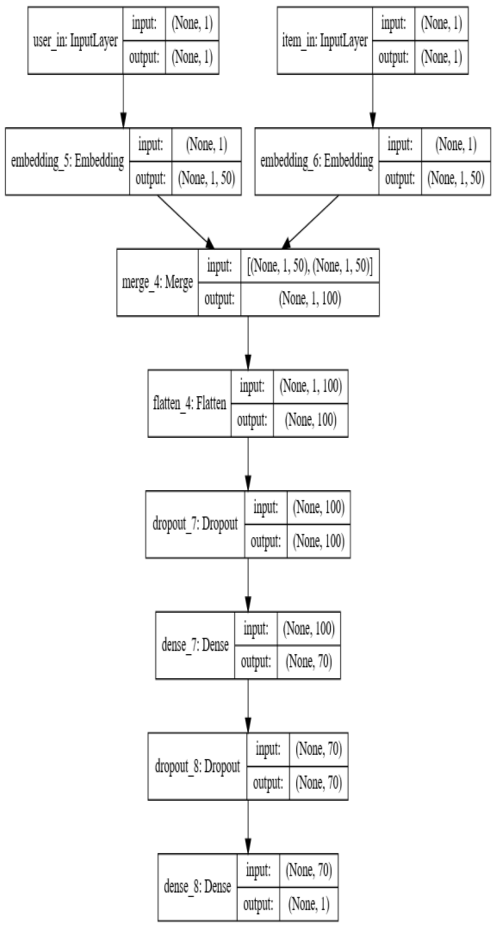

Fig 1: Collaborative Filtering Deep Neural Network Model

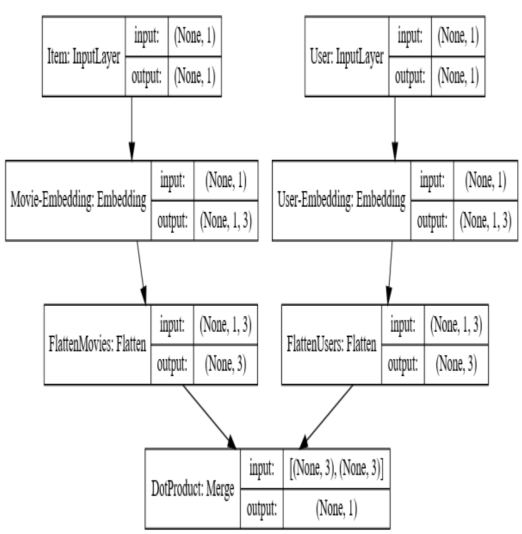

Fig 2 Collaborative Filtering Deep Recommender

A brief description to two deep learning recommender models is provided here:

• User (u) and Item (i) are used to create embedding (low-dimensional) for user and item • Generalized Matrix Factorization (GMF) combines the two-embedding using the dot product. This is our regular matrix factorization. • We combine/concatenate them to create a joint feature vector which can be passed on to the deep neural layers. • For CFDR architecture, the latent feature vector learning based on deep learning happens separately, before matrix factorization, whereas for CFMDNN architecture, in additional latent feature learning through deep learning, additional processing in terms of multiple stages of deep neural network learning allows, joint latent features to be extracted. With an extended processing of joint feature vectors, better discovery and learning of deeper interaction information between users, items and ratings is captured, and hence can improve recommender system performance. Figures 1 and 2 show the two different architectures (CFDR and CFMDNN) • The two models (CFDR and CFMDNN) were built for each dataset using Keras Deep learning tools, and we used an empirical selection of hyper parameters, with leave one out validation with validation set to (80% training data for building the model, and 20% data for testing/validation). • For performance comparison, we used different performance metrics including accuracy in ratings class prediction, as well as precision, recall, True positive and false positive rates, the root means square error RMSE and MAE.

- MAE is computed as follows:

$$MAE = \frac{\sum\{u,i\}|Tu,i - Tu,i|}{n}$$

Where

- n stands for the total number of ratings over all users,
- Tu,i stands for the predicted rating for user u on item i, and
- Tu,i stands for the actual rating.

V. Experimental Results

In this section, we discussed the results of applying two frameworks for all datasets and compared with CF in terms of accuracy, MAE and RMSE.

1. *Ratings Prediction Accuracy*

Figure (3) shows the effectiveness of models using accuracy measure. CFDR and CFMDNN both have excellent accuracy and it showed how the models can improve the accuracy of recommender system from 59.30% to 100% in Amazon Books dataset and 73.80% to 100% in Amazon Digital Music. However, Book-Crossing dataset has improved with 98% in CFDR and got high accuracy in CFMDNN which is better than CFDR. Also, MovieLens 20M dataset got good accuracy in CFMDNN model rather than CFDR model. For baseline comparison, we also show simple Collaborative filtering approach performance as well, which uses a simple matrix factorization using direct dot product of users and items, without deep learning.

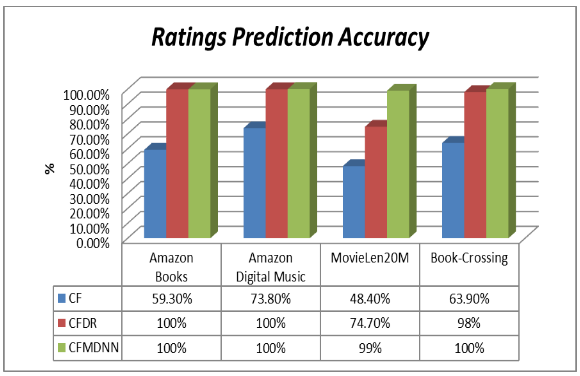

Fig. 3 : Accuracy of CF, CFDR and CFMDNN

2. Mean Absolute Error (MAE)

Instead of classification accuracy or classification error, the most commonly used metric in CF research literature is Mean Absolute Error (MAE) [14],[4], which computes the average of the absolute difference between the predictions and true ratings. From figure (4) it can be seen that the MAE get better in all datasets in CFMDNN models, validating our hypothesis that joint learning space with the CFMDNN model allows better discovery of joint latent features and can lead to a better recommender system rather than non-joint feature learning space as in CFDR model.

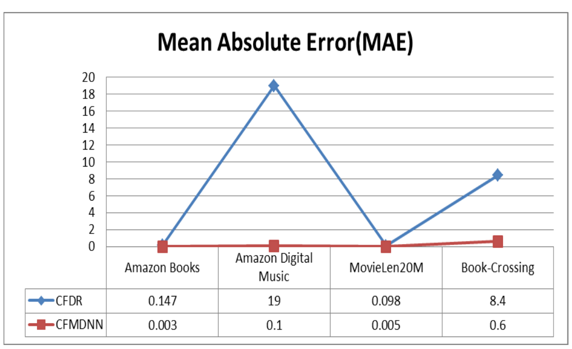

Fig 4: MAE of CF,CFDR and CFMDNN

3. *Root Means Square Error (RMSE)*

Root Means Square Error (RMSE) is considered one of the most important measurements to test the models. The root mean square error (RMSE) has been used in many statistical metric to measure model performance in many field such as climate research studies [4]. It's the square root of the average of squared differences between prediction and exact observation

$$RMSE = \sqrt{\frac{\sum\{u,i\}(Tu,i - Tu,i)^2}{\text{n}}}$$

From figure (5) we can see that improvement RMSE improvement highly in CFMDNN model. For example, MovieLens 20M dataset improve from 35.5 to 1 with CFMDNN model, whereas with CFDR model, the RMSE achieved it 9.71.

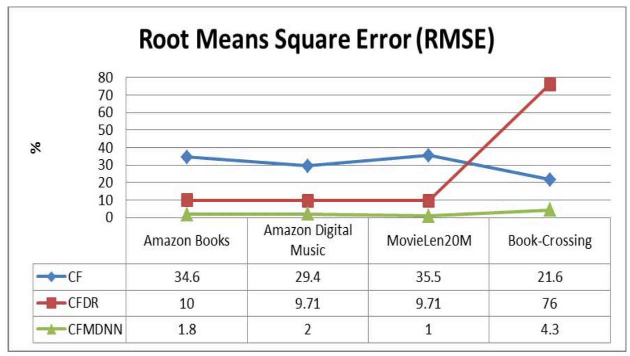

Fig. 5 : Root Mean Square Error (RMSE) for CF,CFDR and CFMDNN

4. CFMDNN Prediction Model

Since CFMDNN model turned out be best performing, we examined the ratings prediction accuracy for this model for all datasets. Tables 5 to 8 show the ratings prediction performance for each dataset based on best performing CFMDNN model

Table 5. Ratings Prediction with CFMDNN prediction for MovieLens 20M dataset.

User_ID	Movie_ID	Rating (Actual Data)	CFMDNN Prediction Model
1012	187	4	4
0365	236	4	4
0237	0277	4	4
0237	0323	4	4

Table 6. Ratings prediction on Amazon digital Music

User_ID	*Music_ID*	*Rating Actual Data*	CFMDNN *Prediction Model*
2	1	4	4
5	0	5	5
6	2	3	3
7	0	5	5
8	0	5	5

Table7. Ratings prediction on Book-Crossing Dataset

User_ID	*Book_Id*	*Rating Actual Data*	CFMDNN *Prediction Model*
15	6	9	9
22	3	8	8
23	5	10	10
29	5	10	10

Table 8. Ratings prediction on Amazon Books Dataset

User_ID	*Book_Id*	*Rating Actual Data*	CFMDNN *Prediction Model*
32	324	4	4
222	55	5	5
232	7	2	2
29	2	4	4

VI. Conclusion

This paper presents a new deep learning-based framework for augmenting the collaborative-filtering based recommendation systems for capturing the deep and hidden interactions between users and items and improves the recommender system performance in terms of ratings prediction of unseen items and users. The evaluation of the proposed framework, with two different neural recommender models, on four different publicly available datasets, has shown improved performance of the system, as compared to traditional collaborative filtering-based approach based on matrix factorization techniques. Particularly, the multistage deep collaborative neural network model gives good performance and robustness due to its ability for joint modelling and discovery to deep interactions. For future work, we plan to extend CFMDNN framework to incorporate novel deep neural network approaches that can work on sparse data. Also, future work planned involves further exploration into deep learning schemes in terms of active and online learning with low or zero resource models, based on small or no training data, particularly unsupervised deep learning models.

VII. References

[1] Billsus, D., and Pazzani, M. Learning collaborative information filters. In 15th International Conference on machine Learning. Madison 1998, pp. 46–54.

[2] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In WWW, pages 285-295, 2001.

[3] Baccouche, Moez, et al. "Sequential deep learning for human action recognition." International Workshop on Human Behavior Understanding. Springer, Berlin, Heidelberg, 2011

[4] Chai, Tianfeng, and Roland R. Draxler. "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature." Geoscientific model development 7.3 ,2014,: 1247-1250.

[5] Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, Georg Lausen; Proceedings of the 14th International World Wide Web Conference (WWW '05), May 10-14, 2005, Chiba, Japan.

[6] Deshpande, Mukund, and George Karypis. "Item-based top-n recommendation algorithms." ACM Transactions on Information Systems (TOIS) 22, no. 1 ,2004,: 143-177.

[7] De Nart, Dario, and Carlo Tasso. "A personalized concept-driven recommender system for scientific libraries." Procedia Computer Science 38 ,2014, : 84-91.

[8] Felfernig, Alexander, et al. "Basic approaches in recommendation systems." Recommendation Systems in Software Engineering. Springer, Berlin, Heidelberg, 2014. 15-37.

[9] Harper, F. Maxwell, and Joseph A. Konstan. "The movielens datasets: History and context." Acm transactions on interactive intelligent systems (tiis) 5.4 ,2016, 19.

[10] https://keras.io/

[11] H. Wang, N. Wang, and D.-Y. Yeung. Collaborative deep learning for recommender systems. In KDD, pages 1235-1244, 2015.

[12] Huang, Junshi, et al. "Cross-domain image retrieval with a dual attribute-aware ranking network." Proceedings of the IEEE international conference on computer vision. 2015

[13] Jahrer, Michael, Andreas Töscher, and Robert Legenstein. "Combining predictions for accurate recommender systems." Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2010

[14] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, “Evaluating collaborative filtering recommender systems,” ACM Transactions on Information Systems, vol. 22, no. 1, pp. 5–53, 2004.

[15] K. Goldberg, T. Roeder, D. Gupta, and C. Perkins, “Eigentaste: a constant time collaborative filtering algorithm,” Information Retrieval, vol. 4, no. 2, pp. 133–151, 2001.

[16] Koren, Yehuda, Robert Bell, and Chris Volinsky. "Matrix factorization techniques for recommender systems." Computer 8 ,2009,: 30-37.

[17] Li, S 2017. https://towardsdatascience.com/how-did-we-build-book-recommender-systems-in-an-hour-part-2-k-nearest-neighbors-and-matrix-c04b3c2ef55c

[18] Mcginty, Lorraine, and Barry Smyth. "Adaptive selection: An analysis of critiquing and preference-based feedback in conversational recommender systems." International Journal of Electronic Commerce 11.2 ,2006,: 35-57.

[19] Refaeilzadeh, Payam, Lei Tang, and Huan Liu. "Cross-validation." Encyclopedia of database systems. Springer, Boston, MA, 2009. 532-538.

[20] Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Recommender systems: introduction and challenges." Recommender systems handbook. Springer, Boston, MA, 2015. 1-34.

[21] Sarwar, Badrul, et al. "Item-based collaborative filtering recommendation algorithms." Proceedings of the 10th international conference on World Wide Web. ACM, 2001

[22] Osoba, Osonde A. and Paul K. Davis, An Artificial Intelligence/Machine Learning Perspective on Social Simulation: New Data and New Challenges. Santa Monica, CA: RAND Corporation,2018.https://www.rand.org/pubs/working_papers/WR1213.html.

[23] Wong, Aaron, and Brittany Factura. "Amazon Digital Music: Sentiment Analysis and Text Mining." Amazon Digital Music: Sentiment Analysis and Text Mining.

[24] X. Wang and Y. Wang. Improving content-based and hybrid music recommendation using deep learning. In MM, pages 627-636, 2014.

[25] X. He, H. Zhang, M.-Y. Kan, and T.-S. Chua. Fast matrix factorization for online recommendation with implicit feedback. In SIGIR, pages 549-558, 2016.

[26] Yi, Xun, and Yanchun Zhang. "Privacy-preserving naive Bayes classification on distributed data via semi-trusted mixers." Information systems 34.3 ,2009,: 371-380.

[27] Y. Koren. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In KDD, pages 426-434, 2008.