

Explore the Personalized Resource Recommendation of Educational Learning Platforms: Deep Learning

Xiaosi Qi¹, Jianwei Zhao² and Guochao Hu^{*3}

Email: qixaiosi_903903@163.com, zhaojianwei515@163.com, cai8828@163.com

¹School of Grammar, Yanching Institute of Technology, Langfang, Hebei 065201, China

²Institute of Marxism, Yanching Institute of Technology, Langfang, Hebei 065201, China

³Accounting Institute, Yanching Institute of Technology, Langfang, Hebei 065201, China

* Corresponding author

Keywords: educational learning platform, personalized resource recommendation, deep learning, recommendation algorithm

Received: Februar 1, 2024

With the development of educational learning platforms, the resources available on them have become increasingly abundant, which has increased the difficulty of personalized resource recommendations. To further improve the effectiveness of personalized recommendation, this paper first analyzed the neural collaborative filtering (NeuCF) algorithm and then improved generalized matrix factorization (GMF) to expanded GMF (EGMF). Furthermore, additional user and project information was incorporated into the input to better capture user preferences and further improve personalized recommendation effects. Experiments were performed using data from a massive open online courses (MOOC) platform. The experimental results demonstrated that the improved NeuCF (INeuCF) algorithm outperformed other algorithms, including the user-collaborative filtering algorithm, in personalized resource recommendation. When the length of the recommendation list was 10, the INeuCF algorithm achieved an F1 value of 0.227 and a normalized discounted cumulative gain (NDCG) value of 0.337. In comparison to the NeuCF algorithm, the EGMF improved the F1 value by 0.008 and the NDCG value by 0.005. Additionally, the incorporation of other information further enhanced the F1 value by 0.01 and the NDCG value by 0.007. These results verify the effectiveness of the proposed improvement to the NeuCF algorithm and suggest that the method can be practically applied to educational learning platforms to achieve more effective personalized resource recommendations.

Povzetek: Članek se ukvarja s priporočanjem virov na izobraževalnih platformah z uporabo globokega učenja. Predlagana je izboljšava algoritma nevronskega kolaborativnega filtriranja (INeuCF), ki vključuje dodatne informacije uporabnika in projekta.

1 Introduction

Under the influence of the continuous development of cloud computing, big data, and other technologies, there are also many new changes in the field of education. More and more education and learning platforms appear [1], which further expands the way learners learn online. However, with the continuous development of education and learning platforms, there are more and more learning resources stored in them. It brings more resources for learners to choose from, but at the same time, it also increases the difficulty for learners to find suitable and interesting resources. At present, most of the education and learning platforms use the method of information retrieval, and learners search resources through keywords. However, this method requires learners to screen resources again in the search process and fails to consider the learners' situation and provide more personalized search results. Personalized recommendation is a method to provide personalized content for users according to their interest characteristics [2]. At present, it has been successfully applied in e-commerce platforms, social

platforms, and other fields [3]. This paper mainly discussed the personalized recommendation of resources on the education and learning platform and designed an improved neural collaborative filtering (INeuCF) algorithm based on deep learning. The recommendation effect of this method was verified through experiments on a dataset. This work presents a novel approach to enhance resource recommendations on the education and learning platform, thereby improving learners' learning experience and promoting further development of educational learning platforms.

2 Related works

	Method	Results and findings
Ohtomo et al. [4]	A supervised multivariate autoencoder considering user preferences	They proved the effectiveness of this approach through experiments on a dataset

		containing six users and 99,844 posts.
Gan et al. [5]	A knowledge-enhanced contextual multiarmed bandits (CMAB) model	This method had better recommendation accuracy and diversity compared to the most advanced metho.
Alnahhas et al. [6]	A method to extend the user model by using the semantic relationship of the knowledge base	This model and the associated recommendation algorithm were superior to all previous methods.
Prathama et al. [7]	A multiple implicit feedback matrix collaborative factorization method	The combination of multiple implicit feedback and matrix cofactorization improved recommendation quality.

3 Resource personalized recommendation algorithm based on deep learning

3.1 Collaborative filtering algorithm

The characteristics of online education learning platforms lie in the abundance and diversity of learning resources. Learners who want to find suitable resources from these platforms inevitably need to invest a significant amount of time. However, online education learning platforms also provide a wealth of learner data, which greatly assists in the design of personalized recommendations.

At present, the collaborative filtering (CF) algorithm is the most commonly used in personalized recommendations, which can be divided into the following two types according to the different objects.

(1) User-based CF (User-CF)

This method computes user similarity to recommend items that similar users like [8]. The steps are as follows.

① Establish a scoring matrix. For M users and N items, the scoring matrix is:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ r_{21} & r_{22} & \cdots & r_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M1} & r_{M2} & \cdots & r_{MN} \end{bmatrix}.$$

② User similarity is calculated based on cosine similarity and so on.

③ The user similarity is sorted, and the previous n users are used as the similar user group of the target user.

④ The items that have not been evaluated by the target user are selected from the items evaluated by the

similar user group as the candidate set to be recommended to the target user.

(2) Item-based CF (Item-CF)

Compared with user-CF, this method changes the calculation of user similarity to the calculation of item similarity and recommends items similar to their favorite items to users [9].

3.2 Neural collaborative filtering algorithm

Based on CF, to better understand the potential relationship between users and items, the NeuCF algorithm [10] has emerged. This method combines generalized matrix factorization (GMF) and multilayer perceptron (MLP). The former learns linear features, and the latter learns nonlinear features, so as to achieve better recommendation effects.

GMF is defined as follows.

By inputting the one hot codes of user and item IDs, latent feature vectors p_u and q_i of the user and item are obtained, and their correlation is written as:

$$\delta(p_u, q_i) = p_u \cdot q_i.$$

The prediction result is obtained by input-layer mapping:

$$\hat{y}_{u,i} = f(w(\delta))$$

where w is the output layer weight and f is the activation function.

MLP is defined as follows.

For input sample (x, y) , its hidden layer and output layer outputs are:

$$a_h = f_h(w_h \cdot x),$$

$$a_{out} = f_{out}(w_{out} \cdot x),$$

where w is the weight and f is the activation function.

Then, NeuCF is defined as follows.

The one-hot codings of the user and item ID are input and mapped into word vectors through the embedding layer. The corresponding user and item representation are obtained through GMF and MLP. The output of GMF and MLP is connected. The final prediction value is obtained through the Softmax function. The formula can be written as:

$$X^{GMF} = p_u^{GMF} \cdot q_i^{GMF},$$

$$X^{MLP} = f_L \left(w_L \left(\cdots f_1 \left(w_1 \left[p_u^{MLP} \right] + b_1 \right) \right) + b_L \right),$$

$$\hat{y}_{u,i} = f_{out} \left(w_{out} \left[X^{GMF} \right] \right),$$

where p_u^{GMF} is the user representation in GMF, q_i^{GMF} is the item representation in GMF, p_u^{MLP} is the user representation in MLP, q_i^{MLP} is the item representation in MLP, f_L is the activation function for the L-th layer of MLP, w_L is the weight for the L-th layer of MLP, b_L is the threshold for the L-th layer of MLP, and $\hat{y}_{u,i}$ is the predicted value.

3.3 Improved neural collaborative filtering algorithm

In the recommendation process, the traditional NeuCF only considers the ID of the user and the item, without

taking into account any additional information. For example, in the educational learning platform resources, user A likes item B but does not like item C, which can only be expressed as user A likes item B because item B is item B and does not like item C because item C is item C. In fact, user A is a student majoring in Chinese language and literature. The student likes item B because it is "Jin Yong's novel research" and dislikes item C because it is "C language programming", i.e., in addition to the ID of the user and the item, there are much other information that will affect the recommendation results. In addition, GMF may produce negative values after decomposition, which has no practical significance in the recommendation. To improve these two points, this paper improves the NeuCF algorithm to obtain the INeuCF algorithm. Firstly, the expanded generalized matrix factorization (EGMF) is obtained by expanding GMF. The formula can be written as:

$$\begin{cases} \delta(p_a, q_b) = p_a \cdot q_b \\ p_a = p v_A, p \geq 0, & v_A \geq 0 \\ q_b = q v_B, q \geq 0, & v_B \geq 0, \\ \hat{y}_{u,i} = f(w(\delta)) \\ st: w \geq 0, \text{diag}(w) = 0 \end{cases}$$

The mathematical principle of EGMF is approximately decompose the original utility matrix into the form of multiplication of two low-rank matrices, $R = P^T Q$, where P and Q satisfy non-negative constraints. In the equations, p_a and q_b are the implicit factor matrices of the user and item, v_A and v_B represent the user and item. In EGMF, it is limited that p , q , and weight w must be non-negative, i.e., the positive correlation between users and items. $\text{diag}(w)$ is the diagonal element of w , and its value is limited to 0, which can avoid the tendency to recommend the original item itself. The other steps are consistent with GMF. The structure of EGMF is shown in Figure 1.

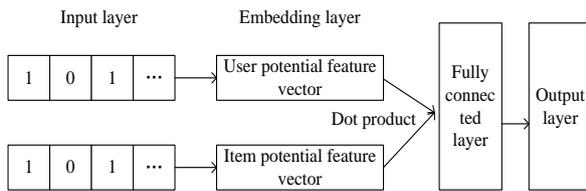


Figure 1: The structure of EGMF

As shown in Figure 1, the input layers are the one-hot codings of users and item IDs, which are then converted into vectors by calling embedding in the embedding layer. The inner product of user and item vectors is calculated, and finally, the prediction result is output based on the sigmoid function in the fully connected layer.

Given the limitations of information, this paper combines some other information based on user and item ID to make some changes to the input of the INeuCF algorithm. In the user (learner) information, in addition to ID, gender, major, and occupation are combined. In the item (resource) information, in addition to ID, resource

categories and subcategories of resources are combined. Examples are shown below.

User information: User ID, male, computer application technology, front-end engineer

Item information: Item ID, computer, programming and development

The calculation process of the INeuCF algorithm can be written as:

$$\begin{aligned} X^{EGMF} &= (l_0^{EGMF} \oplus l_1^{EGMF} \oplus \dots)(c_0^{EGMF} \oplus \\ &c_1^{EGMF} \oplus \dots), \\ X^{MLP} &= f_L \left(w_L \left(\dots f_1 \left(w_1 \left[l_0^{MLP} \oplus l_1^{MLP} \oplus \dots \right] \oplus \right. \right. \right. \\ &\left. \left. \left. c_0^{MLP} \oplus c_1^{MLP} \oplus \dots \right] \right) \right) + b_L, \\ \hat{y}_{u,i} &= f_{out} \left(w_{out} \left[X^{EGMF} \right] \right), \end{aligned}$$

where l_x^{EGMF} represents the user (learner) vector (it is user ID if $x = 0$ and other information such as gender if $x \neq 0$), c_x^{EGMF} represents item (resource) vector (it is resource ID if $x = 0$ and other information such as resource category if $x \neq 0$), and \oplus stands for connection, indicating the linkage between the user vector and other information vectors (allowing for flexible addition of various information).

4 Results and analysis

4.1 Experimental data

The development tool used for the experiment was Pycharm, and the programming language used was Python 3.6. Data were collected from the massive open online courses (MOOC) educational learning platform of China University, including the interaction data of 3,678 users to 812 learning resources from 2018 to 2022. The number of interactions of each user was more than 20. The interaction between users and resources was recorded as 1, indicating positive samples, while the absence of interaction was recorded as 0, indicating negative samples.

The user information included user ID, gender, major, and occupation.

The resource information included resource ID, resource category, and resource subcategory. The resource category included computer, foreign language, science, engineering, and agriculture, etc. The resource subcategory took the computer category as an example, and the computer category included big data and artificial intelligence, program design and development, software engineering, hardware and software systems, and principles.

The five-fold cross-test method was used in the experiment, and the final results were averaged.

4.2 Evaluation index

Suppose that N items recommended by the algorithm for user u is $R(u)$, and the item actually liked by the user is

$T(u)$. The following indicators were used to evaluate the recommendation effect of the algorithm.

(1) Precision: the proportion of the items that the user actually likes in the results:

$$P = \frac{\sum_u |R(u) \cap T(u)|}{\sum_u R(u)}$$

(2) Recall rate: the proportion of items that users like out of all the liked items:

$$R = \frac{\sum_u |R(u) \cap T(u)|}{\sum_u T(u)}$$

(3) F1 value: the combined evaluation of the precision and recall rate:

$$F_1 = \frac{2PR}{P+R}$$

(4) Normalized discounted cumulative gain (NDCG) [11]: it reflects the degree of coincidence between the recommendation list obtained by the algorithm and the actual list:

$$NDCG = \frac{\sum_{n=1}^N \frac{2^{rel_i-1}}{\log_2(n+1)}}{\sum_{n=1}^N \frac{2^{rel_i-1}}{\log_2(n+1)}}$$

where rel_i stands for the relevance of the recommendation result of position i . If it hits, then $rel_i = 1$; if it does not hit, then $rel_i = 0$.

4.3 Recommendation results

The INeuCF algorithm was compared with the following methods:

- (1) user-CF,
- (2) item-CF,
- (3) sparse linear method (SLIM) [12],
- (4) factorization machine (FM) [13],
- (5) nonnegative matrix factorization (NMF) [14],
- (6) NeuCF,
- (7) INeuCF.

Firstly, under different lengths of recommended lists, the results are presented in Table 1.

Table 1: Comparison of recommendation effects

N=5	P	R	F1	NDCG
User-CF	0.167	0.212	0.187	0.207
Item-CF	0.177	0.214	0.194	0.208
SLIM	0.181	0.217	0.197	0.215
FM	0.186	0.223	0.203	0.276
NMF	0.197	0.227	0.211	0.297
NeuCF	0.205	0.232	0.218	0.307
INeuCF	0.219	0.245	0.231	0.315
N=10	P	R	F1	NDCG
User-CF	0.115	0.276	0.162	0.211
Item-CF	0.117	0.287	0.166	0.212
SLIM	0.181	0.217	0.197	0.215
FM	0.143	0.297	0.193	0.289
NMF	0.151	0.307	0.202	0.305
NeuCF	0.162	0.312	0.213	0.321
INeuCF	0.168	0.351	0.227	0.337

1

Table 1 shows that the traditional user-CF and item-CF methods had poor performance in personalized

resource recommendation. When $N=5$, the F1 values were 0.187 and 0.194, respectively, and the NDCG values were only 0.207 and 0.208, indicating that the traditional CF algorithm could not effectively explore learners' preferences for resources, leading to a poor recommendation effect. Compared to the user-CF and item-CF methods, the SLIM showed some improvement in personalized recommendation of resources. However, the improvement was not significant. When $N=5$, its F1 value was 0.197, and the NDCG value was 0.215. When $N=10$, its F1 value was 0.171, and the NDCG value was 0.225. Moreover, the recommendation performance of the FM and NMF was slightly improved. When $N=5$, the F1 values of the FM and NMF methods were 0.203 and 0.211 respectively, and the NDCG values were 0.276 and 0.297 respectively, but the improvement effect was not obvious. The F1 value of the NeuCF algorithm was 0.218, and the NDCG value was 0.307 when $N=5$, which were better than the previous four algorithms, indicating that the recommendation algorithm combined with deep learning showed obvious advantages in personalized resource recommendation and could effectively improve the recommendation effect. Finally, the F1 value of the INeuCF algorithm reached 0.231 when $N=5$, which was 0.013 higher than the NeuCF algorithm, and the NDCG value was 0.35, which was 0.008 higher than the NeuCF algorithm. These results proved the effectiveness of the improvement of NeuCF.

From the point of view of different lengths of recommendation lists, the P value and F1 value of $N=5$ was slightly higher than those of $N=10$, and the R value and NDCG value of $N=10$ was slightly higher than those of $N=5$. With the increase in the length of the recommendation list, the number of resources in the results obtained by the algorithm and the actual results increased, so this change occurred. But in general, under the same N value, the results of the INeuCF algorithm were better.

When $N=10$, the performance of the INeuCF algorithm was further analyzed by comparing the improvement of the recommendation effect after different optimizations, and the results are presented in Table 2.

Table 2: Analysis of different improvements of the algorithm2

Algorithm	Input	F1	NDCG
MLP	User and item ID only	0.185	0.301
GMF	User and item ID only	0.189	0.307
EGMF	User and item ID only	0.205	0.309
NeuCF (MLP+GMF)	User and item ID only	0.213	0.321
INeuCF (MLP+EGMF)	User and item ID only	0.221	0.326
INeuCF (MLP+GMF)	ID+ other information	0.223	0.328

INeuCF (MLP+EGMF)	ID+ other information	0.227	0.337
----------------------	--------------------------	-------	-------

According to Table 2, when only using MLP or only GMF, the recommendation effect was poor, with F1 values below 0.2 and NDCG of 0.301 and 0.307, respectively. The comparison between the GMF and EGMF showed that the F1 value of the EGMF was 0.205, which was 0.016 higher than that of the GMF, and the F1 value of NDCG was 0.309, which was 0.307 higher than that of the GMF, which proved the effectiveness of the improvement of the GMF. The NeuCF algorithm is a recommendation algorithm that combines MLP and GMF, and its F1 and NDCG values were 0.213 and 0.321, respectively, which were significantly higher than those of the MLP, GMF, and EGMF. These results showed that learning linear and nonlinear features simultaneously could significantly improve the performance of recommendation algorithms.

For the improvement of the NeuCF algorithm, if only the GMF was improved to the EGMF and user and item ID were still used as input, the F1 value was 0.221, which was 0.008 higher than the NeuCF algorithm, and the NDCG value was 0.326, which was 0.005 higher than the NeuCF algorithm. If MLP+GMF was used, but the input was improved by combining other information, the F1 value was 0.223, which was 0.01 higher than the NeuCF algorithm, and the NDCG value was 0.328, which was 0.007 higher than the NeuCF algorithm. These results showed that the addition of other information had a more obvious effect on improving the recommendation effect.

Finally, in the INeuCF algorithm, where the GMF was improved to the EGMF and additional information was added to the input, both the F1 value and NDCG values showed improvement compared to the previous methods, with values of 0.227 and 0.337 respectively. The results demonstrated that the designed method was reliable in the personalized recommendation of resources and could be applied on the actual educational learning platform.

5 Discussion

In educational learning platforms, there are abundant learning resources that greatly facilitate students' autonomous learning. However, the existence of a massive amount of resources also poses certain limitations for resource retrieval and recommendation. While recommendation algorithms have been widely applied in fields such as e-commerce and social media, limited research has been conducted on their application in the education sector. Moreover, educational learning platforms are emerging and constantly evolving, rendering traditional recommendation algorithms unsuitable. Therefore, in response to the resource recommendation problem on educational learning platforms, this paper designed an INeuCF algorithm based on deep learning and collected data from actual educational learning platforms to verify the performance of the proposed method.

In order to demonstrate the reliability of the method proposed in this paper, a comparison was made with some commonly used recommendation algorithms on the same

dataset. It can be observed that as the length of the recommendation list increased, different recommendation algorithms showed an increase in precision but a decrease in recall rate. When comparing algorithms under equal recommendation list lengths, it is evident that the designed INeuCF algorithm outperformed others in terms of P value, R value, F1 value, and NDCG value. This result confirmed the reliability of this approach for resource recommendations on educational learning platforms.

Then, a specific analysis was conducted on the proposed improvement method. It was found that the recommendation performance was effectively improved by transforming GMF into EGMF. Furthermore, incorporating additional user information during input further enhanced the reliability of the recommendation results, and outcomes that aligned more closely with the actual recommended list were obtained.

Based on the comprehensive experimental results, the INeuCF algorithm is effective and reliable. It has certain advantages compared to other current recommendation algorithms and can be applied on practical educational learning platforms. Based on the INeuCF algorithm, it can provide more accurate and effective recommendations of resources based on learners' relevant information, thereby better stimulating learners' motivation to learn and improving learning outcomes. In cases where learner satisfaction is high, it can also promote further development of educational learning platforms, attract more learners and educators, expand educational resources further, and provide learners with more learning choices.

6 Conclusion

To solve the problem of resource recommendation on educational learning platforms, this paper designed an INeuCF algorithm, collected data on the actual educational learning platform, and conducted experimental analysis. The results showed that compared with some other recommendation algorithms, the recommendation results obtained by the INeuCF algorithm were closer to the actual recommendation list, and the F1 value and NDCG value were higher. It can be implemented within the existing educational learning platform to facilitate learners in locating resources of their interest more effectively.

References

- [1] Wang W, Guo L, Sun R (2019). Rational herd behavior in online learning: Insights from MOOC. *Computers in Human Behavior*, 92, pp. 660-669. <https://doi.org/10.1016/j.chb.2017.10.009>.
- [2] Gu T, Liang H, Bin C, Chang L (2021). Combining user-end and item-end knowledge graph learning for personalized recommendation. *Journal of Intelligent and Fuzzy Systems*, 40, pp. 9213-9225. <https://doi.org/10.3233/JIFS-201635>.
- [3] Jing Z (2021). Application of Data Mining Based on Computer Algorithm in Personalized Recommendation Service of University Smart

- Library. *Journal of Physics: Conference Series*, 1955, pp. 1-8. <https://doi.org/10.1088/1742-6596/1955/1/012008>.
- [4] Ohtomo K, Harakawa R, Ogawa T, Haseyama M, Iwahashi M (2021). Personalized Recommendation of Tumblr Posts Using Graph Convolutional Networks with Preference-aware Multimodal Features. *ITE Transactions on Media Technology and Applications*, 9, pp. 54-61. <https://doi.org/10.3169/mta.9.54>.
- [5] Gan M, Kwon O C (2022). A knowledge-enhanced contextual bandit approach for personalized recommendation in dynamic domains. *Knowledge-based Systems*, 251, pp. 1-11. <https://doi.org/10.1016/j.knosys.2022.109158>.
- [6] Alnahhas A, Alkhatib B (2020). Augmented Context-Based Conceptual User Modeling for Personalized Recommendation System in Online Social Networks. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 14, pp. 1-19. <https://doi.org/10.4018/IJCINI.2020070101>.
- [7] Prathama F, Senjaya W F, Yahya B N, Wu J Z (2020). Personalized Recommendation by Matrix Co-Factorization with Multiple Implicit Feedback on Pairwise Comparison. *Computers & Industrial Engineering*, 152, pp. 1-14. <https://doi.org/10.1016/j.cie.2020.107033>.
- [8] Houshmand-Nanehkaran F, Lajevardi S M, Mahlouji-Bidgholi M (2022). Optimization of fuzzy similarity by genetic algorithm in userbased collaborative filtering recommender systems. *Expert Systems: The International Journal of Knowledge Engineering*, 39, pp. 1-27.
- [9] Murty C S V V S N, Varma G, Satyanarayana C (2022). Content-Based Collaborative Filtering with Hierarchical Agglomerative Clustering Using User/Item based Ratings. *Journal of Interconnection Networks*, 22, pp. 2141026:1-2141026:18. <https://doi.org/10.1142/s0219265921410267>.
- [10] He X, Liao L, Zhang H, Nie L, Hu X, Chua T S (2017). Neural Collaborative Filtering. *Proceedings of the 26th International Conference on World Wide Web*, 2017, pp. 173-182. <https://doi.org/10.1145/3038912.3052569>.
- [11] Phonexay V, Xinchang K, Park D (2020). Movie Recommendation System Based on Users' Personal Information and Movies Rated Using the Method of k-Clique and Normalized Discounted Cumulative Gain. *Journal of Information Processing Systems*, 16, pp. 494-507. <https://doi.org/10.3745/JIPS.04.0169>.
- [12] Ning X, Karypis G (2011). SLIM: Sparse Linear Methods for Top-N Recommender Systems. *2011 IEEE 11th International Conference on Data Mining*, 2011, pp. 497-506. <https://doi.org/10.1109/ICDM.2011.134>.
- [13] Zhang J, Wu Z, Li F, Li W, Ren T, Li W, Chen J (2021). Deep Attentional Factorization Machines Learning Approach for Driving Safety Risk Prediction. *Journal of Physics: Conference Series*, 1732, pp. 1-8. <https://doi.org/10.1088/1742-6596/1732/1/012007>.
- [14] Wang W, Wang S, Qin D, Fang Y, Zheng Y (2023). Heart-lung sound separation by nonnegative matrix factorization and deep learning. *Biomedical Signal Processing and Control*, 79, pp. 772-784. <https://doi.org/10.2139/ssrn.4017034>.