Hindawi
Scientific Programming
Volume 2022, Article ID 3199134, 10 pages
https://doi.org/10.1155/2022/3199134

Research Article

Deep Collaborative Online Learning Resource Recommendation Based on Attention Mechanism

Cuiping Hao and Ting Yang

Yulin University, Yulin 719000, Shannxi, China

Correspondence should be addressed to Ting Yang; yt1028@yulinu.edu.cn

Received 6 January 2022; Revised 14 February 2022; Accepted 1 March 2022; Published 24 March 2022

Academic Editor: Hangjun Che

In view of the lack of hierarchical and systematic resource recommendation caused by rich online learning resources and many learning platforms, an attention-based ADCF online learning resource recommendation model is proposed by introducing the attention mechanism into a deep collaborative DCF model. Experimental results show that the proposed ADCF model enables an accurate recommendation of online learning resources, reaching 0.626 and 0.339 on the HR and NDCG metrics, respectively, compared to the DCF models before improved, up by 1.31% and 1.25%, and the proposed ADCF models by 1.79%, 2.17%, and 2.32%, respectively, compared to the IUNeu and NeuCF models.

1. Introduction

In recent years, the application of the Internet has promoted education from offline to online. Online education has become a trend in today's education development and occupies a place in the huge education market. Data show that the number of online education users in China's in 2020 grew from 90.992 million in 2015 to 182.492 million, an increase of 50.14%. With the infiltration of online education methods, online education and learning resources are increasingly rich, which not only brings the balanced development and multidirectional development of education but also brings great difficulties to the recommendation of learning resources. On the one hand, there are more online education platforms, and the curriculum types of each platform are complex and the curriculum quality is poor, making it difficult for learners to choose by themselves; on the other hand, online education and learning resources lack understanding of learners, which makes it difficult for the platform to recommend personalized courses to learners. Therefore, it is necessary to integrate and analyze the online learning resources and recommend them according to learners' needs. To achieve this purpose, the relevant scholars conduct research. According to the different characteristics of students, Hu and others select a small piece of content course knowledge points in the learning resources, take the course knowledge as the recommendation point, and realize the recommendation of personalized online learning resources by designing personal personalized learning mechanism recommendation [1]. In order to solve the problem of sparse data and poor scalability in collaborative filtering algorithms, Honggang Wang and others optimized them using dynamic k close neighbors and slope one algorithms, and analyzed the sparsity of learning resource data in the network based on the neighbor selection results. Two-way self-equilibrium of stage evolution is adopted to improve the personalized recommendation of resource push, and the fuzzy adaptive binary particle group optimization algorithm based on evolutionary state judgment is adopted to solve the optimal sequence recommendation problem, thus to realize the personalized recommendation of learning resources and improve the matching degree and recommendation speed of online learning resources [2, 3]. Yuan studies link prediction methods in network education, builds a suitable model for network education, and proposes an improved path sorting algorithm based on the neural network sorting method through an improved analysis of traditional methods.

Meanwhile, the random-walk model and the neural network-path sorting algorithm are used to realize the link prediction problem in the online learning knowledge base [4]. Xie et al. proposed a user interaction-based recommendation framework that explores real-time interest from immediate feedback, and experiments on real datasets show that the algorithm achieves more accurate prediction results and higher recommendation efficiency [5]. Liang et al. proposed a learning style model, AROLS model, to represent the characteristics of online learners, realize learning resource adaptation by mining behavioral data of learners, and improve the recommendation effect of online learning resources [6]. Antequera et al. present a novel approach to providing fast, automatic, and flexible resources for application owners with limited expertise in building and deploying appropriate cloud architectures; compared to existing schemes, the scheme improves resource recommendation accuracy in manufacturing scientific gateway applications by 21% [7, 8]. Through the above research, it can be found that the existing online learning resource recommendation is limited to the recommendation of courses on their own platforms and does not integrate and comprehensively recommend similar courses on other platforms, resulting in often the accurate recommendation of courses to learners. To solve this problem, this paper constructs an ADCF recommendation model based on the existing DCF model, by integrating the whole-platform online learning resources and introducing the attention mechanism to allocate the weight to the resources so as to realize the accurate recommendation of online learning resources.

2. Basic Approach

2.1. Brief Introduction of the DCF Model. The DCF model is a deep learning collaborative filter recommendation model developed based on the neural collaborative filter recommendation (NeuCF) model, which solves the problem of few input feature types of the NeuCF model and effectively improves the feature combination ability and nonlinear ability of the model. Its basic architecture is similar to that of NeuCF model architecture, mainly composed of multilayer perceptron (MLP) and generalized matrix decomposition (GMF), including input layer, coding layer, embedding layer, embedding layer, feature extraction layer, pooling layer, feature splicing layer, neural network layer, and output layer, as shown in Figure 1 [9, 10].

The model input layer is mainly responsible for inputting the relevant information and its auxiliary information into the model for training. The coding layer is responsible for coding input feature information. The embedding layer is responsible for converting the encoding into a corresponding feature representation, often including two types for the linear model GMF and for the nonlinear model MLP. The feature extraction layer is responsible for extracting feature relationships, where GMF is used to extract linear feature relationships, and MLP is used to extract nonlinear relationships [10]. The pooling layer is responsible for adjusting the feature size to facilitate feature splicing, including maximum pooling and mean pooling [11]. The feature splicing layer is responsible for integrating the feature information extracted from GMF and MLP, mainly by adding or splicing the extracted features. The neural network layer is responsible for training the model and model parameter tuning, and uses the cross-entropy loss function to adjust the network weight. The output layer maps the output value to a specific range through the activation function. The sigmoid function is selected as the activation function, and its mathematical expression is as follows [12]:

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{1}$$

In the formula, the value range of $f(x)$ is [0, 1].

The DCF model can extract nonlinear and linear features of information and has strong personalized recommendation ability, but the model believes that all features have the same impact on the final recommendation results and do not have the importance of distinguishing between different features. In practice, different characteristic factors contribute differently to the model prediction results, so it is necessary to improve the model. In this paper, the recommendation effect of the model is improved by introducing an attention mechanism into the model to distinguish between the importance of different features.

2.2. DCF Model Refinement. The improvement of the DCF model in this paper is the introduction of the attention mechanism into the model to improve the model recommendation effect by adding the importance of distinguishing attention layers between the feature splicing layer and the neural network layer. The attention mechanism in the DCF model is mainly attached to the framework of the encoder and decoder, and the essence is a thought model, formulas (2) ~ (4), by allowing the neural network in the DCF model to only focus on the partial information of the input and select specific inputs [13].

$$a = f(X_N), \tag{2}$$

$$Z_a = a \Theta X_n, \tag{3}$$

$$\text{st: } a \in (0, 1). \tag{4}$$

In the formula, X_n represents the n-dimensional eigenvector input attention layer, $f(x)$ indicates the attention mechanism network, a represents the attention corresponding to the n-dimensional eigenvector through $f(x)$, and Z_a represents the output layer output result. At that time, $f(x) = \text{Soft}\max(X_n)$, the value range of a was (0, 1); when a is constant 1, the attention network helps fit the complex function model when the model degenerated into a DCF model.

The attention mechanism layer of the DCF model (hereinafter referred to as the ADCF model) obtains the weight of each feature dimension through equations (5) ~ (7) [14]. Computing the attention of each feature dimension using softmax, then interacting with the corresponding

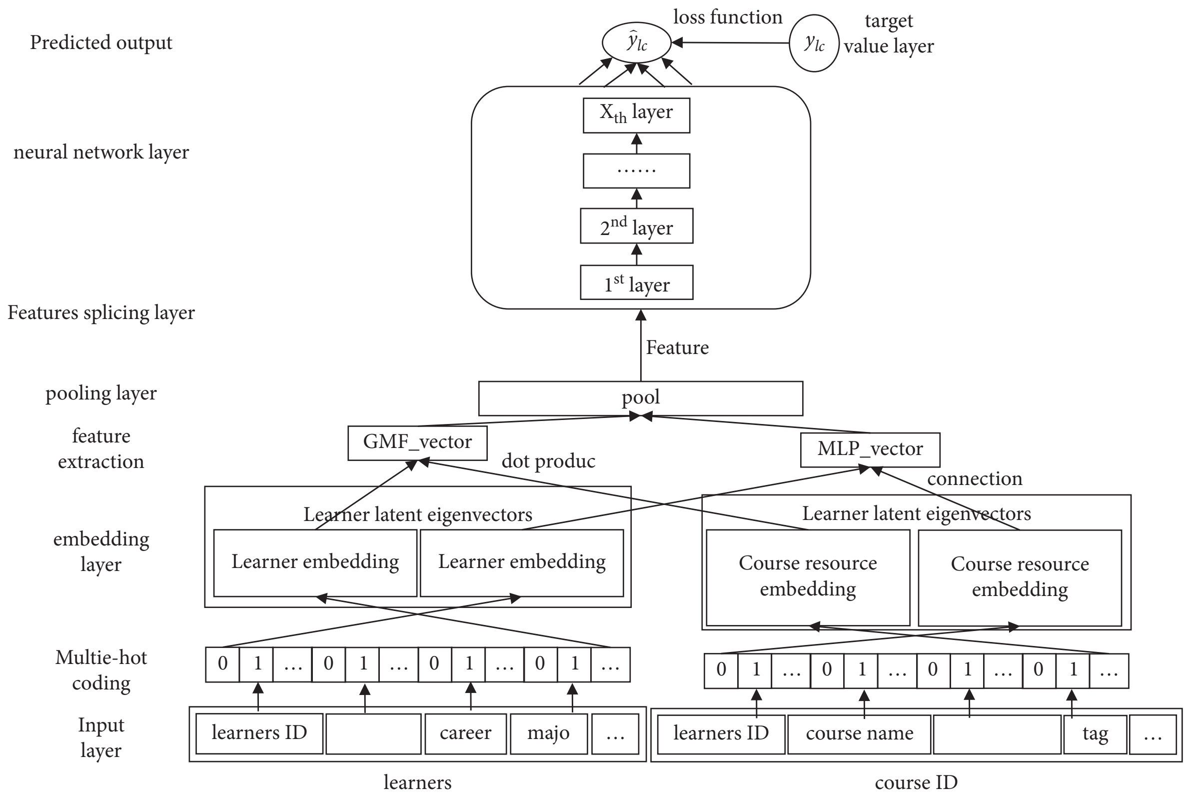

Figure 1: The DCF model structure.

features for a point-multiply input deep neural network, and finally the predictive values can be obtained through iterative training.

$$\text{Soft max}\left(Z_i\right) = \frac{e^{z_i}}{\sum_{i=1}^{k} e^{z_i}}, \tag{5}$$

$$A_n = \text{Soft max}\left(X_n\right), \tag{6}$$

$$A_{\text{out}} = A_n \Theta X_n. \tag{7}$$

In the formula, Z_i represents the ith element input by the softmax function, soft max (Z_i) represents the corresponding softmax value of this element, and X_n represents the output value of the feature splicing layer. ⊙indicates point multiplication; A_n represents the corresponding value calculated by softmax, and A_{out} represents the attention layer output value.

The ADCF model, by introducing the attention layer into the DCF model, can distinguish the degree of contribution of different feature information to the predicted recommendation results and then improve the model recommendation effect. Therefore, this paper proposes a new online resource recommendation method based on the ADCF model.

3. Recommended Methods for Online Learning Resources Based on ADCF

3.1. ADCF Mathematical Model. The mathematical model of the ADCF model is as follows:

$$\begin{aligned}
P_c^{GMF} &= \left(c_0^{GMF} \oplus c_1^{GMF} \oplus c_2^{GMF} \oplus c_3^{GMF}\right),\\
P_l^{GMF} &= \left(l_0^{MLP} \oplus l_1^{MLP} \oplus l_2^{MLP} \oplus l_3^{MLP}\right),\\
q_c^{\text{MLP}} &= \left(c_0^{MLP} \oplus c_1^{MLP} \oplus c_2^{MLP} \oplus c_3^{MLP}\right),\\
\varphi^{GMF} &= P_l^{GMF} \Theta q_c^{GMF},\\
\varphi^{MLP} &= \partial_L W_L^T\left(\partial_{L-1}\left(\ldots\partial_2\left(W_2^r\begin{bmatrix} p_l^{MLP} \\ q_c^{MLP}\end{bmatrix} + b_2\right)\ldots\right)\right) + b_L,\\
A_{\text{out}} &= \text{Soft max}\left(\begin{bmatrix} \varphi^{EGMF} \\ \varphi^{MLP}\end{bmatrix}\right)\Theta\begin{bmatrix} \varphi^{EGMF} \\ \varphi^{\text{MLP}}\end{bmatrix},\\
\hat{y}_k &= \sigma\left(h^T A_{\text{out}} + b\right).
\end{aligned} \tag{8}$$

In the formula, ⊕ represents connections; liGMF and liMLP represent the learner eigenvectors of the GMF and MLP models, $\{l_0, l_1, l_2, l_3\}$, respectively, corresponding to {learner_id, sex, profession, job}; ciGMF and ciMLP represent the course resource eigenvectors of the models, $\{c_0, c_1, c_2, c_3\}$, respectively, corresponding to {curriculum_id, name, complexity, label}; the piGMF, piMLP, qcGMF, and qcMLP represent the learner and course resource feature vector with auxiliary information for the GMF and MLP models; φ^{MLP} and φ^{EGMF}, respectively, represent the underlying feature relationships learned through the MLP and GMF models. From the model, the whole model process uses the learner and course auxiliary information, which jointly determines the prediction recommendation performance of the model.

3.2. Online Learning Resource Recommendation Based on ADCF. From the above analysis, the online learning resource recommendation model based on ADCF is designed, as shown in Figure 2. The model input layer includes the relevant information of the learner and the course, whose input form is the behavioral sequence of the learner-course $X = \{X_1, X_2, X_3, \ldots X_N\}$. Among these, X_i represents the ith behavior $\langle \mathrm{Learner_i}, \mathrm{Course_i} \rangle$. $\mathrm{Learner_i}$ contains information about learners' ID, gender, occupation, and $\mathrm{Course_i}$ contains course ID, label, name, etc.

The encoding layer transforms input information via one hot to embedding feature demo [15]. Taking learner sex as an example, male is coded as [1, 0] and female as [0, 1].

The embedding layer first initializes a embedding matrix about sex, with a matrix size of $2 * d$, where 2 indicates possible sex values and d indicates the embedding dimension. With the above operation, the intractable category feature is able to be converted into tractable vectors. For course and learner features, encoding includes course ID, name, complexity, tag, and learner ID, gender, major, occupation. Embedding corresponds to a length of 16, so the total embedding length is $16 * 8 = 128$ and the output matrix dimension is $1 * m * n$. Then, the auxiliary information $(m - 1)\cdot$n is added through the flattening operation to extend the length to $m \cdot n$, where m represents the input length, n indicates the potential feature length, and the embedding layer input scale extends from (1,1, n) to (1, m, n).

The feature extraction layer uses GMF and MLP to extract linear and nonlinear relationships between the learners and the course, the GMF model takes the implicit feature vector point multiplication of the user and the project as the output result, and the MLP model connects the flattening results beginning to end as the output results into the neural network. The activation function of the neural network adopts the ReLU function [16].

The pooling layer adopts the maximum pooling adjusted feature size so that the output vector of the GMF model and the MLP model reaches the same size through the pooling layer.

The feature splice layer was splicing to splicing the feature information extracted from the GMF and MLP models. The attention layer assigns the splice weights and inputs the assignment results into the neural network layer. The neural network layer loss function follows the cross-entropy loss function of the DCF model, adjusted for by backpropagation of the weights of each layer [17]. The output layer follows the sigmoid function as an activation function and maps the output values to a certain range for recommendation.

4. Simulation Experiment

4.1. Experimental Environment Construction. This experiment was carried out on the NVIDIA distributed framework and CUDA parallel computing platform with Selenium + Web Driver, installation of Chorme80.03987.132, compilation environment Python3.6, compiler PyCharm, and dependent module Requests2.21.0, Beautiful Soup, Selenuim, etc.

4.2. Data Preprocessing. Considering the large number of crawled datasets, computer-class-related data were selected to construct experimental datasets, and learners with interactions greater than 20 were selected as the main study subjects. Through statistical collation, 878 computer courses, 3,066 learners, and 203,987 learner history data were obtained. The distribution of course interaction number and number of learners is shown in Figure 3.

4.3. Dataset Construction. Online learning resource recommendation is actually a disclassification problem of predicting whether a learner will learn course [18]. Online learning resource recommendation is actually a disclassification problem of predicting whether a learner will learn course [18, 19]. If the missing samples in the dataset are taken by default to positive samples, it can easily lead to an unbalanced dataset. Therefore, to solve this problem, this paper selects some samples as negative samples by random uniform sampling from the missing values.

Considering the certain sequence relationship between learner history learning records, the first $n - 1$ records were used as the training set and the last effective interaction n as the test set. The model was trained by using the training set, and its performance was detected using the validation set. The validation set consists of 100 courses and splicing test set data randomly drawn from the no-interaction course. For experimental convenience, the batch size trained by the model is set to the effective number of interactions $n - 1$ per learner.

4.4. Evaluating Indicator. Hit Ratio, HR and Normalized Discounted Cumulative Gain, and NDCG were used as the indicators to evaluate model performance, calculated as formula (9) and formula (10) [20–22]:

$$\mathrm{HitRatio@}K = \frac{\mathrm{Number\ Of\ Hits@}K}{|GT|} \times 100\%, \tag{9}$$

$$NDCG_k = \frac{DCG_k}{IDCG_k}, \tag{10}$$

$$IDCG_k = \sum_{i=1}^{k} \frac{1}{1og_2^{1+i}}, \tag{11}$$

$$DCG_k = \sum_{i=1}^{k} \frac{2^{reli} - 1}{1og_2^{i+1}}, \tag{12}$$

$$CG_k = \sum_{i=1}^{k} reli. \tag{13}$$

In equation (9), HitRatio @K represents the hit rate, the number of tests predicted in the Top-k list per learner, and the denominator represents the number of test sets. In equation (10) ~ equation (13), i represents the position in the recommended list; k represents the k value in Top-k; CGk represents the cumulative gain; and *reli* represents the correlation of the

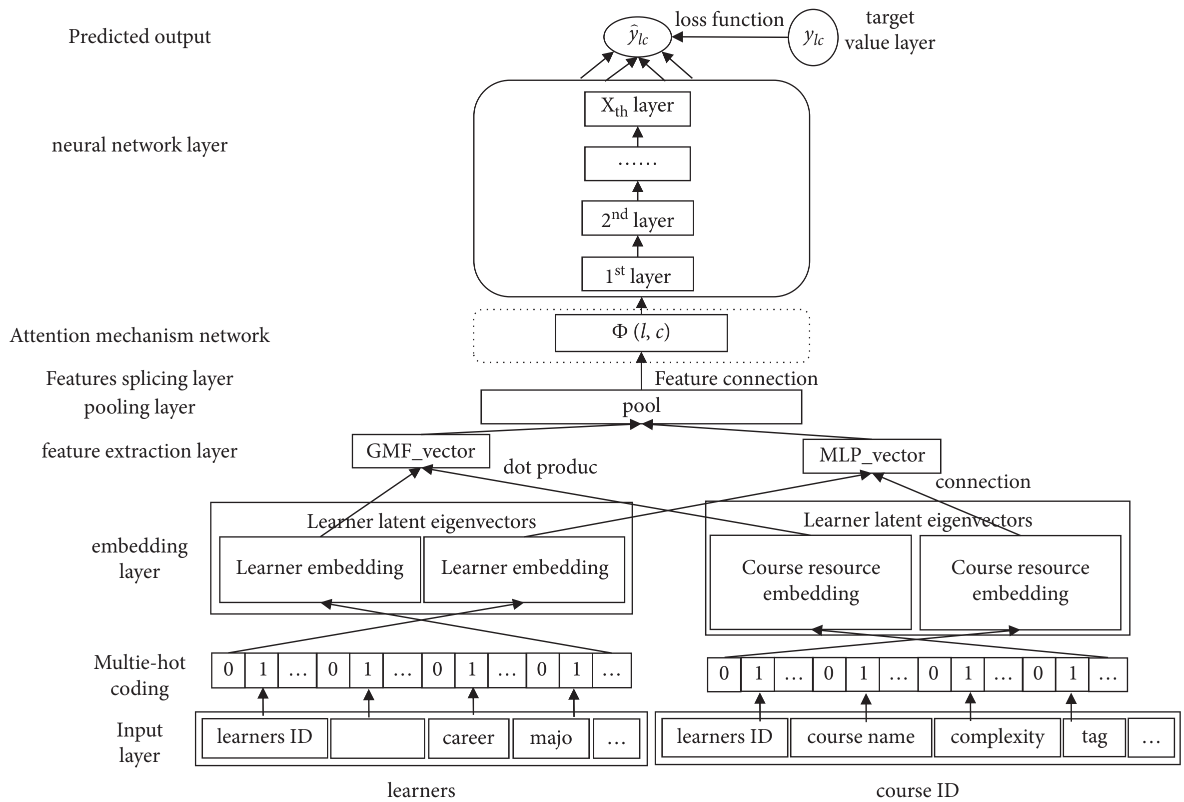

Figure 2: Resource recommendation based on ADCF.

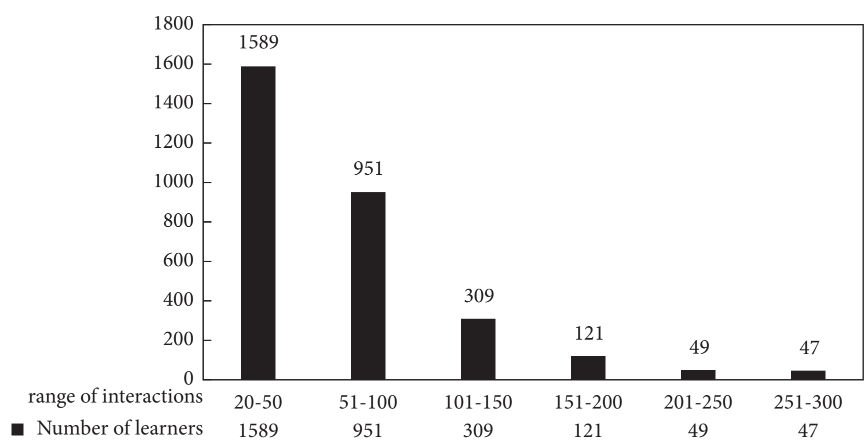

Figure 3: The number of learners distributed across the range of interaction numbers.

recommended result in position *i*. In this paper, *reli* = 1 indicates hits, and reli = 0 indicates misses. DCGk represents the cumulative damage gain, so when *reli* = 1, DCGk is calculated as formula (14) and when *reli* = 0, DCGk is calculated as formula (15). IDCGk represents the idealized loss gain, with all predicted hits, *reli* = 1, so the IDCGk calculation method can be rewritten as in formula (16).

$$DCG_k = \sum_{i-1}^{k} \frac{2^1 - 1}{1\log_2^{i+1}} = \sum_{i-1}^{k} \frac{1\log 2}{1\log(i+1)}, \tag{14}$$

$$DCG_k = \sum_{i-1}^{k} \frac{2^0 - 1}{1\log_2^{i+1}} = 0, \tag{15}$$

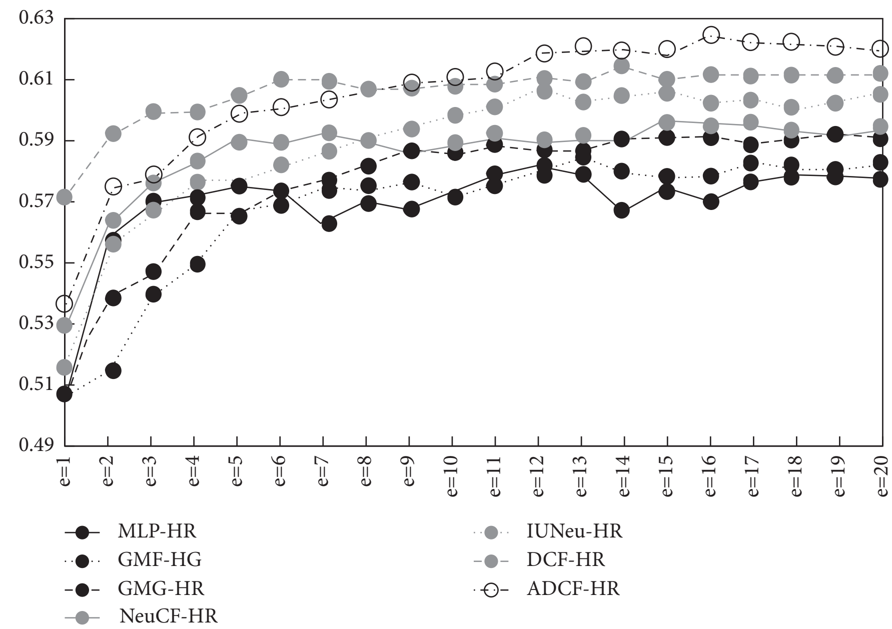

Figure 4: The effect of different iterations on model HR.

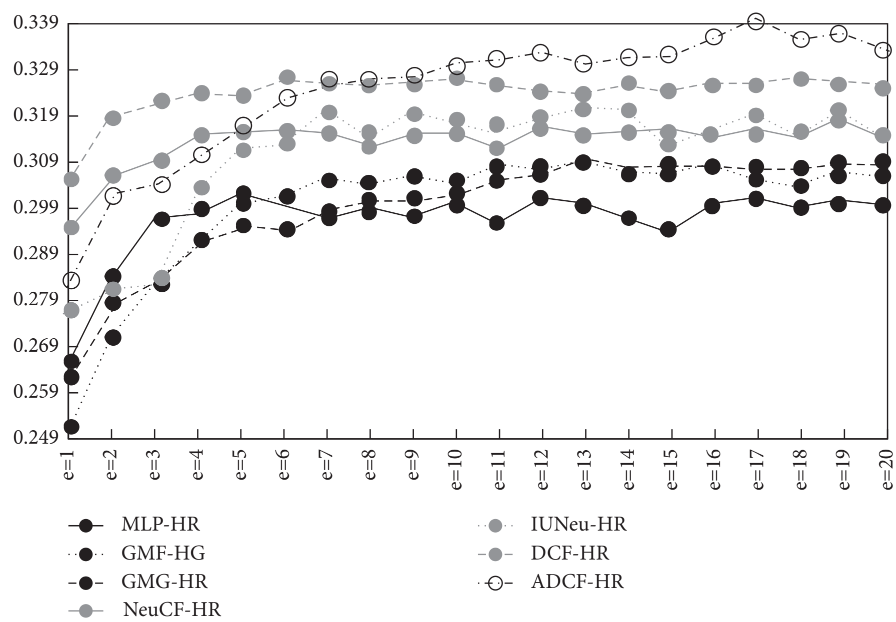

Figure 5: Effect of the different iterations on model NDCG.

$$IDCG_k = \sum_{i-1}^{k} \frac{1}{1\log_2^{1+i}}. \tag{16}$$

4.5. Parameter Setting. In this experiment, the model learning rate was set at 0.001, the number of iterations was 20, and the predicted number was $k = 10$. The model adopts the cross-entropy loss function, which is calculated by the formula as follows [23, 24]:

$$L(Y, P(Y|X)) = -1ogP(Y|X) = -\frac{1}{N}\sum_{i-1}^{N}\sum_{j-1}^{M} y_{ij} 1og\left(P_{ij}\right). \tag{17}$$

In the formula, X represents the input variable, Y the output variable, L the loss function, N the input sample size, and M the number of possible categories. The y_{ij} is a binary indicator indicating whether the category j enters the real category of the instance x_i. The p_{ij} represents the probability that the model predicts that the input instance x_i belongs to

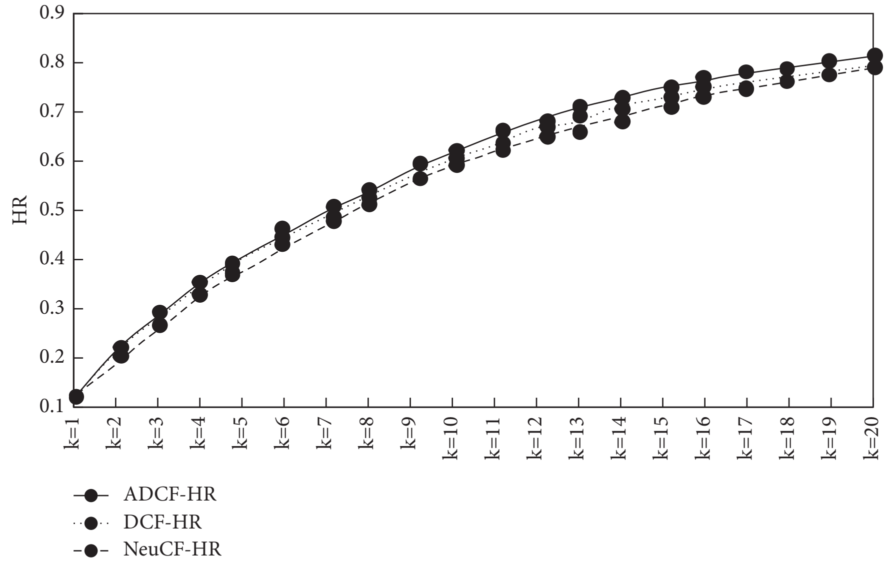

Figure 6: Index comparison of HR Model at different k values.

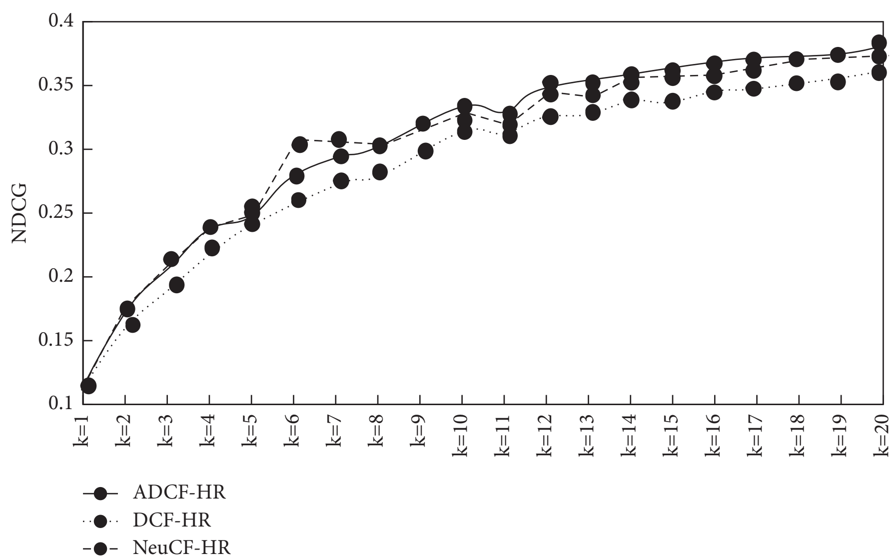

Figure 7: Comparison of the model NDCG indicators at different k values.

the category j. In this paper, the recommendation results include recommendations and disapproval, which are indicated by 1 and not by 0. Thus, formula (17) can be overwritten as

$$L(\text{loss}) = \frac{1}{N}\sum_{i-1}^{N}\left(y_i 1og p_i + \left(1 + y_i\right)1og\left(1 - p_i\right)\right). \tag{18}$$

In the formula, y_i represents the real category of the input instance x_i, p_i indicates the probability that the predicted input instance x_i belongs to category 1, and L (loss) represents the logarithmic loss average for each sample. The cross-entropy function is used to measure the similarity of y_i and p_i.

4.6. Experimental Result

4.6.1. Analysis of Parameters on Model Performance. This experiment was used to explore the influence of different parameters on the model performance. Figures 4 and 5 are the changes in the HR and NDCG indicators at the different number of iterations, respectively. According to the figure, with the same number of iterations, the DCF and ADCF models performed better on the HR and NDCG indicators compared to the IUNeu and NeuCF models, and the DCF and ADCF models also showed better and more stable performance with the number of iterations increasing. The proposed ADCF model adds the attention mechanism to extract the features and perform the best.

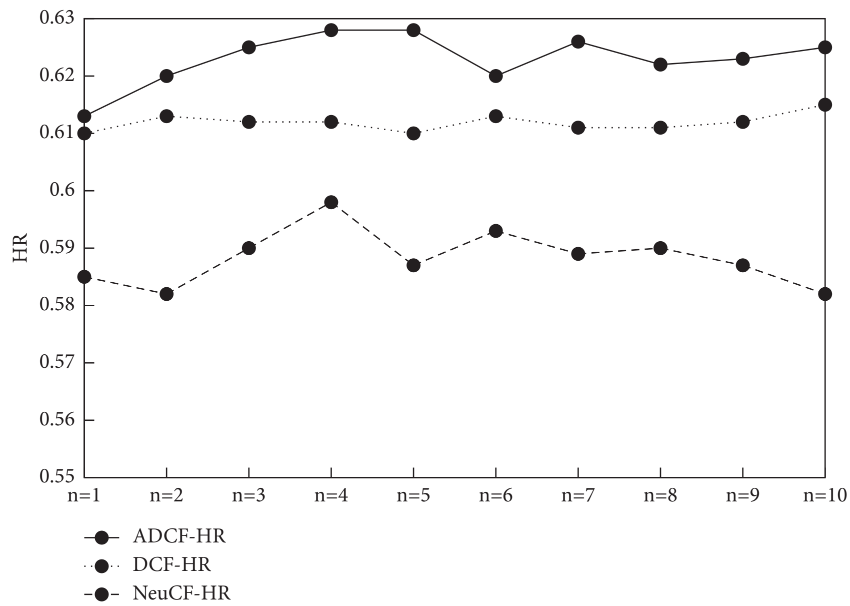

Figure 8: The effect of the different num-neg on model HR.

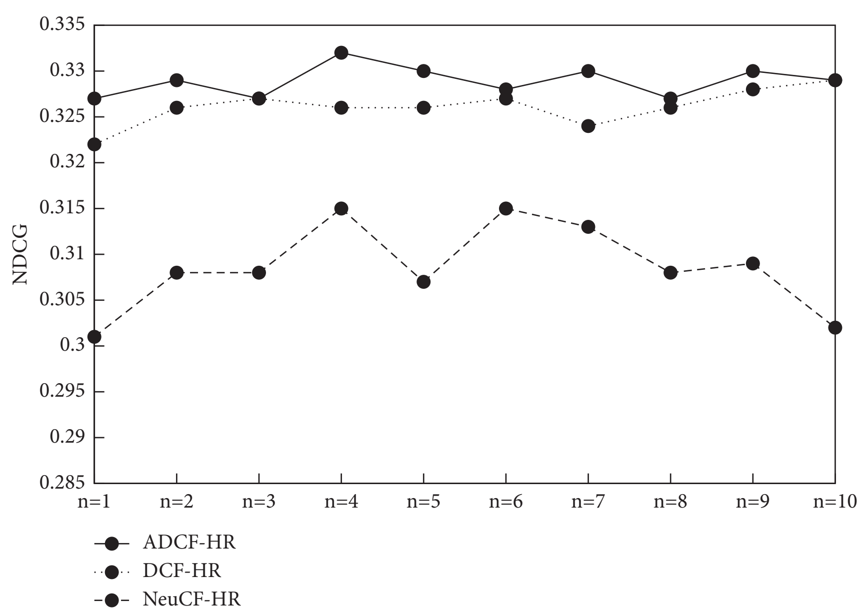

Figure 9: Effects of the different num-neg on model NDCG.

For different recommended courses *k*, the recommended performance of different models is Figures 6 and 7. As shown in the figure, increased with k values, the HR and NDCG metrices of the ADCF, DCF, and NeuCF models were gradually increased, showing that the model recommendation effect is getting better. At the same *k* values, the ADCF model outperformed the DCF model than the NeuCF and the IUNeu model and the ADCF model, showing that the ADCF model recommended the best effect. Thus, it shows that the proposed ADCF model recommends the best performance.

HR and NDCG indicators for different models under different negative sampled num-neg values are compared in Figures 8 and 9, respectively. From the figure, the value range of negative sampling *n* is [1, 10], and the recommended effect of different models is optimal when $n = 4$. Overall, the proposed ADCG model has smaller fluctuations in the HR and NDCG metrics as compared to the DCF and NeuCF models. This shows that the overall ADCF model generally has better recommended performance.

4.6.2. Model Performance Analysis. To validate the performance of the proposed model, the test results of different models such as IUNeu, NeuCF, and DCF were tested with the proposed models on the experimental dataset, as shown in Table 1. According to the table, the

Table 1: Comparison of the results of the different algorithms on the experimental datasets.

Serial number	Algorithm	HR@10	NDCG@10	Loading data time (s)	Time
1	ADCF	0.626	0.339	14.3	12.9 s (training) + 2.3 s (test)
2	DCF	0.613	0.326	12.5	12.4 s (training) + 1.8 s (test)
3	lUNeu	0.608	0.320	12.3	11.4 s (training) + 2.1 s (test)
4	NeuCF	0.598	0.317	11.1	11.2 s (training) + 1.4 s (test)
5	ConvNCF	0.546	0.292	16.9	16.9 s (training) + 2.6 s (test)

proposed model performs best on the HR and NDCG indicators, reaching 0.626 and 0.339, compared to the IUNeu model, respectively; HR and NDCG indicators increased 1.79% and 1.86%, respectively; compared with the NeuCF, it was increased by 2.17% and 2.32% on the HR and NDCG indicators, respectively; compared to the DCF models, it was increased by 1.31% and 1.25%, respectively.

In terms of time indicators, the proposed ADCF model has improved both the training time and the validation time of the dataset compared to the comparison model; compared to the DCF and NeuCF models, the average total training time increased by 1 s and 2.6 s, the average total validation time per iteration was increased by 0.5 s and 0.2 s, respectively. The reason for the analysis is that the proposed model introduces an attention mechanism in the characteristic splicing layer, so its temporal performance decreases, but the overall effect is small.

5. Conclusion

In this paper, an attention-based deep collaborative online learning resource recommendation method is proposed; it can realize the accurate recommendation of online learning resources. Compared with the proposed former DCF model, the proposed ADCF model was improved by 1.31% and 1.25% on the HR and NDCG indicators, respectively; compared with the IUNeu and NeuCF models, the proposed ADCF model was improved by 1.79% and 1.86%, 2.17%, and 2.32% on the HR and NDCG metrics, respectively, which has some practical application value. This paper presents a preliminary study of online learning resource recommendation, but the study is still in its infancy, and there are some problems to be improved. For example, when there is one feature fusion method, a splicing method is adopted, while there are many feature fusion methods, different fusion methods are suitable for different models. Therefore, multiple ways should be explored to choose the best fusion methods. Next step, it will be optimized from the above deficiencies to further improve the recommendation effect of online learning resources.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Hu, Y. Wang, Q. B. Chen, Q. Liu, and X. Fan, “Research on personalized learning based on collaborative filtering method,” *Journal of Physics: Conference Series*, vol. 1757, no. 1, pp. 012050–012055, 2021.

[2] H. Wang and W. Fu, “Personalized Learning Resource Recommendation Method Based on Dynamic Collaborative Filtering,” *Mobile Networks and Applications*, vol. 26, pp. 1–15, 2020.

[3] N. Sokolov Milovančević and A. Gračanac, “Time and ontology for resource recommendation system,” *Physica A: Statistical Mechanics and Its Applications*, vol. 525, pp. 752–760, 2019.

[4] I. Yuan, “Network education recommendation and teaching resource sharing based on improved neural network,” *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 1–10, 2020.

[5] Q. Xie, Y. Zhu, F. Xiong, L. Li, Z. Bao, and Y. Liu, “Interactive resource recommendation with optimization by tag association and significance analysis,” *Neurocomputing*, vol. 391, pp. 210–219, 2020.

[6] Q. Liang, X. Zheng, Y. Wang, and M. Zhu, “O3ERS: an explainable recommendation system with online learning, online recommendation, and online explanation,” *Information Sciences*, vol. 562, pp. 94–115, 2021.

[7] R. B. Antequera, P. Calyam, A. A. Chandrashekara, and R. Mitra, “Recommending heterogeneous resources for science gateway applications based on custom templates composition,” *Future Generation Computer Systems*, vol. 100, pp. 281–297, 2019.

[8] D. Mo, X.-G. Chen, S. Duan et al., “Personalized resource recommendation based on collaborative filtering algorithm,” *Journal of Physics: Conference Series*, vol. 1302, no. 2, p. 022025, 2019.

[9] Y. He, G. Xia, Y. Luo et al., “DVFENet: dual-branch voxel feature extraction network for 3D object detection,” *Neurocomputing*, vol. 459, pp. 201–211, 2021.

[10] L. Dai, X. Liu, Y. Liu et al., “Enhancing two-view correspondence learning by local-global self-attention,” *Neurocomputing*, vol. 459, pp. 176–187, 2021.

[11] Y. L. Wang, Z. J. Zhao, S. Y. Hu, and F. L. Chang, “CLCU-Net: cross-level connected U-shaped network with selective feature aggregation attention module for brain tumor segmentation,” *Computer Methods and Programs in Biomedicine*, vol. 207, pp. 106154-106155, 2021.

[12] Y. Shi, Q. Huang, and T. Hain, “H-VECTORS: improving the robustness in utterance-level speaker embeddings using a hierarchical attention model,” *Neural Networks*, vol. 142, pp. 329–339, 2021.

[13] W. Zheng, L. Yan, C. Gou et al., “Pay attention to doctor-patient dialogues: multi-modal knowledge graph attention image-text embedding for COVID-19 diagnosis,” *Information Fusion*, vol. 75, pp. 168–185, 2021.

[14] M. Chen, Y. Li, and X. Zhou, "CoNet: Co-occurrence neural networks for recommendation," *Future Generation Computer Systems*, vol. 124, pp. 308–314, 2021.
[15] J Zhenyan, M. Wu, H. Yang, and J. E. A. Iningo, "Temporal sensitive heterogeneous graph neural network for news recommendation," *Future Generation Computer Systems*, vol. 125, pp. 324–333, 2021.
[16] J. Wang, "An intelligent computer-aided approach for atrial fibrillation and atrial flutter signals classification using modified bidirectional LSTM network," *Information Sciences*, vol. 574, pp. 320–332, 2021.
[17] F. Xue, S. Pan, M. Yan et al., "Dynamic prediction of late noninvasive ventilation failure in intensive care unit using a time adaptive machine model," *Computer Methods and Programs in Biomedicine*, vol. 208, pp. 106290–106298, 2021.
[18] H. Lin, S. Xie, Z. Xiao, X. Deng, H. Yue, and K. Cai, "Adaptive recommender system for an intelligent classroom teaching model," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 14, no. 5, pp. 51–63, 2019.
[19] Y. Oh and Y. Kim, "A resource recommendation method based on dynamic cluster analysis of application characteristics," *Cluster Computing*, vol. 22, no. 1, pp. 175–184, 2019.
[20] H. Zhang, T. Huang, Z. Lv, S. Liu, and H. Yang, "MOOCRC: a highly accurate resource recommendation model for use in mooc environments," *Mobile Networks and Applications*, vol. 24, no. 1, pp. 34–46, 2019.
[21] H. Liu, "Resource recommendation via user tagging behavior analysis," *Cluster Computing*, vol. 22, no. 1, pp. 885–894, 2019.
[22] *Studies from Wuhan University of Technology Yield New Information about Algorithms (Interactive Resource Recommendation Algorithm Based on Tag information)*, Internet Weekly News, UK, pp. 382–386, 2018.
[23] Q. Xie, F. Xiong, T. Han, Y. Liu, L. Li, and Z. Bao, "Interactive resource recommendation algorithm based on tag information," *World Wide Web*, vol. 21, no. 6, pp. 1655–1673, 2018.
[24] J. H. Park, "Resource collaboration system based on user history and the psychological mode," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 5, pp. 1683–1691, 2018.