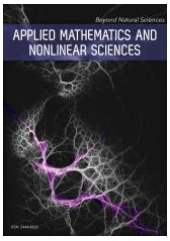
sciendo Applied Mathematics and Nonlinear Sciences, 9(1) (2024) 1-16



Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

The Effectiveness of Educational Resource Optimization Based on Big Data Technology on Enhancing Students' Creative Ability

Lijie Jiang[1, 2,†]

1. Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 210000, China.
2. Nanjing Vocational University of Industry Technology, Nanjing, Jiangsu, 210000, China.

Submission Info

Communicated by Z. Sabir
Received June 2, 2024
Accepted September 1, 2024
Available online October 4, 2024

Abstract

The personalized recommendation method is an efficient and effective solution to optimize educational resources, which can effectively extract learners' preference characteristics and select suitable high-quality educational resources for them to improve teaching efficiency. In this paper, we construct a confidence-aware recommendation model for educational resources. In order to reduce the sparsity problem of rating data, a constraint method based on the implicit factor of user-object interaction is designed to construct the implicit factor of item-user interaction by using the information of a single review. Meanwhile, a confidence matrix is introduced to improve the accuracy of recommendations, and a loss function is constructed using the maximum a posteriori estimation theory. The loss function is optimized by introducing a small batch gradient descent algorithm at the end. Two datasets are used to verify the model's performance, and the impact on students' innovation ability is assessed in teaching practice. It is found that the model in this paper outperforms the other three classification models, and the test results on the two datasets of online_shopping_10_cats and Course Classify achieve an accuracy of more than 0.95. The experimental class's innovation ability increased by 37.6% from 46.25 to 63.63, while the control class's pre- and post-scores of 46.36 and 47.88 were not significantly different. It can be seen that the experimental class has progressed differently in the four dimensions after applying the model of this paper. This study proposes new ideas and feasible paths for combining the optimal allocation of educational resources with cutting-edge information technology, empowering students' innovation ability through big data technology.

Keywords: Confidence perception; Interaction implied factor; Confidence matrix; Loss function.
AMS 2010 codes: 00A35

†Corresponding author.
Email address: jianglijie1232024@126.com

ISSN 2444-8656
https://doi.org/10.2478/amns-2024-2792

sciendo
OPEN ACCESS © 2024 Lijie Jiang, published by Sciendo.

This work is licensed under the Creative Commons Attribution alone 4.0 License.

1 Introduction

Online education in the era of big data relies on Internet technology to fully integrate network services with learning and education, breaking the time and space limitations of traditional school education, utilizing the client port to provide rich learning information, prompting students to actively participate in, timely feedback, and continuous innovation of the education model [1-2]. Online education is able to maximize the absorption and pooling of resources while having the technical skills to provide one-on-one services truly tailored. The development of online education is based on the autonomy of learning and the systematic nature of resources [3]. In the era of big data, educational resources become information flows, and data packets and personality demands become data ports. To adapt to the analysis and sharing of data as a means of speculation and prediction, it is necessary to increase the research and improvement of personality demands and resource optimization [4-5]. Innovative ability refers to the ability of an individual to use all known information (including existing knowledge and experience, etc.) to produce some unique, novel, socially or personally valuable product [6]. In innovation education, it is important to focus on the stimulation of students' learning interest, the cultivation of innovative spirit, the development of potential, the development of personality, and the improvement of practical ability, which is important for improving the comprehensive quality of students and increasing their employment opportunities [7].

On the one hand, data resources need to be updated in real-time. Online education has an immediate nature and can make adjustments at any time with the constant changes in students' demands. This requires that the resources provided are dynamic and updated and can be adjusted in real-time [8]. On the other hand, data resources need to be managed carefully. Online education resources belong to the full sample collection. In this way, the data targeting requirements are very high, which is directly related to the targeted education of the resources, and it is necessary to optimize the volume of resource data in order to enhance the innovative ability of students [9].

Literature [10] determines the teaching status quo of culture and art disciplines, analyzes the innovative ways to optimize the teaching and development of modern cultural education, and the results show that more attention needs to be paid to distance teaching methods and the rational use of online educational resources to carry out innovation. Literature [11] uses a multi-objective ant colony improvement algorithm to optimize the cloud computing platform and optimizes the training system of innovation education in colleges and universities through OpenStack cloud computing, which verifies the effectiveness of the method and promotes the development of innovation education. Literature [12] developed a blended learning design to optimize the teaching curriculum for educational technology majors, and the validation showed that an efficient learning process was supported through blended learning, creating a more innovative and interactive learning process. Literature [13] aims to describe an optimizable design for innovative learning in the era of the Industrial Revolution 4.0, considering active learning models that can foster creativity in students, packaging the learning design into interactive media that can be used to design more interesting and interactive learning that provokes students' interest in learning. Literature [14] integrates the cultivation of college students' comprehensive ability of innovation and entrepreneurship with the "Internet +" technology in depth, continuously improves the quality of college students' innovation and entrepreneurship teaching, and cultivates more and more excellent innovative talents. Literature [15] optimizes the allocation of resources for innovation education under the environment of big data by combining the new situation in depth, and the multimodal teaching based on computer vision creates a new pattern of education, fully mobilizes students' enthusiasm for learning, and meets the needs of college students for innovation education.

In this paper, we construct a confidence-aware recommendation model for optimizing educational resources based on confidence perception to address the problems in general data recommendation

algorithms while taking into account the review information as an interactive relationship and the damage of misleading reviews to recommendation performance. Firstly, the interaction implied factors of users and objects are constructed, and then the relationship between rating outliers and misleading comments is measured by the confidence matrix to further improve the performance of the model so as to construct more accurate implied factors of users and objects and provide more accurate educational resources recommendation services for the educated. The model is tested, and compared with baseline models such as GRU on two datasets, and finally, the model is applied in teaching practice and analyzed for its effect on improving students' innovation ability.

2 Method

2.1 The significance of optimizing the allocation of innovative educational resources

2.1.1 Imbalanced allocation of educational resources

The equalization of basic education resource services means breaking through the existing configuration bottlenecks, giving full play to the characteristics of educational resource sharing, maximizing the effectiveness of quality educational resources, and enabling more people to enjoy better quality educational resources at a lower cost. The equalization of basic education resource services can be measured in terms of the accessibility of education resources and the level of application: accessibility reflects whether users in the jurisdiction can have equal access to education resources, focusing on access to quality education resources. The application level is a reflection of how users apply the acquired educational resources to the teaching process, which is directly related to their performance in using the resources. The equalization of basic education resource services requires that users in the jurisdiction should enjoy quality education resources in roughly equal quantity and quality, which contributes to the fair distribution of quality education resources. It can be seen that the means of solving the problem of non-equalization of traditional teaching resources cannot be applied to the field of educational resource allocation, and the mode of "building an educational resource bank in schools" will inevitably lead to inefficiency in resource allocation and loss of utility, triggering more social costs. In order to truly realize the equalization of basic education resource services, we cannot follow the traditional teaching resource allocation model but must build a new model for equalizing basic education resource services based on the characteristics of sharing, openness, and interaction of educational resources.

2.1.2 Rationale for sharing and interchange of educational resources

Educational resources are characterized by the richness of their knowledge content, the simplicity of their use, the interactivity of the delivery process, the value address of their dissemination and use, and, in particular, the shared nature of the right to use the resources, which can be shared with others, which determines that educational resources can be built and shared regionally in the network environment. Differences in regional economic development can lead to gaps in the level of regional education development. The gap can be increased or expanded by relying on the development of the region itself to solve the regional education imbalance. The exchange and sharing of educational resources can achieve complementary advantages and a dynamic balance of regional educational resources. Regional educational resource sharing and interchange can expand the beneficiary area of high-quality resources, which is conducive to improving economic efficiency, thus realizing the equalization of basic education resource services and helping to promote the fairness of basic education.

1) The characteristics of educational resources determine

The biggest characteristic of educational resources lies in sharing. That is, with the support of information technology, the same educational resources can be possessed and used by different people at the same time or at different times without changing their value and nature, but instead can realize value-added in the process of dissemination and use. The goal of the optimal allocation of educational resources is to realize the great sharing of high-quality educational resources, stemming from the non-depletion nature of the consumption of educational resources: educational resources in the use and consumption of the information content is manifested in the transfer of the information content from one kind of material carriers to another, from a certain user to another user, for any aspects of teaching and learning, the value of the use of educational resources and utility will not disappear.

2) The need for innovative learning

Education is no longer limited to a certain stage in a person's life but throughout their or her life. Innovative learning is characterized by the breadth, autonomy, and creativity of learning: the breadth of learning is reflected in the fact that the school is no longer the only place for learning but that learning takes place in a variety of environments and that various forms of learning form a good connection with school education. Learning autonomy is reflected in the fact that anyone can choose to learn whatever they are interested in at any time they need to. The creativity of learning requires learners to summarize their experiences during the learning process and create new knowledge using different ways of thinking.

Provide users with personalized resource services, realize the reasonable transformation, integration, and consolidation of high-quality educational resources, improve the efficiency of users in acquiring knowledge, and maximize the value and utility of high-quality resources. Collaborative development and application of educational resources is the development trend of resource construction: it is necessary to actively mobilize the power of the government, schools, research institutions, enterprises, and other aspects of society and to adopt a variety of construction methods, such as self-construction, joint construction, introduction and self-build, etc., so as to meet people's learning and life needs, and to build content-rich resources suitable for operation by a variety of mobile devices and capable of meeting people's practical needs. 2) To build lifelong learning resources that are rich in content, suitable for the operation of a variety of mobile devices, and able to meet people's actual needs.

3) Complementary advantages of regional resources

Differences in regional education development determine the certainty of the exchange and sharing of educational resources. The development of regional education is always closely related to regional social development, especially economic development.

The sharing of educational resources can achieve complementary advantages and a dynamic balance of regional educational resources. Basic education has the attribute of being a public product, which is a foundation project organized by the state aimed at improving national quality. The optimal allocation of basic education resources should follow the principle of fairness, strengthen the government's guidance, and guarantee equal opportunities, processes, and results in obtaining quality education resources so as to achieve a relatively balanced state of public education resources allocation. Quality educational resources as public goods have positive externalities, i.e., resource sharing and exchange behaviors of resource providers or

resource users will have a positive impact on other resource providers or resource users, and the beneficiaries do not pay the full price for the behaviors.

4) Helping to promote equity in education

Educational resource sharing is first and foremost initiated by the government and given sustained funding. The construction of educational resources necessitates the full cooperation of many parties, and the public should enjoy the results of its construction to the fullest extent possible. The sharing of educational resources plays an important role in the development of education for all: the purpose of resource sharing is to enable citizens from all walks of life and in all regions to enjoy the same high-quality educational resources, thereby promoting educational equity.

2.2 Personalized Recommendation Model for Educational Resources

In this paper, we use deep learning technology to model learners' feedback information and construct efficient educational resource recommendation algorithms so that learners can have a more personalized learning experience during the learning process.

2.2.1 A Framework for confidence-aware recommendation modeling

The confidence-aware recommendation model can be divided into three main modules: confidence matrix, review implied factor representation learning, and confidence-aware recommendation model.

In the confidence matrix module, a confidence matrix is designed to measure the relationship between rating outliers and misleading reviews, which improves the performance and robustness of the recommendation system. In the review implied factor representation learning module, the review text is input into the convolutional divine meridian, through which the convolutional neural network can capture the contextual information, learn the weight value of each word's influence on the final result, and then generate the implied inner sub-feature vectors of the review text [16]. In the confidence-aware model, a zero-mean Gaussian prior is applied to the implied factor vectors of the user and the object, respectively.

2.2.2 Comments on Implicit Factor Representation Learning

The goal of the Comment Implicit Factor Representation learning module is to go through the convolutional neural network to extract the feature vector representation of user comments. The whole module is divided into four layers: weighted word vector representation layer, convolutional layer, pooling layer, and output layer.

The first layer is the weighted word vector representation layer. First, the embedding matrix converts the input text into a dense real-valued matrix, where each column vector of the matrix is a representation of the corresponding word vector in the text, and this matrix is used as the input to the subsequent convolution layer. The comment matrix $D \in \square^{1<p}$ is represented as follows:

$$D = \begin{bmatrix} & | & | & | & \\ \cdots & w_{c,1} & w_c & w_{,-1} & \cdots \\ & | & | & | & \end{bmatrix} \tag{1}$$

where l is the length of the input comment text and p is the size of the embedding vector dimension of the word. Each word vector w_i is assigned a weight of value ϕ_i, so the output of this layer is as follows:

$$H = \begin{bmatrix} & | & | & | & \\ \cdots & \phi_{i-1} w_{i-1} & \phi_i w_i & \phi_{i+1} w_{i+1} & \cdots \\ & | & | & | & \end{bmatrix} \tag{2}$$

For the convolutional layer, the main objective is to extract contextual features from the comments. The contextual feature vector $c_j^i \in \square$ is extracted through the j nd convolutional kernel $W^j \in \square^{t \times p}$ with shared weights. The window size t defines the number of surrounding words:

$$c_i^j = \mathrm{Re}\,lu\left(W^j * H_{(:,i:(i+t-1))} + b^j\right) \tag{3}$$

The feature representation $c^j \in \square^{l-t+1}$ of the review text extracted by the j st convolutional kernel is shown in equation (4):

$$c^j = \begin{bmatrix} c_1^j & c_2^j & \cdots & c_i^j & \cdots & c_{l-t+1}^j \end{bmatrix} \tag{4}$$

The third layer is the pooling layer, also known as the downsampling layer. The output of the final pooling layer is shown in equation (5):

$$d = \left[\max\left(c^1\right), \max\left(c^2\right), \ldots, \max\left(c^j\right), \ldots\right] \tag{5}$$

The fourth layer is the output layer. The output of the pooling layer d is mapped into a k -dimensional space by a nonlinear space to obtain the final implicit factor feature representation of the comments, and the final result is shown as follows.

$$s = \tanh\left(W_2\left\{\tanh\left(W_1 d + b_1\right)\right\} + b_2\right) \tag{6}$$

where $W_1 \in \square^{h \times m}$, $W_2 \in \square^{k \times m}$ are the weight matrices, $b_1 \in \square^{h}$, $b_2 \in \square^{k}$ are the bias vectors corresponding to the W_1 and W_2 matrices, and $s \in \square^{k}$ is the output. The vector of implied factors of comments as a function of the output is shown in equation (7):

$$s_{ij} = CNN_{\Theta}\left(X_{ij}\right) \tag{7}$$

Where Θ represents the set of all parameters in the convolutional neural network, X_{ij} denotes the text of the original review of item j by user i, and s_{ij} denotes the final feature representation of a single review, i.e., the implicit factor of the review.

2.2.3 Confidence-aware recommendation models

Assuming that there are m objects and n users, the rating prediction is therefore shown in Equation (8):

$$R_v = u_i v_j^T + \gamma_i + \psi_j + \tau \tag{8}$$

To get the representation of u_i and v_j implicit factors will be able to predict the scores of unknown objects and recommend the objects with ungoverned scores to the user, completing the whole derivation process. Next, the target model CARM will be derived from the perspective of probability.

1) Scoring modeling

According to CARM's score top recommendation model (8), the error between the real score and the predicted score is modeled as Gaussian noise: then, the probability function for the observed score in the training data can be written as equation (9):

$$p\left(R|U,V,\gamma,\psi,\tau\right) = \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathrm{N}\left(r_{,} \mid \dot{r}_{v,}, \sigma^2\right)\right]^1 \tag{9}$$

Where $\mathrm{N}\left(x|\mu,\sigma^2\right)$ denotes a Gaussian probability density function with mean variance μ and σ^2. Symbol 1_{ij} is the indicator function, which is 1 if user i has rating data for object j and 0 otherwise.

2) User and object modeling

The probability distribution of the matrix V consisting of the implied factors of all objects is shown in equation (10):

$$p\left(V|\sigma_v\right) = \prod_{j=1}^{n}\mathrm{N}\left(v_j|0,\sigma_v^2 I\right) \tag{10}$$

The same assumption can be made for the implicit factors of the users so that the probability distribution of the matrix U consisting of the implicit factors of all users is as in equation (11):

$$p\left(U|\sigma_U\right) = \prod_{i=1}^{m}\mathrm{N}\left(u_i|0,\sigma_U^2 I\right) \tag{11}$$

Where $\sigma_V^2 I$ and $\sigma_U^2 I$ are the covariance matrices of the object-implicit factor matrix V and the user-implicit factor matrix U, respectively.

3) Modeling of a single comment

The k-dimensional vector $u_i \circ v_j$ represents the Hadamard product of user-implicit factors u_i and object-implicit factors v_j, and can be regarded as a representation of the information about the interaction between the user i and the object j [17]. The difference between $u_i \circ v_j$ and the implicit factor representation of user i's comments on object j is small. It is assumed that the difference between them obeys a Gaussian distribution. Therefore, the Hadamard product of u_i and v_j satisfies equation (12):

$$u_i \circ v_j = CNN_\Theta\left(X_{ij}\right) + \in_R, \in_R \sim \mathrm{N}\left(0, \sigma_R^2 I\right) \tag{12}$$

Then, the conditional probability distribution of $u_i \circ v_j$ is as follows:

$$p\left(u_i \circ v_j \middle| X_{ij}, \Theta, \sigma_R^2\right) = \mathrm{N}\left(u_i \circ v_j \middle| CNN_\Theta\left(X_{ij}\right), \sigma_R^2 I\right) \tag{13}$$

Thus, for all observable comment text data in the training set, the conditional probability is as in equation (14):

$$p\left(U, V \middle| X, \Theta, \sigma_R^2\right) = \prod_i^m \prod_j^n \mathrm{N}\left(u_i \circ v_j \middle| CNN_\Theta\left(X_i\right), \sigma_R^2 I\right) \tag{14}$$

Integrating the above equations, the maximum a posteriori probability of the model is as in equation (15):

$$\begin{aligned} & p\left(U, V \middle| R, X, \gamma, \psi, \Theta, \sigma, \sigma_U, \sigma_V, \sigma_R\right) \\ & = \frac{p\left(R \middle| U, V, \gamma, \psi, \sigma\right) p\left(U \middle| \sigma_U\right) p\left(V \middle| \sigma_V\right) p\left(U, V \middle| X, \Theta, \sigma_R\right)}{p\left(R, X, \gamma, \psi, \Theta, \sigma, \sigma_U, \sigma_V, \sigma_R\right)} \end{aligned} \tag{15}$$

Equation (16) is obtained by introducing the Maximum A Posteriori Estimate (MAP) into equation (15):

$$\begin{aligned} & \max_{U,V} p\left(U, V \middle| R, X, \gamma, \psi, \Theta, \sigma, \sigma_U, \sigma_V, \sigma_R\right) \\ & = \max_{U,V} \underbrace{p\left(R \middle| U, V, \gamma, \psi, \sigma\right)}_{Likelihood\ probability} \cdot \underbrace{p\left(U \middle| \sigma_U\right) p\left(V \middle| \sigma_V\right)}_{User\ and\ object\ prior\ probabilities} \\ & \cdot \underbrace{p\left(U, V \middle| X, \Theta, \sigma_R\right)}_{A\ priori\ probability\ of\ comments} \end{aligned} \tag{16}$$

To facilitate subsequent calculations to prevent numerical overflow, a negative logarithm is introduced to Equation (16) to obtain Equation (17):

$$\begin{aligned} L = & \log p\left(R \middle| U, V, \gamma, \psi, \sigma\right) + \log p\left(U, V \middle| X, \Theta, \sigma_R\right) \\ & + \log p\left(U \middle| \sigma_U\right) + \log p\left(V \middle| \sigma_V\right) \\ = & -\frac{1}{2\sigma^2} \sum_i^m \sum_j^n 1_{ij} \left(R_{ij} - u_i v_j^T - \gamma_i - \psi_j - \tau\right)^2 \\ & - \frac{1}{2\sigma_R^2} \sum_i^m \sum_j^n \left(u_i \circ v_j - CNN_\Theta\left(X_{ij}\right)\right)^T I^{-1} \left(u_i \circ v_j - CNN_\Theta\left(X_{ij}\right)\right) \\ & + \frac{1}{2\sigma_U^2} \sum_i^m \left(u_i\right)^T I^{-1} u_i - \frac{1}{2\sigma_V^2} \sum_j^n \left(v_j\right)^T I^{-1} v_j - C \end{aligned} \tag{17}$$

After some simplification operations, the constant term C can be omitted, so the final objective function of Equation (17) after removing irrelevant variables is Equation (18):

$$
\begin{aligned}
E_{CARM-c} &= \sum_{i}^{m}\sum_{j}^{n}\frac{1_{ij}}{2}\left(R_{ij}-R_{ij}\right)^2 \\
&+\frac{\lambda_R}{2}\sum_{i}^{m}\sum_{j}^{n}\left\|\left(u_i\circ v_j - CNN_{\Theta}\left(X_{ij}\right)\right)\right\|_F^2 \\
&+\frac{\lambda_V}{2}\sum_{i}^{m}\left\|u_i\right\|_F^2+\frac{\lambda_V}{2}\sum_{j}^{n}\left\|v_j\right\|_F^2
\end{aligned}
\tag{18}
$$

where λ_U, λ_V, and λ_R are set to σ^2/σ_U^2, σ^2/σ_V^2, σ^2/σ_R^2, and $\|\square\|_F^2$ representing the Frobenius norm of the vector, respectively.

A confidence matrix is introduced to measure how much each comment contributes to the model. The values of the elements in the confidence matrix represent the level of confidence in each comment, and comments corresponding to outliers in the scoring matrix are given a low level of confidence and vice versa. Therefore, the corresponding energy function is rewritten as Equation (19):

$$
\begin{aligned}
E_{CARM} &= \sum_{i}^{m}\sum_{j}^{n}\frac{1_{ij}}{2}\left(R_{ij}-R_{ij}\right)^2 \\
&+\sum_{i}^{m}\sum_{j}^{n}Q_{ij}\frac{\lambda_R}{2}\left\|\left(u_i\circ v_j - CNN_{\Theta}\left(X_{ij}\right)\right)\right\|_F^2 \\
&+\frac{\lambda_V}{2}\sum_{j}^{m}\left\|u_i\right\|_F^2+\frac{\lambda_V}{2}\sum_{j}^{n}\left\|v_j\right\|_F^2
\end{aligned}
\tag{19}
$$

2.2.4 Confidence matrix

The values of the elements in the confidence matrix are computed from the confidence function $\mathrm{F}\left(\cdot\right)$, the value of $\mathrm{F}\left(\cdot\right)$ represents the level of contribution of the different reviews to the model, and each element value in the matrix represents the derivative of the outlier rating value [18]. Matrix Q can be seen as a regularization adjustment for each review text. Where the element values in the confidence matrix Q are scaled to a range of $[0,1]$.

The confidence function $\mathrm{F}\left(\cdot\right)$ of Equation (20) determines the element values in matrix Q:

$$
\begin{aligned}
Q_{ij} &= \mathrm{F}\left(R_{ij},\beta\right) \\
&= \begin{cases}
e^{-\mathrm{Re}lu\left(\sqrt{\sum_{N=\square}^{\infty}R_{\infty}}-R_{\theta}|-\beta\right)-\mathrm{Re}lu\left(\sqrt{\sum_{\theta-i.\square}^{\theta\to\infty}R_v}-R_i|-\beta\right)}, R_{ij}\le 3 \\
e^{-\mathrm{Re}lu\left(\sqrt{\sum_{N=\square}^{\infty}R_{\infty}}-R_{\theta}|-\beta\right)-\mathrm{Re}lu\left(\sqrt{\sum_{\theta-i.\square}^{\theta\to\infty}R_v}-R_i|-\beta\right)}, R_{ij}>3
\end{cases}
\end{aligned}
\tag{20}
$$

2.3 Model Optimization Algorithm

In order to de-solve the problem of feature representation learning in recommendation algorithms, a small batch more stochastic gradient algorithm is used to optimally solve Eq. (19), and the update rules for all variables in CARM are shown below:

Update of U : Estimation The subproblem of U correspondingly minimizes Eq:

$$\begin{aligned} E_{CARN} &= \sum_{i}^{m}\sum_{j}^{n}\frac{1_{ij}}{2}\left(R_{ij}-R_{v}\right)^{2}+\frac{\lambda_{\omega}}{2}\sum_{j}^{m}\left\|u_{i}\right\|_{F}^{2} \\ &+\frac{\lambda_{n}}{2}\sum_{j}^{m}\sum_{j}^{n}Q_{ij}\left\|\left(u_{i}\circ v_{j}-CNN_{\Theta}\left(X_{ij}\right)\right)\right\|_{F}^{2} \end{aligned} \tag{21}$$

The update rule for parameter u_i is then Eq:

$$\begin{aligned} u_{i} &\leftarrow u_{i}-\alpha\frac{\partial E_{CARy}^{U_{j}}}{\partial u_{i}} \\ &= u_{i}-\alpha\left\{\sum_{j}^{n}1_{ij}\left(R_{ij}-R_{ij}\right)\cdot v_{i}+\lambda_{U}u\right. \\ &\left.+\sum_{j}^{n}1_{ij}Q_{j}\lambda_{R}\left(u_{i}\circ v_{j}-CNN_{\Theta}\left(X_{ij}\right)\circ v_{j}\right)\right\} \end{aligned} \tag{22}$$

Update of V : Since U and V are equivalent, the objective function with respect to V corresponds to the following equation:

$$\begin{aligned} E_{CARM}^{v} &= \sum_{i}^{m}\sum_{j}^{n}\frac{1_{ij}}{2}\left(R_{ij}-R_{ij}\right)^{2}+\frac{\lambda_{V}}{2}\sum_{j}^{n}\left\|v_{i}\right\|_{F}^{2} \\ &+\frac{\lambda_{R}}{2}\sum_{i}^{m}\sum_{j}^{n}Q_{ij}\left\|\left(u_{i}\circ v_{j}-CNN_{\Theta}\left(X_{V}\right)\right)\right\|_{F}^{2} \end{aligned} \tag{23}$$

Thus, the update rule for parameter v_j is:

$$\begin{aligned} v_{j} &\leftarrow v_{j}-\alpha\frac{\partial E_{CARM}^{v}}{\partial v_{j}} \\ &= v_{j}-\alpha\left\{\sum_{i}^{m}1_{ij}\left(R_{ij}-R_{ij}\right)\cdot u_{i}+\lambda_{V}v_{j}\right. \\ &\left.+\sum_{i}^{m}1_{ij}Q_{ij}\lambda_{R}\left(u_{i}\circ v_{j}-CNN_{\Theta}\left(X_{ij}\right)\circ u_{i}\right)\right\} \end{aligned} \tag{24}$$

Updates for γ and ψ : Similarly, for γ the update rule is:

$$\gamma_{i}\leftarrow\gamma_{i}-\alpha\frac{\partial E\gamma_{ARM}}{\partial\gamma_{i}}=\gamma_{i}-\alpha\sum_{i}^{m}1_{iv}\left(R_{i}-R_{i}\right) \tag{25}$$

The update rule for pair ψ is:

$$\psi_j \leftarrow \psi_i - \alpha \frac{\partial EE_{iRM}}{\partial \psi_j} = \psi_j - \alpha \sum_j^n 1_{ij} \left(R_{ij} - R_v \right) \tag{26}$$

Update for Θ: Because the internal parameters of the convolutional neural network are too complex, the gradient for Θ is replaced by the partial derivative of $CNN_\Theta \left(X_{ij} \right)$ and is not solved further:

$$E_{CARM}^{\Theta} = \frac{\lambda_R}{2} \sum_i^m \sum_j^n Q_{ij} \left\| \left(u_i^{\circ} v_j - CNN_\Theta \left(X_{ij} \right) \right) \right\|_F^2 \tag{27}$$

Thus, the Θ-update rule is written as:

$$\begin{aligned} \Theta &\leftarrow \Theta - \alpha \frac{\partial E_{CARM}^{\Theta}}{\partial \Theta} \\ &= \Theta - \alpha \left\{ \lambda_R \sum_i^m \sum_j^n Q_{ij} (u_i \circ v_j - CNN_\Theta \left(X_{ij} \right) \cdot \frac{\partial CNN_\Theta \left(X_{ij} \right)}{\partial \Theta} \right\} \end{aligned} \tag{28}$$

3 Results and Discussion

3.1 Model Performance Analysis

Classification recommendation experiments are conducted on educational resource datasets using the model. In this paper, we evaluate the experimental results using accuracy rate and recall (Recall). The recall rate is the ratio of samples that are judged as positive examples among all samples to all positive examples samples. Two datasets, the online_shopping_10_cats dataset and the Course Classify dataset, are chosen for the experiment.

The online_shopping_10_cats dataset is a public text categorization dataset, and the dataset has 10 categories and more than 60,000 text data. The experiments in this chapter divide the online_shopping_10_cats dataset into a training set, a validation set, and a test set, with a division ratio of 8:1:1.

The CourseClassify dataset is made up of 680-course information from the MOOCCube dataset, which is then enhanced with 14,103 items using crawling techniques. Due to the uneven distribution of data, data enhancement is performed on the dataset using data enhancement methods such as back-translation and synonym replacement in order to solve the distribution problem. The total number of data sets after enhancement operation and data preprocessing is 31852 articles. The data is divided into training sets, validation sets, and test sets with a division ratio of 8:1:1.

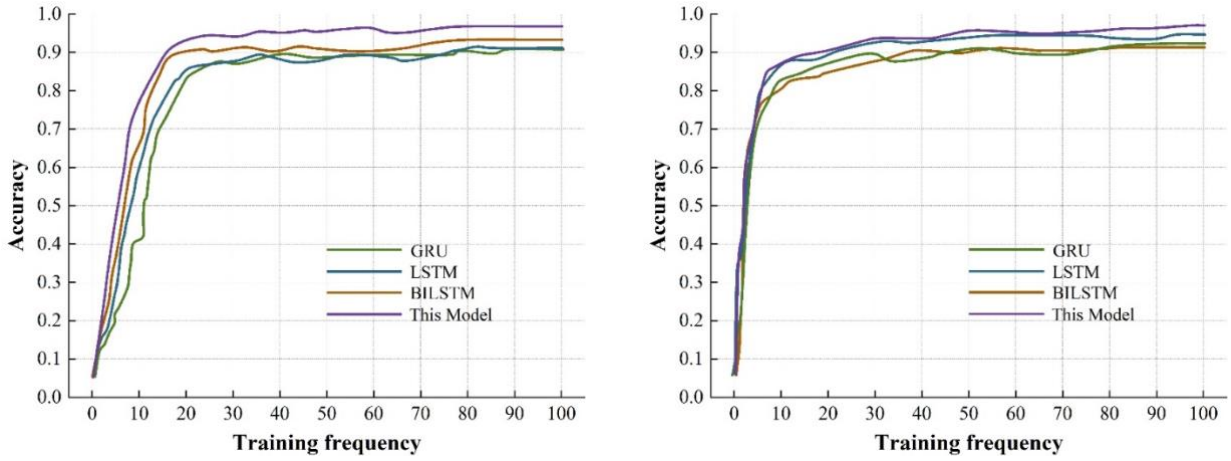
In order to evaluate the performance of the BILSTM-ATTENTION proposed in this paper, comparative experiments are conducted, and the following comparative models are used:

GRU: This model is a specific form of RNN, based on the LSTM model, changing the three gates in the LSTM model to two gates (update gate, reset gate), which has a relative improvement in computational efficiency compared to the LSTM model due to its simple structure and fewer parameters. However, it also loses some useful information in some cases.

LSTM: This model is a specific form of RNN. Unlike the traditional RNN, LSTM has three gates - the forgetting gate, the updating gate, and the output gate to control the cell state.

BILSTM: This model is a combination of a forward LSTM and a backward LSTM. Compared to the unidirectional LSTM model, the BILSTM model is able to capture semantic dependencies in both forward and backward directions more accurately.

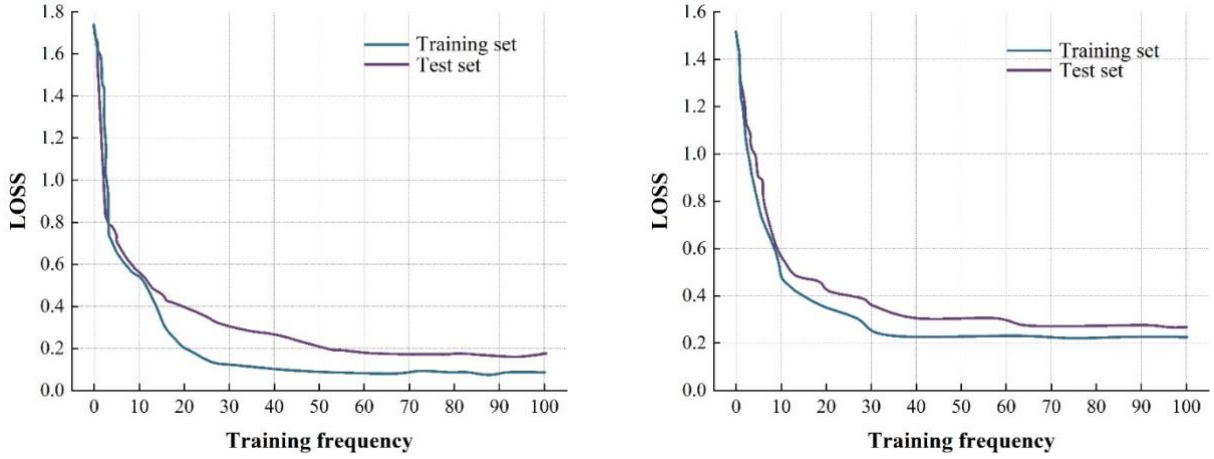
Figure 1 shows the accuracy graphs of each model after 100 rounds of training on two datasets, where (a) and (b) are the test results of two datasets, the online_shopping_10_cats dataset and the Course Classify dataset, respectively. As can be seen from the graphs, the model in this paper outperforms the other three classification models, achieving an accuracy of more than 0.95 on both datasets.



(a) Test results on online_shopping_10_cats (b) Test results on Course Classify

Figure 1. Test accuracy of two data sets

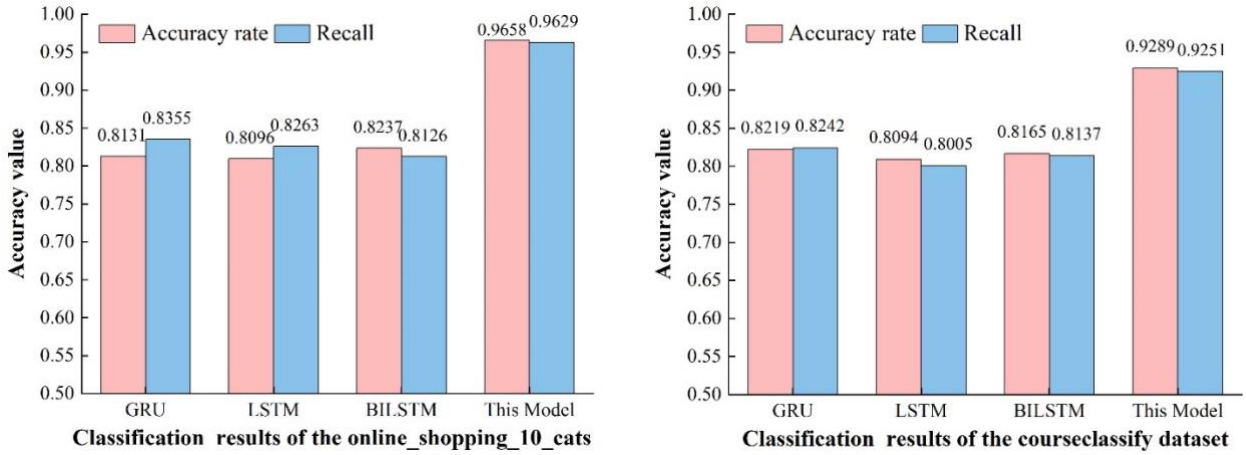
After 100 rounds of training on both datasets, the accuracy and loss of the model in this paper change, as shown in Fig. 2. The horizontal coordinate in the figure is the number of training rounds. As the number of rounds increases, the accuracy of the training set and validation set increases, the loss decreases, and the change decreases after about 20 rounds, and the overall convergence phenomenon is shown.



(a) Test Loss on online_shopping_10_cats (b) Test Loss on Course Classify

Figure 2. Test Loss of two data sets

Figure 3 shows the results of the experiments of each model using the CourseClassify dataset with the online_shopping_10_cats dataset. The models in this paper generally outperform the LSTM model and the GRU model, indicating that the design of the models in this paper is able to learn the text content better and improve the model performance. Among them, the accuracy on online_shopping_10_cats is higher than that of GRU, LSTM, and BILSTM models by 0.1527, 0.1562, and 0.1421, respectively.



(a) Results on online_shopping_10_cats (b) Results on Course Classify

Figure 3. Multiple models test the accuracy of two data sets

3.2 Effectiveness of Students' Creative Ability Enhancement

The model was applied to a school in Guangzhou for a semester-long teaching experiment to analyze the effect of the model on the enhancement of students' innovation ability. The experimental class uses this paper's model to recommend personalized educational resources to students and complete the teaching tasks. The control class was taught in accordance with the original plan. Except for this, the two classes had the same teaching conditions. The innovation ability test was divided into four modules: critical thinking, observation, imagination, and creativity, with 20 points for each module and a total score of 100 points. The Creative Competence Test was administered before the beginning and after the end of the semester, respectively.

Critical thinking requires students to observe problems from multiple perspectives and then improve their ability to recognize and detect problems and fallacies and dare to put forward subversive viewpoints on old traditions, methods, and thinking. Observativeness requires students to be keen enough to spot problems in detail, thus stimulating enthusiasm for innovation. Imagination is the depiction of future unknown problems, which is vital to students' innovative ability. Creativity is the core of innovative ability, which requires students to have a strong pioneering spirit and innovative courage.

The critical thinking module's results are compared in Table 1. There is no significant difference between the two classes' innovative ability in the pre-semester test, and after one semester of experiment, the experimental class is 3.59 points higher than the control class with 17.21 points, and the experimental class has obvious progress in critical thinking.

Table 1. Contrast with critical thinking dimension achievement

	Class	N	Average score	σ	Z	P
Pre-application	Laboratory class	49	13.25	6.48	-0.082	0.479
	Control class	50	13.43	5.49		
Application	Laboratory class	49	17.21	3.35	0.056	0.008***
	Control class	50	13.62	4.49		

The observation module scores are compared in Table 2. The experimental class improved from 9.76 points on the pre-test to 15.83 points on the post-test, while the control class scored 9.54 and 10.24 points on the pre and post-scores, which was not a significant change. The model's use contributed significantly to the improvement of students' observational skills.

Table 2. Observation strength dimension achievement

	Class	N	Average score	σ	Z	P
Pre-application	Laboratory class	49	9.76	4.16	0.108	0.215
	Control class	50	9.54	4.18		
Application	Laboratory class	49	15.83	5.41	0.044	0.017**
	Control class	50	10.24	3.25		

A comparison of scores on the imagination module is shown in Table 3. Before the semester started, the experimental and control classes had scores of 8.92 and 8.62, with no significant difference between them. In the post-test, the experimental class improved to 13.56 points in the imagination module, while the control class improved to 9.07 points. The experimental class made a significant improvement over the control class.

Table 3. Imagination strength dimension achievement

	Class	N	Average score	σ	Z	P
Pre-application	Laboratory class	49	8.92	6.08	0.075	0.374
	Control class	50	8.62	3.81		
Application	Laboratory class	49	13.56	4.25	0.039	0.014**
	Control class	50	9.07	6.2		

A comparison of the creativity module scores is shown in Table 4. The creativity score of the experimental class increased from 14.32 to 17.03, an improvement of 2.71 points. The control class's pre-and post-test scores were 14.77 and 14.95, with no significant differences.

Table 4. Creativity strength dimension achievement

	Class	N	Average score	σ	Z	P
Pre-application	Laboratory class	49	14.32	3.28	0.062	0.298
	Control class	50	14.77	3.02		
Application	Laboratory class	49	17.03	5.5	0.054	0.023**
	Control class	50	14.95	3.76		

Table 5 demonstrates the comparison of the total scores for the innovative ability test. The innovative ability of the experimental class increased by 37.6% from 46.25 to 63.63, and there is no significant difference between the before and after scores of 46.36 and 47.88 in the control class. It can be seen

that after applying the model of this paper, the experimental class has different degrees of progress in the four dimensions of critical thinking, observation, imagination and creativity, and the model has an excellent effect on enhancing students' innovative ability.

Table 5. Comparison of test results of innovative ability

	Class	N	Average score	σ	Z	P
Pre-application	Laboratory class	49	46.25	15.53	0.094	0.206
	Control class	50	46.36	15.22		
Application	Laboratory class	49	63.63	16.5	0.088	0.011**
	Control class	50	47.88	15.05		

4 Conclusion

In this paper, we use deep learning algorithms to construct an efficient, personalized recommendation model for educational resources based on feedback information algorithms and validate the model's performance on the dataset. The model is used in teaching practice to evaluate the impact on the improvement of students' innovation capability. It is found that the model in this paper outperforms the other three classification models, and the test results on the online_shopping_10_cats and Course Classify datasets achieve an accuracy rate of more than 0.95. After 100 rounds of training, the accuracies of the training and validation sets increase, and the losses decrease. However, the change decreases after about 20 rounds, and the overall convergence phenomenon is observed. The innovation ability of the experimental class increased 37.6% from 46.25 to 63.63, while there was no significant difference between the before and after scores of 46.36 and 47.88 in the control class. It can be seen that after applying the model of this paper, the experimental class has different degrees of progress in the four dimensions of critical thinking, observation, imagination, and creativity, and the model has an excellent effect on improving students' innovation ability, which is in line with the design expectation.

References

[1] Indrawati, C. D. S. (2021). The Effectiveness of Archiving Videos and Online Learning on Student's Learning and Innovation Skills. International Journal of Instruction, 14(4), 135-154.

[2] Ma, Z., Guan, J., & Li, R. (2021). Research on innovative teaching mode of art education in the age of convergence of media. International Journal of Emerging Technologies in Learning (iJET), 16(2), 272-284.

[3] Wang, L., Zhou, J., & Li, X. (2024). Influence of College Teacher's Instructional Design on the Development of College Students' Thinking of Innovation Multi-Algorithms Perspective Analysis. IEEE Access.

[4] Rii, K. B., Edastama, P., & Nabilah, N. F. (2022). Study on Innovation Capability of College Students Based on Extenics and Theory of Creativity. Startupreneur Business Digital (SABDA Journal), 1(2), 134-142.

[5] Xiong, W., & Zhang, Y. (2024). The path of optimizing and improving the teaching innovation ability of visual communication based on the SWOT model. Applied Mathematics and Nonlinear Sciences.

[6] Han, X. (2024). Integration and Optimization of Teaching Resources for Electronics and Communications Courses in Higher Vocational Education. Transactions on Comparative Education, 6(3), 157-163.

[7] Chootongchai, S., & Songkram, N. (2018). Design and development of SECI and moodle online learning systems to enhance thinking and innovation skills for higher education learners. International Journal of Emerging Technologies in Learning (Online), 13(3), 154.

[8] Huang, M., Yuan, H., & Sun, B. (2022, April). Offline teaching method in blended teaching mode of improving the innovation ability of students in IT majors. In Proceedings of the 7th International Conference on Information and Education Innovations (pp. 30-34).

[9] Zhao, Q., Xiong, C., Liu, K., Zhang, X., & Liu, Z. (2022). Cultivation design of applied undergraduates' engineering innovation ability based on virtualization technology. Wireless Communications and Mobile Computing, 2022(1), 5500021.

[10] Kuzmenko, T., Panarin, O., Parkhomenko, I., Soroka, I., & Salata, H. (2022). Innovation to optimize the teaching and development of modern education in the field of culture and art: vectors of development. Revista Eduweb, 16(2), 107-118.

[11] Xu, C., & Song, C. (2022). Optimization of innovation and entrepreneurship education and training system in colleges and universities based on OpenStack cloud computing. Scientific programming, 2022(1), 2868499.

[12] Pattaufi, P., Aswan, D., & Hakim, A. (2023). The development of teaching material for blended learning: A strategy to improve students' creativity and innovation in the 21st century. Journal of Educational Science and Technology, 9(1), 19-27.

[13] Indriani, F., Hidayah, N., AlQisan, A., & Okfitri, N. N. (2021, September). Innovative learning era industrial revolution 4.0 in optimizing 21st century competence elementary school students. In Proceedings of the Second Asia Pacific International Conference on Industrial Engineering and Operations Management (pp. 3147-3157).

[14] Wang, Y. (2022). Optimizing the Cultivation Path of College Students' Innovation and Entrepreneurship Ability from the Perspective of the Internet. Wireless Communications & Mobile Computing (Online), 2022.

[15] Ren, Z. (2023). Optimization of Innovative Education Resource Allocation in Colleges and Universities Based on Cloud Computing and User Privacy Security. Wireless Personal Communications, 1-15.

[16] Sifang Long,Yanjun Zhang,Shuo Kang,Boliao Li & Jun Wang. (2024). Parameter calibration of the angle of repose of particle materials based on convolutional neural network. Biosystems Engineering152-163.

[17] Hodgson Liam,Li Yue,Iturria Medina Yasser,Stratton Jo Anne,Wolf Guy,Krishnaswamy Smita... & Bzdok Danilo. (2024). Supervised latent factor modeling isolates cell-type-specific transcriptomic modules that underlie Alzheimer's disease progression. Communications Biology(1),591-591.

[18] Chai TongYuen,Goi BokMin & Yap WunShe. (2021). Towards Better Performance for Protected Iris Biometric System with Confidence Matrix. Symmetry(5),910-910.