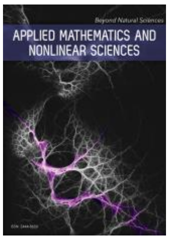
sciendo Applied Mathematics and Nonlinear Sciences, 9(1) (2024) 1-17



Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

Innovation of educational management paths in higher education based on LSTM deep learning model

Qi Zhou[1,†]

1. Science and Technology Department, Yancheng Institute of Technology, Yancheng, Jiangsu, 224000, China.

Submission Info

Communicated by Z. Sabir
Received January 30, 2023
Accepted May 15, 2023
Available online November 4, 2023

Abstract

In this paper, based on the data of the Riiid education platform, the LSTM deep learning model is used to provide accurate prediction and guidance for the education management of colleges and universities. The Gini coefficient is also introduced to simplify the calculation process, focusing on predicting the development of students' careers. To achieve this goal, the online education platform provided a dataset that was carefully pre-processed and cleaned of data, and feature engineering was performed to obtain more informative features. Comparing the AUC value of the offline area of the ROC curve, the AUC value of the LSTM deep learning model can reach 0.758, and the training time of a single model is about 41.8 seconds. Therefore, a deep learning model based on the LSTM algorithm can be used for innovation research.

Keywords: Riiid education platform; LSTM deep learning; University education management; Gini coefficient; Career development.
AMS 2010 codes: 97D60

†Corresponding author.
Email address: hjh13819@163.com

ISSN 2444-8656
https://doi.org/10.2478/amns.2023.2.00972

1 Introduction

Education is related to the country's future development and has long been closely watched by the whole society. In the context of the booming development of the Internet, the amount of data in the field of education has also risen exponentially. The development of artificial intelligence has also changed the field of education from the perspectives of teaching environment, learning mode, education governance and teaching content []-[3]. The current research on the innovation of educational management paths in higher education can be summarized as precision education, an emerging field []-[6]. Due to the development of mobile Internet, the popularity of education has greatly increased, which also leads to an explosive increase in the number of users and the amount of data, and the combination of machine learning and deep learning in the field of education can realize the prediction of students' performance points, the prediction of the accuracy of answering questions, the prediction of students' attendance and so on, which is of great practical significance []-[8].

Predicting students' accuracy in answering questions can effectively track the learning process and adjust the learning plan accordingly. By achieving personalized and customized education, the efficiency of education can be greatly improved, and social resources can be saved. Historical studies have used algorithms such as neural networks, support vector machines, and random forests to achieve academic performance prediction with good results []-[11].

The Riiid dataset selected for this paper was firstly modeled and trained for predictive modeling of students' career development using the LSTM deep learning model to predict the likelihood and trend of students' career direction.

Secondly, the evaluation indexes of university management pathway innovation were investigated to assess the accuracy and performance of the model by comparing the LSTM deep learning model with other machine learning algorithms. Next, the dataset of student career development was preprocessed, including data cleaning and feature engineering, to ensure the completeness and accuracy of the data input into the model and to extract more meaningful features. The exploratory data analysis phase involved detailed data analysis to uncover potential correlations and patterns between the data, guiding subsequent modeling. Then, data cleaning and feature engineering was carried out to deal with the missing values and outliers in the data through data cleaning to ensure the completeness and accuracy of the data. The model's predictive performance can be improved by selecting the most representative and influential features in feature engineering. Finally, the performance of the LSTM deep learning model is compared, and its advantages are analyzed with traditional machine learning algorithms and artificial intelligence to assess the superiority and applicability of the LSTM model in the research of education management path innovation in colleges and universities.

2 Literature review

Noroozi, O et al. propose multidisciplinary innovations and technologies that promote self-directed learning, focusing on technology-enhanced learning environments. Provides learners with rich opportunities to regulate learning processes and activities autonomously. Achieve desired learning outcomes across multiple subject areas, from soft to hard sciences and humanities to natural and social sciences [12]. Denisov, I et al. delve into the impact of educational paradigm shifts and students' emotional states to provide useful references and guidance for policymakers and educators in the education field and promote continuous development and innovation in distance learning [13]. Cavus N et al. identify the factors that led to the total closure of schools during COVID-19, but future research should continue to explore the application of other linear and nonlinear AI techniques to further enhance the understanding and solutions to the problem of educational sustainability. An in-depth exploration of these studies can help to provide developing countries and other affected regions with adaptive and flexible educational models for various challenges that may arise in the future [14]. Suanj, Z et al. found, through descriptive indicators, that there are variations in practices adopted by higher education institutions regarding the effectiveness of academic management and professional development. The need to adapt current human resource management practices in Croatian higher education was examined, emphasizing the need for a performance management system appropriate to the academic environment [15].

Preudhikulpradab S et al. aimed to analyze the impact of online teaching and learning systems on the teacher and student communities. The study addressed faculty perspectives, which included content delivery, technological literacy barriers, and infrastructure issues, as well as student perspectives, which included cost-effectiveness, infrastructure issues, and network connectivity. Educational institutions must be committed to providing quality education that is responsive to market needs to meet their students' employment and development needs [16]. Okada, A et al. analyzed by surveying the attitudes and experiences of 328 students using adaptive trust-based e-learning assessment systems using a mixed-methods approach, and the results showed that the acceptance level of these e-credentialing technologies by distance education students was high. Promoting and improving e-authentication systems must help develop more targeted measures to enhance students' trust and acceptance of the system [17]. Noroozi, O et al. demonstrated the importance of self-regulated learning in modern education and emphasized the positive role of technology in facilitating the learning process. This has a positive impact on improving academic achievement and motivation to learn and promoting innovation in the field of education [18].

3 Deep learning-based prediction model for college education management

3.1 Riiid Education Platform Data

The purpose of the dataset used in this paper is to construct a suitable algorithmic model from an online education platform. The education platform's historical data predicts innovative pathways for education management development in colleges and universities. Riiid is a leading global provider of AI tutor solutions in South Korea, and the data it generates in its operation is a mature and well-established open dataset. Providing innovative pathways to the education market through its high-end AI technology, Riiid Labs is committed to providing AI solutions to the traditional education market. The design idea of its platform data processing is shown in Figure 1, which is very much in line with the needs of deep learning modeling for innovation research on college education management pathways. By fully utilizing the dataset of the Riiid education platform, this study aims to provide feasible and innovative solutions for university education management and promote the development of the education field.

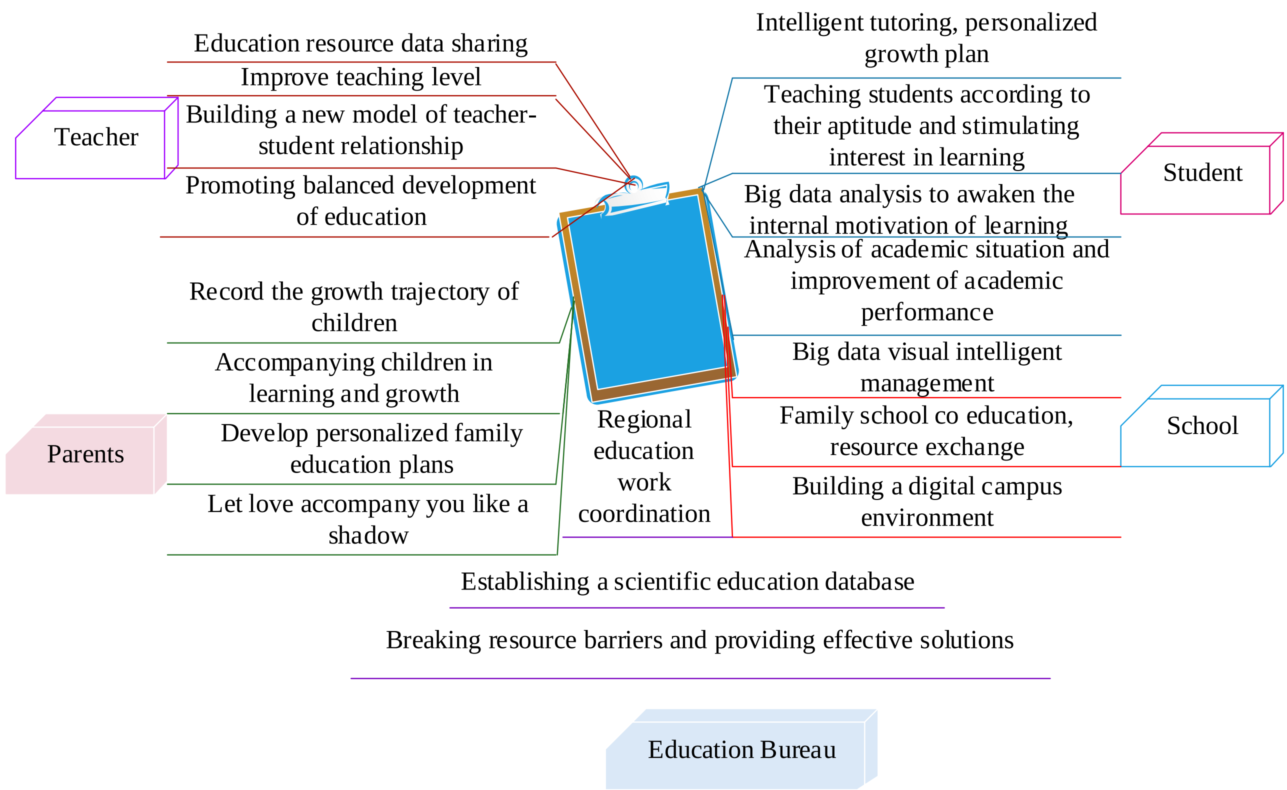


Figure 1. Riiid Design ideas for platform data processing

This paper addresses the challenges of cross-disciplinary machine learning by focusing on regression and classification problems in supervised learning. Classification problems aim to determine the class of a given dataset, whereas each data point has multidimensional features that may be continuous or discrete. Logistic Steele Regression and Support Vector Machines are traditional algorithms that perform well on small-scale classification problems but ineffective on large-scale problems. For this reason, LSTM is a deep learning model for time-series data and is suitable for dealing with historical student data [19]. This paper aims to forecast the progress of educational management innovations in higher education and encourage ongoing advancement in education. It offers accurate predictions and guidance for managing university education and introduces new opportunities and breakthroughs in the education system.

3.2 Application of LSTM deep learning models

Long Short-Term Memory Network System LSTM is a deliberate product designed to solve long-time dependent questioning, a special recurrent neural network RNN, a neural network with long and short-term memory capability. Like other neural networks, RNNs have the concept of implicit states and can receive messages from other neurons and themselves. Multiple and ordered inputs are used in an RNN to simulate the order in which people read text and other sequential data and, after decoding, the implicit layer neurons. The corresponding memory capacity to better understand sequential data can be created by transmitting signals from the previous implicit layer neuron to the next implicit layer neuron.

The unit structure of LSTM is shown in Figure 2. The long and short-term memory network system LSTM, a variant of recurrent neural networks, introduces intrinsic states c_t and a gating mechanism, with three kinds of gates as input gate i_t, forgetting gate f_t, and output gate o_t. Forgetting gate Controls how much information needs to be forgotten from the previous intrinsic state c_t. Input gate i_t controls how much information needs to be saved for the current candidate state $\bar{c}$. Output gate o_t controls how much information is output from the current internal state c_t. By fully utilizing the memory ability and gating mechanism of LSTM, it provides accurate prediction and guidance for the education management of colleges and universities and brings new opportunities and breakthroughs to the education system.

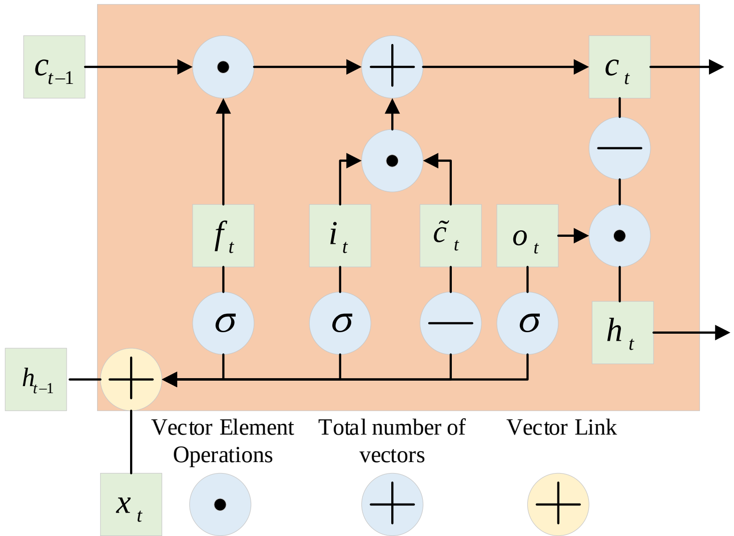


Figure 2. LSTM the module structure

The three gates take values between (0,1) for a certain percentage of the run-through, which is calculated by the formula:

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{1}$$

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{2}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{3}$$

Where $W*$, $U*$ and $b*$ are network parameters. The long distance temporal dependencies can be established by the above network unit LSTM, which is formulated as:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{4}$$

$$c_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{5}$$

During training, the LSTM is computed using the algorithm of backpropagation over time BPTT. Remember that the loss function of the LSTM network at moment t is $L_t = L\left(y, g\left(h_t\right)\right)$, where y_t is the supervised information at moment t and $g\left(h_t\right)$ is the output at moment t. Thus the loss function for the whole sequence $x_1 : T$ is $L = \sum_{t=1}^{T} L_t$.

According to the chain law, the LSTM model can be studied more deeply and comprehensively in the field of university education management through the back propagation process of the BPTT algorithm. First, the partial derivatives of the loss at moment t with respect to parameter U are computed $\frac{\partial L_t}{\partial U}$. For each hidden layer, the net input of parameter U is $Z_k = Uh_{k-1} + Wx_k + b$. Thus the gradient at moment t is:

$$\frac{\partial L_t}{\partial u_{ij}} = \sum_{k=1}^{t} \frac{\partial z_k}{\partial u_{ij}} \frac{\partial L_t}{\partial z_k} \tag{6}$$

Define the error term $\delta_{t,k} = \frac{\partial L_t}{\partial z_k}$ as the derivative of the loss at moment t with respect to the net input z_k of the hidden layer at moment k, then:

$$\begin{aligned} \delta_{t,k} &= \frac{\partial L_t}{\partial z_k} \\ &= \frac{\partial h_t}{\partial z_k} \frac{\partial z_{k+1}}{\partial h_t} \frac{\partial L_t}{\partial z_{k+1}} \\ &= diag\left(f'\left(z_k\right)\right) U^T \delta_{t,k+1} \end{aligned} \tag{7}$$

In summary, the above equation can be written in the form of a matrix:

$$\frac{\partial L_t}{\partial U} = \sum_{k=1}^{t} \delta_{t,k} h_{k-1}^T \tag{8}$$

The application of such deep learning algorithms is expected to provide colleges and universities with more accurate and efficient management decisions and promote innovation in higher education.

3.3 Predictive modeling for student career development

Gradient ascent is efficiently implemented and improved to enhance the operational efficiency of the student career development prediction model. The Gini coefficient is introduced to simplify the model calculation, which serves as an indicator for assessing modeling impurity and can help to identify valid features, thus enhancing the performance of the prediction model [20]. The arithmetic is as follows:

$$Gini\left(D\right) = \sum_{k=1}^{k} \frac{\left|C_k\right|}{\left|D\right|} \left(1 - \frac{\left|C_k\right|}{\left|D\right|}\right) \tag{9}$$

Where C_k represents the subset of samples in set D that belong to sample category k. The conditional Gini coefficients for a given feature A, for dataset D, are as follows:

$$Gini\left(D\middle|A\right) = \sum_{i=1}^{n} \frac{\left|D_i\right|}{\left|D\right|} Gini\left(D_i\right) \tag{10}$$

Where D_i denotes the subset of samples in D where feature λ takes the ih value. For the regression problem, CART uses the sum-variance measure to find the minimum mean squared error of the divided set, as shown in equation (11):

$$L = \min_{a,s}\left[\min_{c_1}\sum_{x_i\in D_1}(y_t-c_1)^2+\min_{c_2}\sum_{x_i\in D_2}(y_t-c_2)^2\right] \tag{11}$$

Where feature $a \in A$, corresponding to any division point s divides the dataset into D_1 and D_2, c_1 and c_2 are the sample output means of sets D_1 and D_2, and y_t is the model output value.

The final predicted value is obtained by accumulating the calculation results of multiple trees, and the residuals obtained from the previous tree are used to update the target value to realize the accurate prediction of career development. As shown in equation (13):

$$F_k(x) = F_{k-1}(x) + f_k(x) \tag{12}$$

Where $F(x)$ is the predicted value, $f(x)$ is the predicted residual value. At step t, the model predicted value $\hat{y}_i'$ of model f at the ith sample can be written as $\hat{y}_i' + f_t(x_i)$. Then, the objective function of the algorithm is formulated as follows when considering the regular term Ω:

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^t\right) + \sum_{i=1}^{t}\Omega(f_i) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{t-1} + f_t(x_i)\right) + \sum_{i=1}^{t}\Omega(f_i) \tag{13}$$

Realizing the second-order Taylor expansion of the objective function in $\hat{v}_{i=1}'$ out, the objective function can be written:

$$Obj^{(t)} = \sum_{i=1}^{n}\left[l\left(y_i, \hat{y}_i^{t-1}\right) + g_i f_t(x_t) + \frac{1}{2}h_i f_t^2(x_i)\right] + \sum_{i=1}^{t}\Omega(f_i) \tag{14}$$

Where g_i is the first order derivative of the loss function and h_i is the second order derivative of the loss function. The regular term of the objective function is calculated by equation (15):

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \tag{15}$$

Where T is the number of leaves, w is the weight of leaf nodes, and γ and λ are hyperparameters. In summary, the objective function can be written as:

$$\begin{aligned} Obj^{(t)} &\approx \sum_{i=1}^{n}\left[g_i f_t(x_t) + \frac{1}{2}h_i f_t^2(x_i)\right] + \Omega(f_i) \\ &= \sum_{i=1}^{n}\left[g_i w_{q(x_i)} + \frac{1}{2}h_i w_{q(x_i)}\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \\ &= \sum_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)w_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)w_j^2\right] + \gamma T \end{aligned} \tag{16}$$

Where $I=\{i|a(x_i)=j\}$ is the sample set of the j nd leaf node and $a(x)$ is the leaf node serial number. Let $G_j=\sum i\in I_j g_i, H_j=\sum i\in I_j h_i$ be a constant since G_j and H_j are the results of step $t-1$, and the optimal weight value is $w^*=-G_j/H_j+\lambda$. In summary, in order to get more accurate results in predicting students' career development. The objective function can be abbreviated as equation (17):

$$Obj^{(t)}=\sum_{j=1}^{T}\left[G_j w_j+\frac{1}{2}\left(H_j+\lambda\right)w_j^2\right]+\gamma T \tag{17}$$

4 Evaluation indicators for innovations in university management pathways

As a dichotomous classification problem, this paper adopts the ROC curve, which characterizes subjects' workability, as the main measure of the model. In the case of a dichotomous classification problem, instances can be categorized into positive and negative classes, and in the actual classification, the following four classes of results can be obtained:

1) True class TF, predicted positive class and actual positive class.
2) False negative class FN is predicted as a negative and positive class.
3) False positive class FP, predicted as a positive class, actual negative class.
4) True negative class TN, predicted as a negative class, actual negative class.

Table 1 shows how to obtain the confusion matrix based on the results, which can be used to obtain multiple evaluation labels.

Table 1. Confusion matrix for dichotomy problems

	Predicted As Positive Class	The Prediction Is Negative
True Is Positive	TP	FN
True Is Negative	FP	TN

The model's accuracy is obtained from equation (18), combining the four categories of results described above. However, the accuracy rate does not effectively reflect the model's performance when the samples are unbalanced. Therefore, TPR and FPR values are introduced, and the true case rate TPR and false positive rate FPR can be calculated by Eq. (19) and Eq. (20), and the AUC is used as a metric to measure the model's effectiveness. The higher the TPR and the lower the FPR, the better the classifier is proven to categorize the cases. The AUC value is the area under the line of the ROC curve, and the ROC uses the TPR as the vertical axis and the FPR as the horizontal axis.

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \tag{18}$$

$$TPR=\frac{TP}{TP+FN} \tag{19}$$

$$FPR=\frac{FP}{FP+TN} \tag{20}$$

Figure 3 Comparison of ROC curves and curves under the random selection strategy. First, the classification threshold is adjusted to zero, and then the classification threshold is set to the prediction value of each sample in turn. That is, the prediction model judges each sample as a positive example each time, obtains the coordinate points in turn, and finally connects all the points to obtain the ROC curve. The area under the ROC curve was calculated as the AUC value. The AUC value is less than or equal to 1. Larger values indicate that the model is more effective in classification.

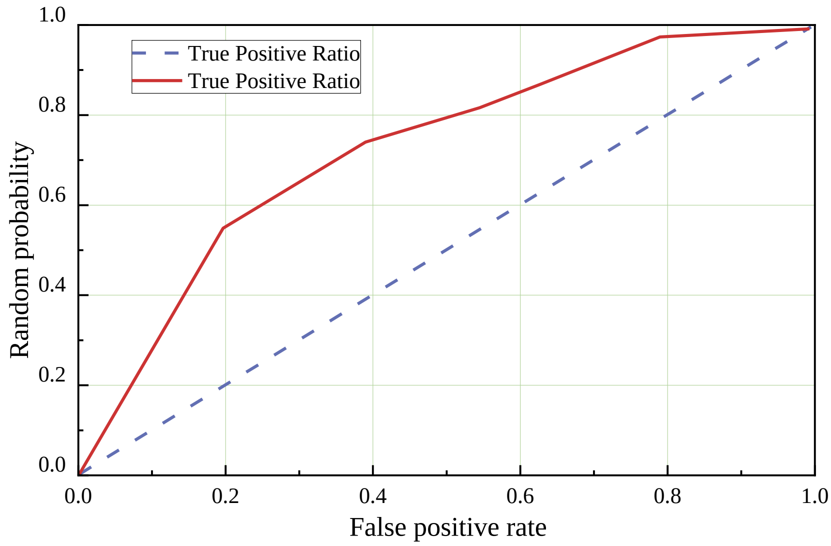


Figure 3. The ROC curve versus the curve under the random selection strategy figure

5 Student career development data collection and pre-processing

5.1 Exploratory data analysis

The data generated by Riiid is a very large dataset, containing approximately one million pieces of data, including information such as a student's historical performance, other students' performance under the same problem, and metadata about the problem. Exploratory data analysis can give an initial understanding of the data and greatly improve the efficiency of subsequent work. Table 2 displays the characteristics of the dataset. Features may play different roles in modeling students' career development prediction; some features may be used as input variables to the model to help predict students' career paths, while others may be used to assess student's learning behaviors and abilities. The rational use of these features during the modeling process can improve the accuracy and interpretability of the model.

Table 2. Data set characteristics

Feature Name	Type	Represent
Row Id	Int64	Data Line Number
Timestamp	Int64	Time Stamp
User__Id	Int32	User Number
Contented	Inti 6	User Interaction Number
Content_Type_Id	Int8	0 For Questions, 1 For Class
Task_Container_Id	Inti 6	Batch Number Of The question or course
User_Answer	Int8	User's Answer To The question
Answeredcorrectly	Int8	Do You Answer Correctly
Prior_Question_Elapsed_Time	Float32	Average Time To Solve problems
Prior_Question_Had_Explanation	Bool	Whether To View The Answer Or Resolve The Question
Questioned	Int64	Question Number
Bundle_Id	Int64	Problem Lot Number
Correct_Answer	Int8	Correct Response
Lecture_Id	Int32	Course Number
Tag	Int32	Problem Label
Part	Int8	Question Type Of Question
Type_Of	Int32	Course Purpose

This paper analyzes the relationship between user registration time and the role of deep learning models based on the dataset collected from the adopted platform. Figure 4 depicts the user registration time and interaction rate, with the highest frequency of interaction for users who have just registered.

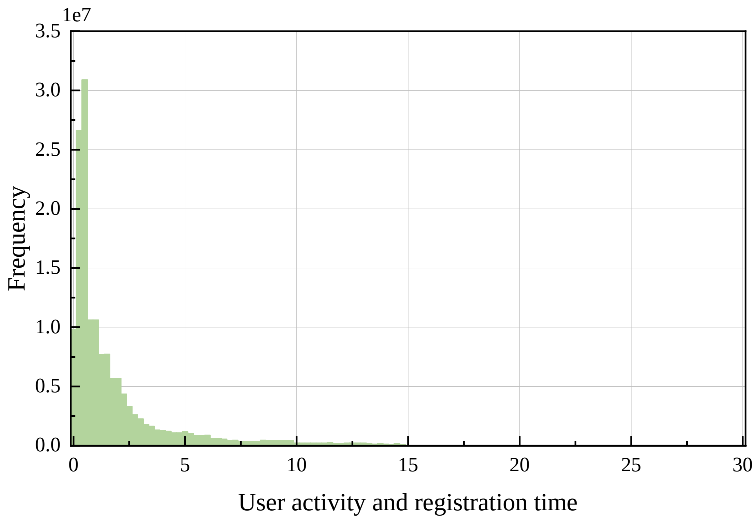


Figure 4. User registration time and interaction rate

Figure 5 displays user registration time and learning efficiency, categorizing users into five sections based on their registration length, with registration time increasing from left to right. Although the frequency of interaction is high, the learning efficiency for users who have just registered is relatively low.

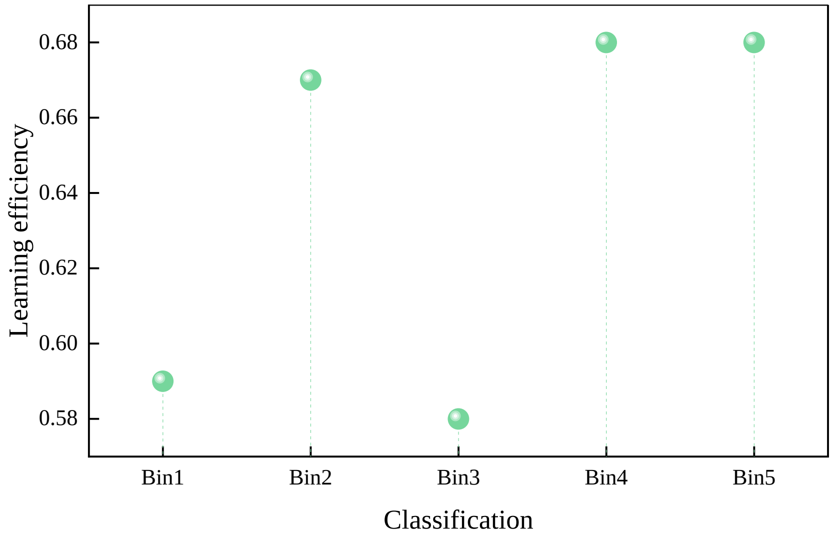


Figure 5. User registration time and learning efficiency

To improve the learning efficiency, the data about question-answering in the platform are aggregated based on the accuracy and the number of answers, from which 1000 data points are randomly sampled, and Figure 6 shows the relationship between the number of user answers and the accuracy rate. The user's accuracy rate varies when the number of answers is small. However, the accuracy rate shows an upward trend with the increase in the number of answers, indicating that the user's answer accuracy rate positively correlates with the answer volume to some extent.

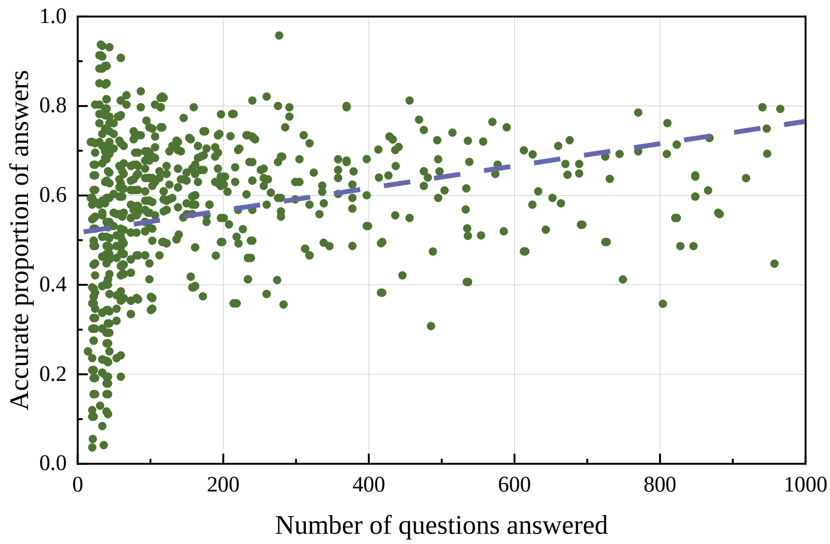


Figure 6. The relationship between the user's answer volume and the accuracy rate

Based on the data processing results above, the labeling data for the questions is analyzed below. Figure 7 depicts the 10 question labels with the highest and lowest correct rates. The analysis of these question labels can help gain insight into students' performance at different stages of career development, thus providing more effective pathway innovation solutions for educational management in higher education.

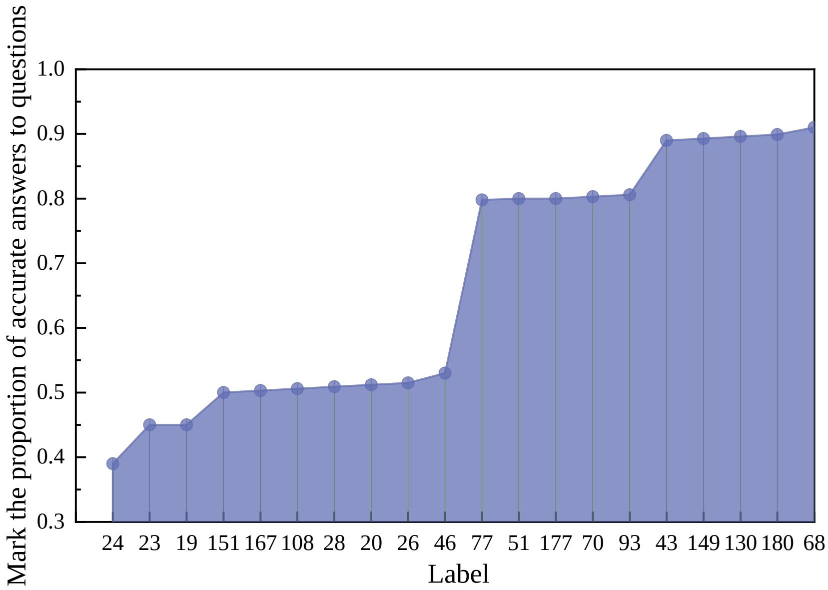


Figure 7. 10 problem labels with the highest and lowest accuracy

Figure 8 shows the 30 most common question labels, which help to generate more effective data features. Using LSTM, more accurate predictive models can be built to help students make more informed career development decisions.

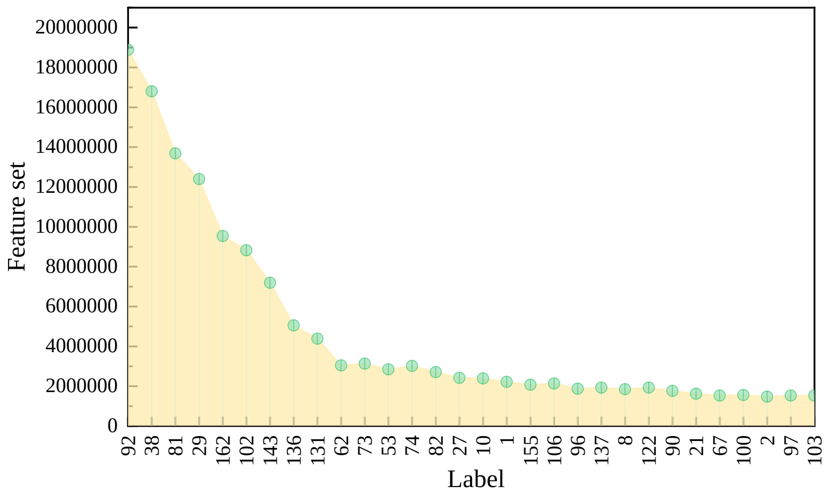


Figure 8. The 30 most common problem labels

The effect of course learning on the instant answer rate is shown in Figure 9. The users' class information is initially analyzed, and those who have undergone course learning are assigned the label True, and vice versa. The results show that the accuracy rate of the users who have undergone course learning is 13% higher than that of the users who have not, so whether or not they have gone through the course, learning can be used as a valid feature of the model. Students who have undergone coursework may pay more attention to their career development and proactively plan and prepare for their careers, thus getting higher prediction accuracy in the model. Introducing this feature provides more valuable information for the LSTM deep learning model and helps improve its predictive performance.

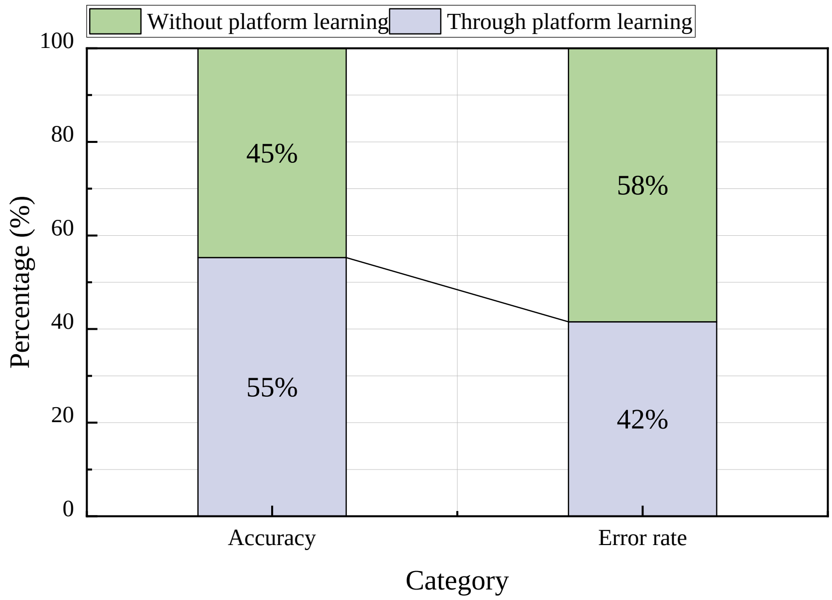


Figure 9. Effect of course learning on the instantaneous rate of answers

5.2 Data Cleaning and Feature Engineering

To fully utilize the data information for training, filling in the vacant data first is necessary since the data contains null values. Here are the exact steps:

1) Fill the null data in the field of priority_question_had_explanation as False. The student didn't check the explanation before answering the question, so fill in the missing data for answered_correctly as 0.5.

2) To prevent outliers from affecting the model accuracy, outlier detection was carried out on the data to remove data that was not three times the standard deviation.

3) Since the priority_question_had_explanation feature is a category feature, LabelEncoder encodes it so the model can process it.

The above data cleaning and feature engineering steps can make the dataset more complete and reliable and provide more effective and accurate training data for the research of educational management path innovation in universities based on the LSTM deep learning model. Table 3 shows the importance of features. The larger the weight, the more important the features are in the model, and features with negative weight can be eliminated. All the data features are retained because they contribute positively to the model. Among the features, the average accuracy of the question type and the average accuracy of the users themselves are the features that contribute the highest to the model.

Table 3. Permutation importance weight of the features

Feature Name	
Answered_Correctly_Content	0.1522±0.0011
Answered_Correctly_User	0.0743±0.0012
Sum	0.0380±0.0005
Prioi_Question_Elapsed_Time	0.0334±0.0007
Bundleid	0.0271±0.0011
Tagl	0.0260±0.0004
Tag2	0.0192±0.0005
Tag3	0.0181±0.0003
Part	0.0146±0.0004
Prior_Question_Had_Explanation	0.0002±0.0001

6 Analysis of the application of the LSTM deep learning model in the management of university education

6.1 Model Performance Comparison and Advantage Analysis

LSTM is a feasible algorithm choice in the innovation of educational management paths in colleges and universities, and the AUC performance is shown in Table 4. The AUC of LSTM is 0.756, while the AUC of machine learning is 0.749. It can be seen that LSTM slightly outperforms machine learning, with a slightly higher AUC value, which implies that the LSTM has a slightly better performance on the classification task. The average training time of LSTM is the average training time for LSTM is 41.8 s, while the average training time for machine learning is 41.8 s. The training speed of LSTM is much faster than machine learning, which may be because LSTM is an integrated learning algorithm based on deep learning, which is more efficient in training. The average model size of LSTM is 500 M, while the average model size of machine learning is 850 M. The model of machine learning is larger, which may be related to its decision tree structure and model parameters of the integration. In contrast, LSTM's models are relatively small, which helps to reduce resource consumption and model storage costs.

Table 4. A comparison of the results in the algorithm

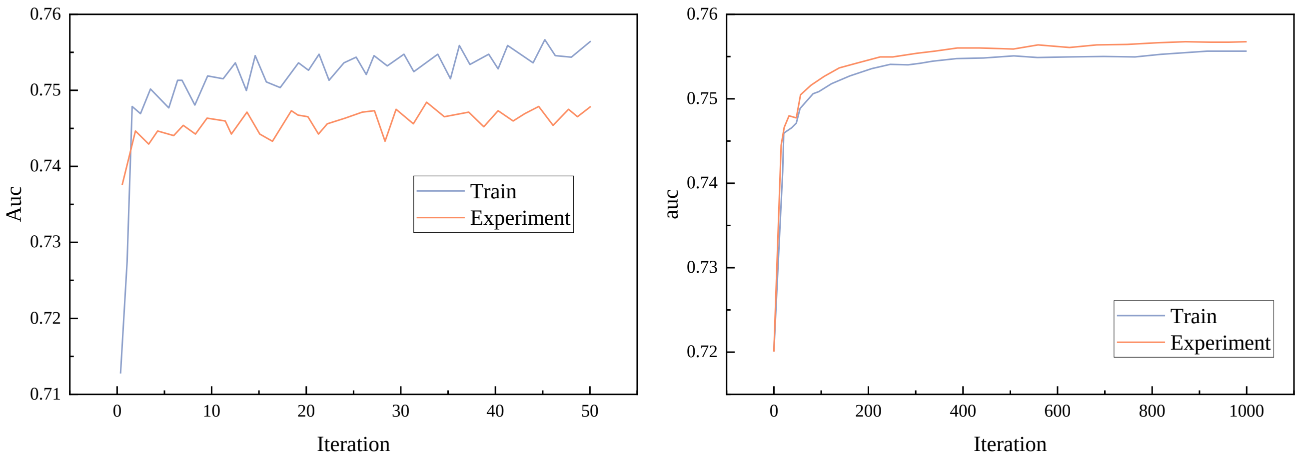
Algorithm	AUC Price	Average Training Duration	Average Model Size
LSTM	0.756	41.8s	500M
Machine learning	0.749	587s	850M

6.2 Impacts and comparisons of educational management

Introducing new features is very important to improve the accuracy to verify the effectiveness of introducing new features. This paper also compares the AUC values before and after the engineering of XGBoost features, and Figure 10 shows the results of the comparison between the LSTM deep learning model and artificial intelligence techniques. The training AUC curve of LSTM is shown in Figure 10(a), which shows that the algorithm has a value of 0.749 on the test set, the training time is 58.7 seconds, and the model size is about 500 M. The training AUC curve of the model of AI is shown in Figure 10(b), which shows that the AUC value of 0.756 on the test set, the training time of a single model is about 41.8 seconds, and the size of a single model is about 850 M.

The comparison experiment of the two algorithms reveals that the AI training AUC value is 0.749, while the AUC value of LSTM is 0.756, which indicates that the accuracy and efficiency of using the LSTM algorithm are higher, and it is more suitable to be used in the deep learning model for the research of education management path innovation in colleges and universities.

It should be noted here that before the introduction of feature engineering, the AUC value of the LSTM deep learning model was 0.732, which is a large difference in value and will affect the experimental results, so the introduction of feature engineering is an important step in processing data, which helps to be better applied to the deep learning research on the innovation of educational management paths in colleges and universities and to help college and university educational administrators to better understand the learning needs of the students and their behavioral patterns. Provide a more reliable basis for optimizing educational management and teaching methods. Therefore, introducing the combination of the LSTM deep learning model and feature engineering in the research of educational management path innovation in colleges and universities is a respectable method that can help achieve higher educational quality and student satisfaction.



(a) The AUC change curves for the LSTM training process

(b) Artificial Intelligence AUC change during training

Figure 10. Comparison of LSTM and Artificial Intelligence Techniques

7 Discussion

The research on developing educational management paths in colleges and universities using the LSTM deep learning model has broad development potential in the future. Through continuous technology optimization, data fusion and policy application, this research will provide more scientific and intelligent decision support for educational management in colleges and universities and promote students' career development and continuous progress in education.

The LSTM deep learning model can be further optimized in future research by adding more feature engineering and data preprocessing steps to improve the precision and accuracy of the prediction model. Exploration of more complex neural network structures and model architectures can lead to more refined prediction results. Or consider integrating datasets from different online education platforms and education data sources to expand the data size and increase the generalization ability and robustness of the model. Cross-platform dataset integration can aid in gaining a more complete understanding of students' learning and career development behaviors.

8 Conclusion

The research on college education management path innovation based on the LSTM deep learning model has achieved good results, which improves the accuracy and efficiency of the model through the introduction of new features and feature engineering and provides useful references for education management and decision-making in colleges and universities. LSTM, as a viable algorithmic option, excels in innovating education management paths in colleges and universities. Compared with machine learning, LSTM performs slightly better on the classification task, with an AUC value of 0.756, which is higher than the AUC value of 0.749 for machine learning. LSTM has a faster training speed, with an average training duration of 41.8 seconds, much faster than that of machine learning, which may be attributed to the higher training efficiency of LSTM as an integrated learning algorithm based on deep learning. LSTM's average model size of LSTM is about 500 M, compared to the average model size of machine learning, which is 850 M. The LSTM's small model size helps to reduce resource consumption and model storage costs.

However, the proposed study still has the abovementioned limitations that need to be addressed and improved in subsequent studies. The data sample used may be from a specific university or a specific region and, therefore, has limited applicability to other universities or regions. The limitations of the sample may result in limited generalization of the findings and the need for more data from the sample to improve the study's reliability and usefulness and promote the further development of deep learning applications for educational management in higher education.

Reference

[1] Liu, X., Zhang, X., Chen, W. W., & Yuan, S. M. (2020). Eye movement analysis of digital learning content for educational innovation. Sustainability, 12(6), 2518.

[2] Zhang, F., & She, M. (2021). Design of english reading and learning management system in college education based on artificial intelligence. Journal of Intelligent and Fuzzy Systems(2), 1-10.

[3] Dong, C., & Guo, Y. (2021). Improved differentiation classification of variable precision artificial intelligence higher education management. Journal of Intelligent and Fuzzy Systems(3), 1-10.

[4] Sue, B., Lori, L., & Shirley, A. (2018). Towards sustainable technology-enhanced innovation in higher education: advancing learning design by understanding and supporting teacher design practice. British Journal of Educational Technology.

[5] Silvia Monteiro, Isusi-Fagoaga, R., Almeida, L., & Adela Garcia-Aracil. (2021). Contribution of higher education institutions to social innovation: practices in two southern european universities. Sustainability 13.

[6] Rof, A., Bikfalvi, A., & Pilar Marquès. (2020). Digital transformation for business model innovation in higher education: overcoming the tensions. Sustainability, 12.

[7] Butler, Keith, T., Davies, Daniel, & W., et al. (2018). Machine learning for molecular and materials science. Nature.

[8] Caleb, Robinson, Bistra, Dilkina, Jeffrey, & Hubbs, et al. (2017). Machine learning approaches for estimating commercial building energy consumption. Applied Energy.

[9] Francisco, H. L., Mingorance-Estrada ángel, Juan, T. T., Aznar-Díaz Inmaculada, & Cáceres Reche María. (2018). Incidence of the flipped classroom in the physical education students' academic performance in university contexts. Sustainability, 10(5), 1334.

[10] Arunthavanathan, R., Khan, F., Ahmed, S., & Imtiaz, S. (2022). Autonomous fault diagnosis and root cause analysis for the processing system using one-class svm and nn permutation algorithm. Industrial & Engineering Chemistry Research(3), 61.

[11] Héctor Javier Fuentes López, & Bravo, L. E. C. (2020). A systematic review of techniques and classical machine learning factors used to predict academic performance. Sylwan, 164(7).

[12] Noroozi, O., Jrvel, S., & Kirschner, P. A. (2019). Multidisciplinary innovations and technologies for facilitation of self-regulated learning. Computers in Human Behavior, 100, 295-297.

[13] Denisov, I., Petrenko, Y., Koretskaya, I., & Beni, S. (2021). The gameover in universities education management during the pandemic covid-19: challenges to sustainable development in a digitalized environment. Sustainability, 13.

[14] Cavus, N., Mohammed, Y. B., & Yakubu, M. N. (2021). Determinants of learning management systems during covid-19 pandemic for sustainable education. Sustainability, 13(9), 5189.

[15] Suanj, Z., Jakopec, A., & Ori, A. (2020). Academics' effectiveness and professional development in croatia: challenges for human resource management in higher education institutions. European Journal of Education, 55(4).

[16] Kariapper, R., Samsudeen, S. N., & Fathima, S. M. B. (2020). Quantifying the impact of online educational system in teaching and learning environment among the teachers and students.Solid State Technology,63(6), 12118-12132.

[17] Okada, A., Whitelock, D., Holmes, W., & Edwards, C. (2018). E-authentication for online assessment: a mixed-method study. British Journal of Educational Technology.

[18] Noroozi, O., Jrvel, S., & Kirschner, P. A. (2019). Multidisciplinary innovations and technologies for facilitation of self-regulated learning. Computers in Human Behavior, 100, 295-297.

[19] Li, X., Dong, Y., Jiang, Y., & Ogunmola, G. A. (2022). Analysis of the teaching quality of college ideological and political education based on deep learning. Journal of Interconnection Networks, 22(Supp02).

[20] Schneider, M. (2020). The discovery of the gini coefficient. History of Political Economy(1), 115-141.