# A Study on the Evaluation of AI's Impact on University Students' Learning Based on Data Analysis and Machine Learning

Huihao Chang[1, *, #], Wenqian Liu[2, #], Lidong Chu[1, #], Xiaoyue Li[2, #]

[1]School of Architecture and Engineering, Northeast Electric Power University, Jilin, China, 132012

[2]School of Economics and Management, Northeast Electric Power University, Jilin, China, 132012

* Corresponding Author Email: 13955772268@163.com

#These authors contributed equally.

**Abstract.** With the rapid development of artificial intelligence (AI) technology, its application in the field of education has significantly affected the learning style of college students. This article explores the impact of artificial intelligence on college students' learning. It uses optimized evaluation algorithms and survey analysis to demonstrate AI's positive effects. The study begins with cleaning and correcting survey data, identifying textual features, and encoding variable types. After standardizing the data using the Min-max method, validity tests and descriptive statistical analysis were conducted. The results indicate high reliability and correlation in students' attitudes towards AI tools and their learning styles. An evaluation system was constructed based on these findings, selecting 11 key indicators through the Fisher algorithm model and assessing them with the LightGBM quantitative model. The model shows that AI influences college students' learning primarily through factors like learning purposes, safety of tool use, motivation, usage time, and satisfaction. This evaluation system not only confirms AI's positive impact but also offers optimization goals and methods to enhance students' learning abilities in the AI era.

**Keywords:** Data Encoding, Validity Test, Bartlett Sphericity Test, Fisher Algorithm, LightGBM Quantitative Model.

## 1. Introduction

With the rapid development and widespread application of artificial intelligence technology, the role of AI in the field of education is becoming increasingly prominent, especially in the learning and teaching process of college students[1]. AI tools such as intelligent tutoring systems, adaptive learning platforms, and online educational resources are changing the way students learn and the interactive mode of education. However, while AI technology improves educational efficiency and personalized learning experiences, it also raises concerns about the impact on learning outcomes, student psychology, and behavioral patterns[2]. Therefore, accurately assessing the extent and methods of AI's impact on college students' learning has become an important research topic in the field of educational technology[3].

The Bartlett sphericity test and LightGBM quantification model have been widely applied in many fields such as economics, sports, and transportation[4-6].In this context, this study aims to explore the deeper applications of data analysis and machine learning methods in specific fields, to provide more accurate decision support and predictive analysis.

To study the effectiveness and practicality of artificial intelligence learning in improving educational quality, optimizing educational resource allocation, promoting educational policy reform, and promoting educational innovation, investigators designed a survey questionnaire based on this study and obtained feedback results (survey data). This study preprocessed and analyzed the survey data, and then quantitatively evaluated the impact of AI on college student learning and constructed a model, proving the effectiveness and practicality of artificial intelligence learning in improving educational quality, optimizing educational resource allocation, promoting educational policy reform,

and promoting educational innovation. Propose optimization goals and innovative ways to achieve them, promote the development of artificial intelligence, and promote the vigorous development of college students' learning abilities in the wave of artificial intelligence.

## 2. Data Preprocessing and Feature Analysis.

To address the issue of how to analyze and process survey data, it is necessary to analyze and digitize the survey data. Firstly, conduct a preliminary analysis of the survey data, clean and correct the data, and identify the data characteristics of the survey data. Secondly, data encoding is applied to categorical variables, numerical variables, and sequential variables for each variable. Direct assignment and binary hot coding methods are used to encode data for both individual and multivariate variables. Furthermore, the Min-max method is used to standardize the survey data and output preprocessed sample data. Finally, discrimination tests and mathematical statistics are conducted on the preprocessed sample data, and the characteristics and analysis conclusions of the data are obtained based on the output results.

### 2.1. Data Analysis

This question is about text data processing and data analysis. Firstly, data preprocessing is required, so data processing is particularly important. In response to another issue in this article, it is necessary to consider the dimension of data quantification to extend other comprehensive analyses. However, due to the large amount of text and sample size in the survey data, preliminary screening and analysis of the data are required, and data preprocessing is also required. The data processing method is shown in the following figure.

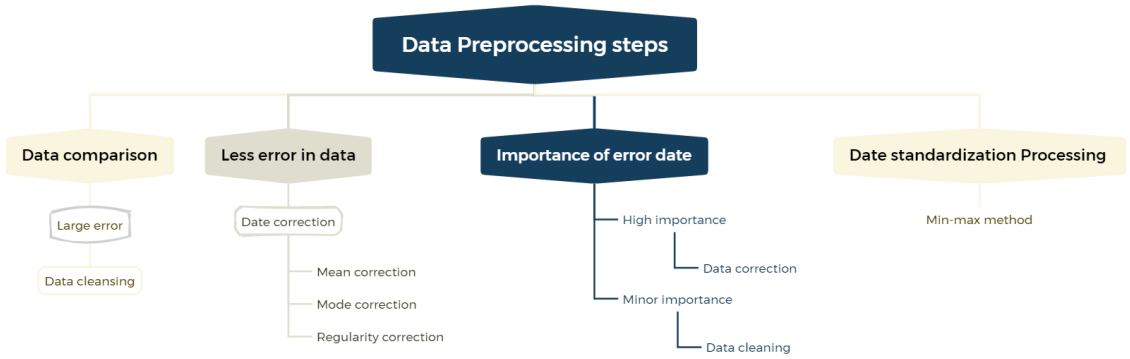



**Figure 1.** Data preprocessing steps

According to Figure 1, a preliminary analysis was conducted on the data, and it was found that there were no missing values or error values found in the questionnaire data. Therefore, further analysis can be conducted on the data, and descriptive statistics can be performed. Some of the results are shown in Figure 2.

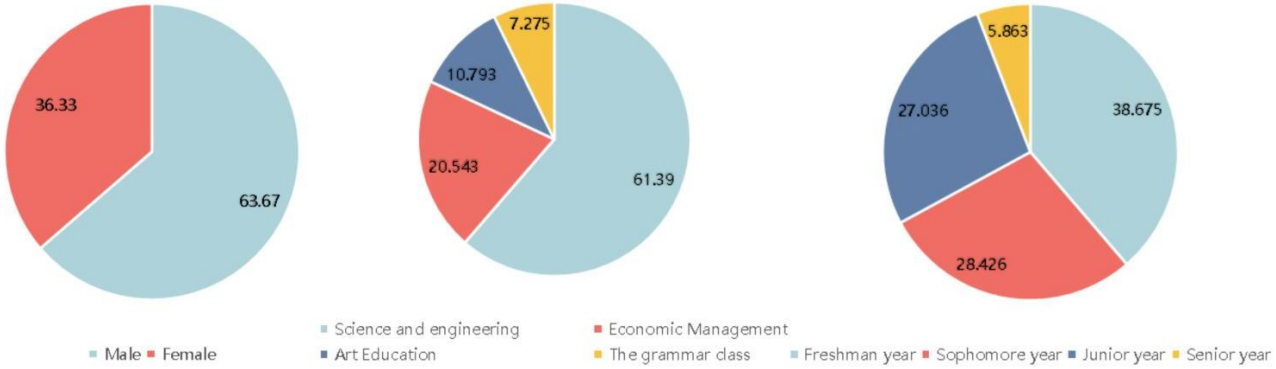



**Figure 2.** Descriptive statistical visualization image of partial data

According to Figure 2, we can obtain the gender ratio and subject distribution ratio in the questionnaire data, with males and science and engineering being the majority.

### 2.2. Data Encoding

Based on the analysis above, we can conclude that the survey data is composed of a large amount of textual data. As the data is textual, to achieve analysis of the problem, we need to encode the data. The rules for encoding the data are shown in the following figure3.

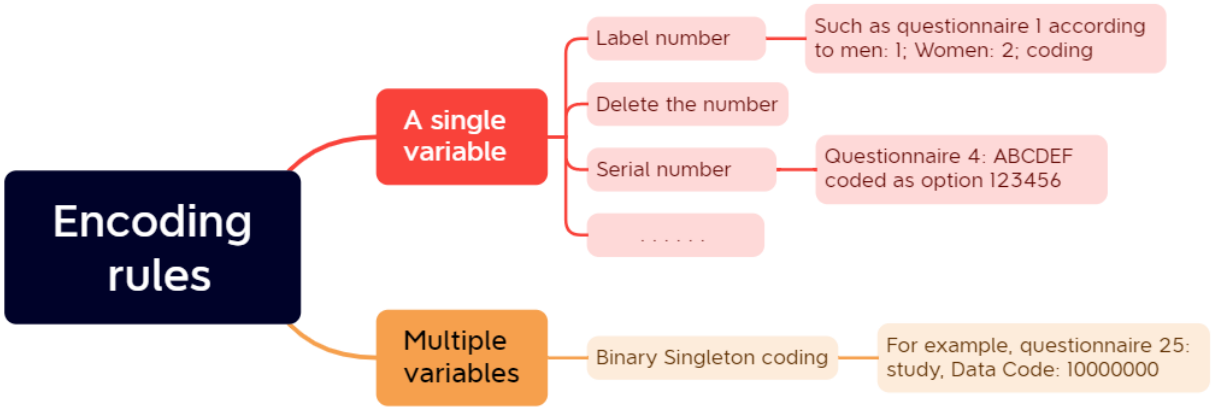



**Figure 3.** Encoding Rules

### 2.3. Data Standardization

This question is a data analysis question, and the dimensions of each indicator are different. Standardization (normalization) processing is a basic work of data mining. Different evaluation indicators often have different dimensions and units, which can affect the results of data analysis. The dimensions of indicator data in the survey data are different. To eliminate the dimensional impact between indicators, data standardization is needed to solve the comparability between data indicators. After standardizing the raw data, all indicators are in the same order of magnitude, making it suitable for comprehensive comparative evaluation.

This article applies the Min-max normalization data processing method for data standardization. The Min-max normalization method, also known as Min-max scaling, is a linear transformation of the original data, mapping the resulting values to between [0,1]. The conversion function is as follows[7]:

$$x_{ij} = \frac{\max\{x_{1j},.....,x_{nj}\} - x_{ij}}{\max\{x_{1j},.....,x_{nj}\} - \min\{x_{1j},.....,x_{nj}\}} \quad (1)$$

In equation (1) above, max is the maximum value of the sample data, and min is the minimum value of the sample data. We apply equation (1) to standardize the encoded survey data, and the processed data is shown in the following figure.

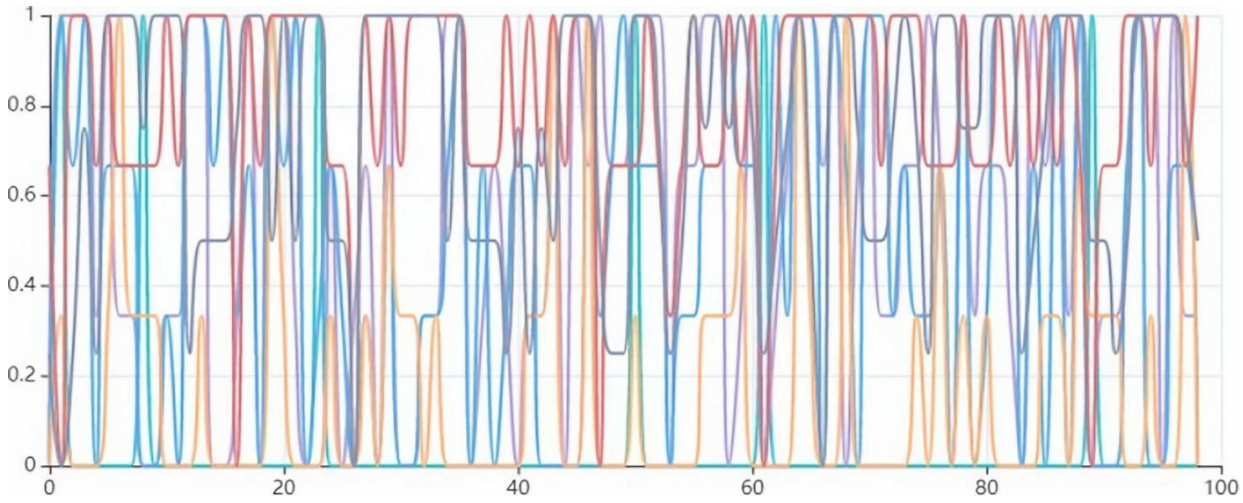



**Figure 4.** Partly visualized images of data standardization processing

According to Figure 4, the standardized data after encoding can be obtained, and then these data can be used for the next step of data feature analysis.

### 2.4. Data Feature Analysis

To explore the characteristics of survey data, we conduct validity and discriminability tests on the preprocessed survey data, reasonably evaluate the validity and enforceability of the data in the survey data, and output the characteristics of the data.

#### 2.4.1 Descriptive Statistics

Descriptive statistics were conducted on all indicators using SPSS, and the results obtained are shown in the following table.

**Table 1.** Descriptive statistical results of some data

| Variable Name | Sample Variable | Max | Min | Average | Standard Deviation | Median | Variance | Kurtosis | Skewness | Coefficient of Variation(CV) |
|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 4605 | 3 | 1 | 2.02 | 0.825 | 2 | 0.68 | -1.528 | -0.036 | 0.035 |
| 21 | 4605 | 4 | 1 | 2.051 | 1.157 | 2 | 1.338 | -1.304 | 0.507 | 0.049 |
| 22 | 4605 | 4 | 1 | 2.888 | 0.483 | 3 | 0.233 | 7.208 | -2.237 | 0.014 |
| 30 | 4605 | 4 | 1 | 2.147 | 1.112 | 2 | 1.237 | -1.18 | 0.454 | 0.045 |
| 25 | 4605 | 4 | 1 | 2.31 | 1.173 | 2 | 1.375 | -1.381 | 0.322 | 0.044 |
| 19 | 4605 | 3 | 1 | 1.879 | 0.538 | 2 | 0.29 | 0.24 | -0.09 | 0.025 |

According to Table 1, the anomaly coefficient of the data is relatively small, with a maximum not exceeding 0.05, indicating that the questionnaire data indicators are more accurate and have fewer outliers.

#### 2.4.2 Test of Distinction

Through our understanding, the purpose of discriminant analysis is to study whether data can effectively distinguish high and low levels, to evaluate the quality of a specific item. It usually divides a piece of data into low, medium, and high groups using 27% and 73% as distinguishing criteria. Data below the 27% percentile is referred to as the low group, while data between 27% and 73% is referred to as the medium group; If it is higher than 73%, it is called a high group. Use a T-test to determine whether the P-value is significant ($P<0.05$). If it is significant, perform a difference analysis based on the mean and test value, and describe the size of the difference. If there is a difference, it indicates that the scale item design is appropriate. Otherwise, it indicates that the scale item cannot distinguish information, and if the design is unreasonable, it should be deleted. The following table2 shows the discriminability test results of our data.

**Table 2.** Differentiation Results

| Number | Group (mean ± standard deviation) | | | t | p |
|---|---|---|---|---|---|
| | 0%-27%(n=1244) | 27%-73%(n=2117) | 73%-100%(n=1244) | | |
| 22 | 2.75±0.656 | 2.908±0.429 | 2.992±0.297 | -11.846 | 0.000*** |
| 21 | 2.233±1.264 | 1.93±1.103 | 2.076±1.108 | 3.289 | 0.001*** |
| 20 | 2.756±1.406 | 3.399±1.142 | 3.82±0.667 | -24.117 | 0.000*** |
| 19 | 1.719±0.613 | 1.898±0.508 | 2.007±0.465 | -13.223 | 0.000*** |
| 17 | 1.343±0.599 | 1.505±0.664 | 1.736±0.731 | -14.672 | 0.000*** |
| 15 | 1.506±0.701 | 2.042±0.801 | 2.495±0.666 | -36.099 | 0.000*** |
| 14 | 1.453±0.67 | 2.021±0.803 | 2.511±0.652 | -39.897 | 0.000*** |
| 13 | 1.359±0.606 | 1.831±0.785 | 2.367±0.732 | -37.372 | 0.000*** |
| 12 | 1.122±0.328 | 1.176±0.381 | 1.249±0.433 | -8.253 | 0.000*** |
| 8 | 1.523±0.824 | 1.677±0.932 | 1.871±1.048 | -9.208 | 0.000*** |
| 7 | 1.08±0.271 | 1.045±0.208 | 1.047±0.211 | 3.387 | 0.001*** |
| 6 | 3.018±1.272 | 3.621±1.227 | 4.009±1.131 | -20.515 | 0.000*** |
| 23 | 2.755±1.076 | 3.334±0.839 | 3.645±0.622 | -25.284 | 0.000*** |
| 25 | 1.603±0.846 | 2.31±1.127 | 3.016±1.102 | -35.874 | 0.000*** |
| 30 | 1.483±0.773 | 2.125±1.061 | 2.85±1.064 | -36.652 | 0.000*** |

Note: * * *, * *, * * represent significance levels of 1%, 5%, and 10%, respectively

According to the above table, it can be concluded that most indicators show significant levels, rejecting the null hypothesis, indicating that the scale item design has high discrimination and is relatively reasonable.

### 2.4.3 Validity Test

Validity test refers to the analysis of the accuracy of a scale in achieving measurement indicators. There are many methods for validity testing, and commonly used in research are item analysis, independent criterion validity analysis, and factor analysis. Validity analysis is generally divided into three dimensions analysis, namely the KMO test, gravel plot analysis, and factor loading coefficient statistics[7].

Firstly, a KMO test is conducted, and the KMO test results for randomly selected indicators are shown in Table 3.

**Table 3.** KMO Inspection Results Table

| KMO Value | | 0.765 |
|---|---|---|
| Bartlett Sphericity Test | Approximate Chi-squared value | 14614.7 |
| | df | 105 |
| | P | 0.000*** |

Note: * * *, * *, * * represent significance levels of 1%, 5%, and 10%, respectively

The results of the KMO test showed a value of 0.765, the value of KMO is 0.765, while Bartlett's sphericity test showed a significant P-value of 0.000 * * *, indicating significance at the horizontal level. The null hypothesis was rejected, and there was a correlation between the variables. Factor analysis was effective, with a moderate degree.

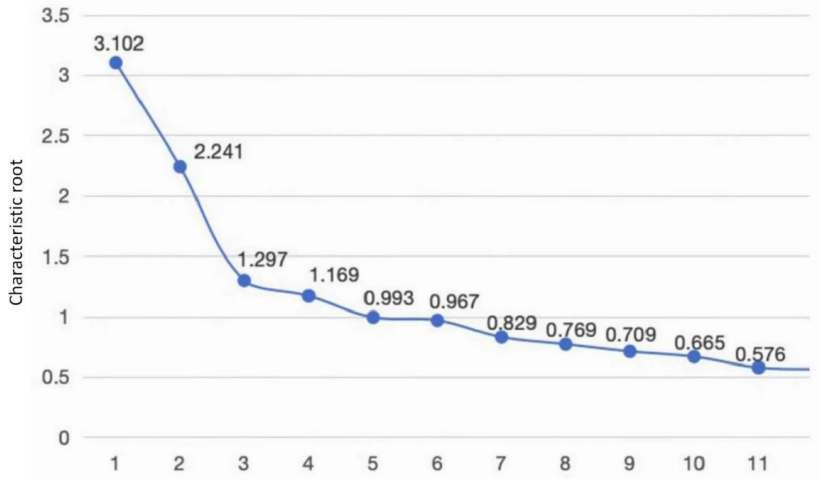



**Figure 5.** Crushed Stone Diagram

From the Crushed Stone Diagram in Figure 5, it can be concluded that the characteristic values of the indicators tend to flatten after the first five factors. Therefore, we will use the five dimensions as principal component analysis, laying the foundation for subsequent evaluation analysis.

So the factor Heat map of the factor load matrix for each indicator output is shown in Figure 6.

| | | | | | | |
|---|---|---|---|---|---|---|
| 30、Which learning link … | -0.032 | -0.008 | 0.805 | 0.077 | -0.023 | 0.656 |
| 25、After entering the … | -0.023 | -0.057 | 0.818 | 0.029 | 0.025 | 0.674 |
| 23、What are your main … | 0.135 | -0.490 | 0.508 | 0.181 | 0.028 | 0.550 |
| 6、How much time do you … | 0.030 | -0.448 | 0.071 | -0.101 | 0.611 | 0.591 |
| 7、Have you used any … | 0.096 | 0.725 | -0.093 | 0.052 | 0.053 | 0.548 |
| 8、When do you use the … | 0.124 | 0.324 | -0.082 | 0.055 | 0.672 | 0.582 |
| 12、If there are artificial … | 0.265 | 0.600 | -0.063 | -0.087 | -0.006 | 0.442 |
| 13、 Do you have the idea … | 0.821 | 0.108 | -0.028 | 0.030 | -0.083 | 0.694 |
| 14、 Do you have the idea … | 0.875 | 0.023 | -0.029 | 0.074 | -0.058 | 0.775 |
| 15、 Do you have any idea … | 0.868 | 0.054 | -0.051 | 0.052 | -0.043 | 0.763 |
| 17、Do you agree with the … | 0.391 | 0.478 | 0.040 | -0.271 | -0.035 | 0.457 |
| 19、What is your opinion … | 0.418 | 0.430 | 0.112 | -0.159 | 0.208 | 0.441 |
| 20、If you use AI learning … | 0.024 | -0.143 | 0.131 | 0.630 | -0.268 | 0.506 |
| 21、Do you think AI tools … | -0.507 | 0.042 | -0.156 | 0.059 | -0.380 | 0.431 |
| 22、When AI tools are … | 0.043 | -0.002 | 0.057 | 0.814 | 0.161 | 0.693 |

**Figure 6.** Heat map of factor load matrix

From Figure 6, it can be concluded that the importance of latent variables in each principal component is significant. In summary, questionnaire indicators such as whether college students agree with the use of artificial intelligence tools and whether they learn through artificial intelligence methods are of high importance[8].

## 3. Quantitative Evaluation and Modeling of AI's Impact on University Students' Learning AI

Regarding the rationality of selecting evaluation indicators based on the analysis results of data, and constructing an evaluation indicator system, to construct an effective evaluation system, this article first applies the Fisher algorithm model for semantic detection of all indicator data and

conducts priority selection to determine which aspects have the most important impact on the learning of college students, to discuss the rationality of selecting evaluation indicators based on the analysis results of data and discuss the construction of an evaluation indicator system. Secondly, to achieve a quantitative evaluation of the impact of AI on the learning of college students, the indicators defined by labels are bi-directional encoded. The positive encoding has a positive impact on the learning of college students, while the negative encoding represents a negative impact on their learning. Furthermore, by constructing a LightGBM quantitative model, the differentially encoded indicators are comprehensively scored, and the indicators with higher scores are output. Finally, by conducting statistical tests on the model, the scientificity and rationality of the model are outputted, and an evaluation index system for the impact of AI on the learning of college students is also outputted.

### 3.1. Fisher Semantic Detection Model

Fisher is a text analysis algorithm tool based on NumPy, which was created to solve text data analysis. Fisher incorporates a large number of libraries and some standard data models, providing the tools needed to efficiently manipulate large text sets. Fisher semantic analysis provides a plethora of functions and methods that enable us to quickly and conveniently filter data. With well-defined semantic information conditions, it can quickly achieve text data classification. We used Fisher in Python to perform priority evaluation and analysis on the data and filtered the data. Finally, we identified a total of 11 indicators with higher priority, which can represent the impact of AI on the learning of college students.

### 3.2. Bidirectional Encoding

To better study the impact of artificial intelligence on the learning of college students, we performed secondary encoding on the survey data. This secondary encoding was to construct a bidirectional encoding, with positive encoding promoting college students' learning and negative encoding inhibiting their learning. So the rules for bidirectional encoding are shown in the following figure[9].

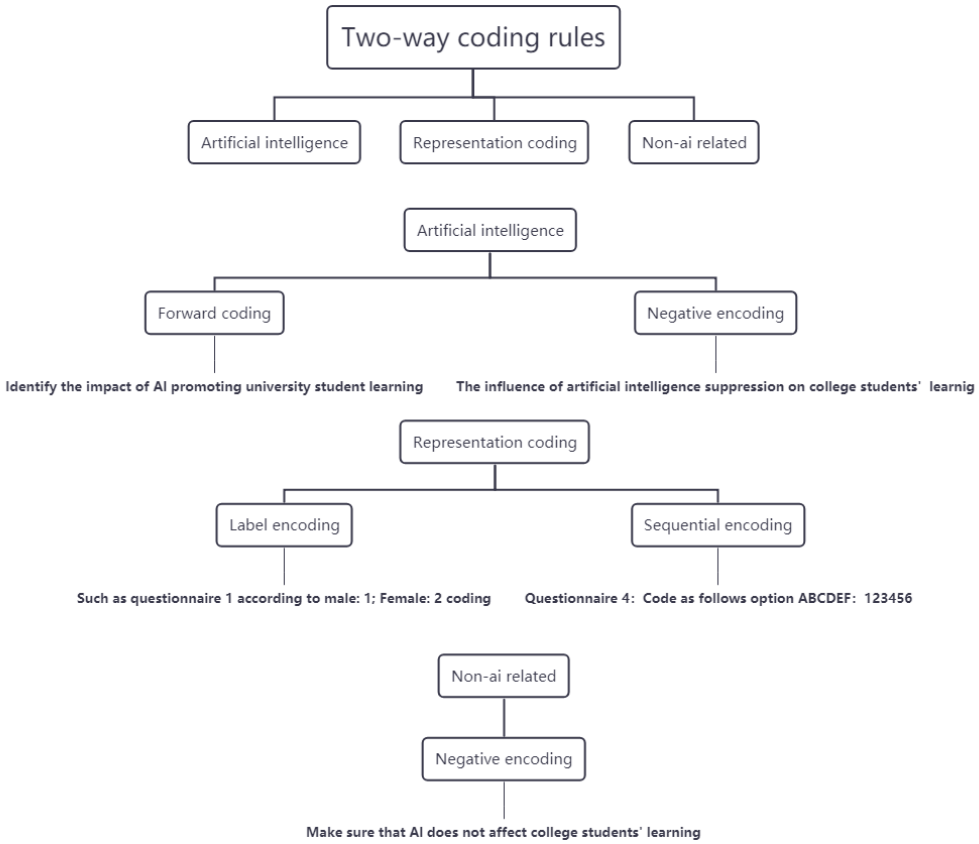



**Figure 7.** Bidirectional Encoding Rules

According to the bidirectional encoding rule in Figure 7, the indicators are subjected to secondary encoding, and the output encoding data is shown in the following figure.

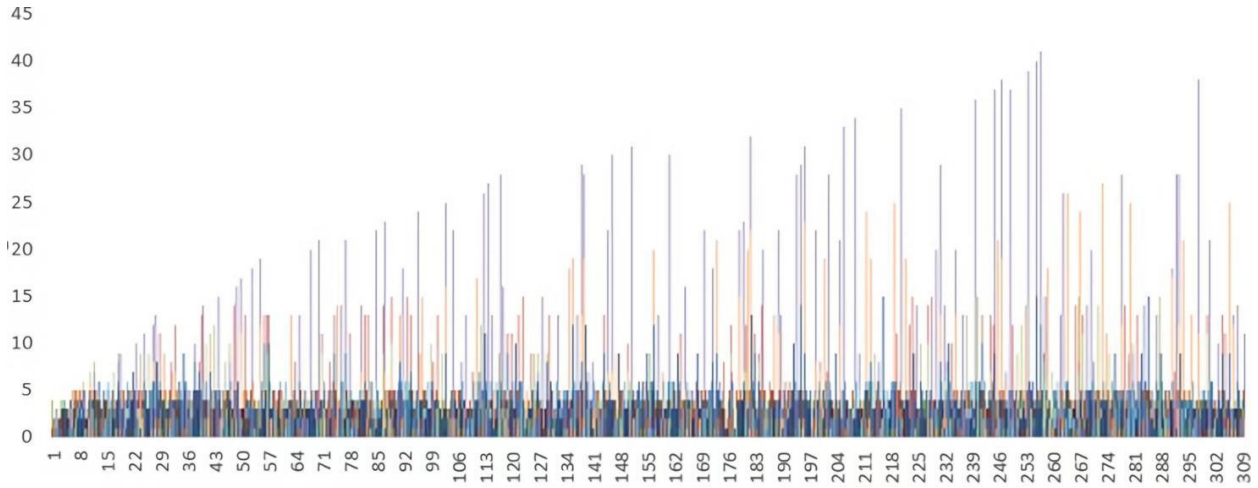



**Figure 8.** Visualization of new encoded data

According to Figure 8, the differential encoding results of all indicators can be obtained, which distinguish whether artificial intelligence has an impact on the learning of college students through bidirectional encoding.

### 3.3. Construction of LightGBM Quantitative Indicator Evaluation Model

#### 3.3.1 Establishment of LightGBM model

This question will establish a quantitative indicator evaluation model based on LightGBM. The effective features extracted after data preprocessing are used as input data for the model, and the evaluation values of each indicator are used as output data. LightGBM is an efficient implementation of GBDT, which generally involves iteratively fitting the residuals of the current model and combining weak learners into strong learners. LightGBM has multiple steps, such as second-order Taylor expansion of the objective function, calculation of leaf node values, and expression of tree complexity. The following will provide a specific analysis of its establishment process[10].

Assuming there are n samples, the dataset with m-dimensional features for each sample is $D=\{(x_i,y_i)\}(|D|=n,x_i\in\Re^m,y\in\Re)$, and the prediction function composed of K basis functions is[10]:

$$\hat{y}_i=\phi\ (X_i)=\sum_{k-1}^{K} f_k(X_i),\ f_k\in F \tag{2}$$

Among them, $D=\{(x_i,y_i)\}(|D|=n,x_i\in\Re^m,y\in\Re)$ is the function space containing all regression trees, and q represents the mapping from the sample space to the leaf nodes, which is the structural parameter of the regression tree; $w_i$ represents the value of the i-th leaf node, T is the total number of leaf nodes in the regression tree, and each weak learner $f_k$ has its own independent tree structure parameter q and complete leaf node w. Given the sample point $x_i$, the order is to use K regression tree rules (determined by q) to divide the sample points into different leaf nodes and add the w values obtained in each regression tree to obtain the predicted value $\hat{y}_i$[11].

Input: Training dataset $D=\{(x_i,y_i)\}(|D|=n,x_i\in\Re^m,y\in\Re)$ with L iterations;
Output: The trained LightGBM algorithm model;
The steps are as follows:

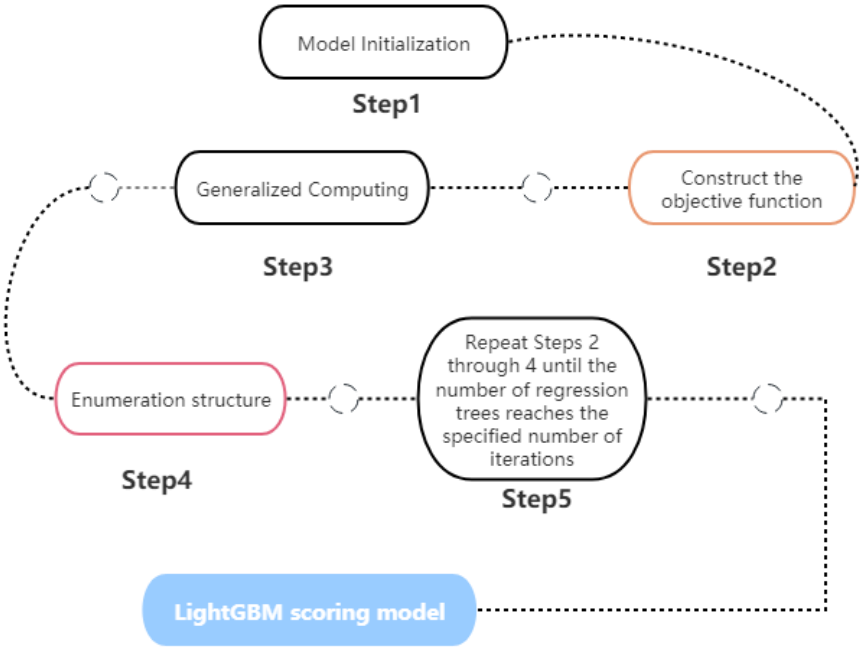



**Figure 9.** LightGBM scoring model diagram

Through the above steps (Figure 9), the final LightGBM scoring model is obtained, which is:

$$\hat{y}_i = \sum_{t=1}^{l} f_t(a_i) \tag{3}$$

### 3.3.2 Answer to the model

Build the model using Matlab, import the bi-directional encoded data, and output the rating of the indicators as shown in the following figure.

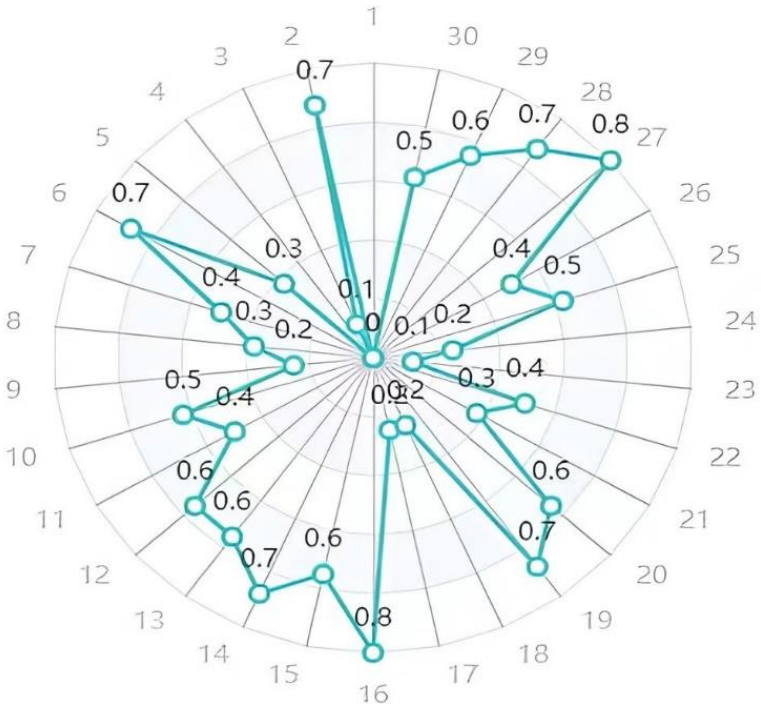


**Figure 10.** LightGBM Quantitative Score

According to Figure 10, the dimensions of the impact of artificial intelligence on the learning of college students are shown in Table 4.

**Table 4.** Dimensions of the Impact of AI on the Learning of College Students

| Questionnaire questions | Score | Effect |
|---|---|---|
| 16、If there are artificial intelligence learning tools, what are the reasons that affect your use of artificial intelligence learning software? | 0.8 | Forward |
| 27、What security have you considered using AI tools? | 0.8 | Forward |
| 6、How much time do you spend online each week? | 0.7 | Forward |
| 10、Do you want to access the learning resources of colleges and universities across the country? | 0.7 | Forward |
| 28、Which of the following aspects do you think are important for AI learning tools? | 0.7 | Forward |
| 21、Do you think artificial intelligence tools can replace teachers in the future? | 0.3 | Reverse |
| 9、Will you share your information with others on the internet? | 0.2 | Reverse |

According to Figure 10 and Table 4, the quantified values of 16, 27, 6, 10, 28, 21, and 9 are 0.8, 0.8, 0.7, 0.7, 0.7, 0.3, and 0.2, respectively. In summary, the evaluation system for the impact of AI on the learning of college students is shown in Table 5. Artificial intelligence mainly affects college students' learning by influencing their learning purposes, AI tool usage safety, learning motivation, AI usage time, and satisfaction with usage.

## 4. Conclusions

In summary, questionnaire indicators such as whether college students agree with the use of artificial intelligence tools and whether they learn through artificial intelligence methods are of high importance; At the same time, it can be inferred that artificial intelligence mainly affects the learning of college students by influencing their learning purposes, AI tool safety, learning motivation, AI usage time, and satisfaction with usage.

To address the issue of how to analyze and process survey data, first conduct a preliminary analysis of the survey data, clean and correct the data, and obtain the textual data features of the survey data. Secondly, data encoding is applied to categorical variables, numerical variables, and sequential variables for each variable. Direct assignment and binary hot coding methods are used to encode data for both individual and multivariate variables. Furthermore, the Min-max method is used to standardize the survey data and output preprocessed sample data. Finally, validity and discriminability tests, as well as descriptive statistical analysis, were conducted on the preprocessed sample data. The KMO value of the questionnaire data was 0.765, and Bartlett's sphericity test showed a significant P-value of 0.000 * * *, showing significance at the horizontal level. There was a correlation between the variables, and questionnaire indicators such as whether college students agreed to use artificial intelligence tools and whether they learned through artificial intelligence were of high importance.

To address the issue of selecting evaluation indicators based on the analysis results of data and constructing an evaluation indicator system, this article first applies the Fisher algorithm model for semantic detection on all indicator data and conducts priority selection to determine a total of 11 priority indicators, including 17, 19, 20, 21, 22, 27, 30, 12, 13, 14, and 15, representing the impact of AI on college students' learning. Secondly, to achieve a quantitative evaluation of the impact of AI on the learning of college students, the indicators defined by labels are bi-directional encoded. The positive encoding has a positive impact on the learning of college students, while the negative encoding represents a negative impact on their learning. Furthermore, by constructing a LightGBM quantification model and comprehensively scoring the differentially encoded indicators, the quantified values of 16, 27, 6, 10, 28, 21, and 9 were obtained as 0.8, 0.8, 0.7, 0.7, 0.7, 0.3, and 0.2, respectively. Finally, through statistical testing of the model, it was found that the accuracy of the model is high. The constructed evaluation system can conclude that artificial intelligence mainly

affects the learning of college students by influencing their learning purposes, AI tool usage safety, learning motivation, AI usage time, and usage satisfaction.

This article provides a research approach and framework applied to various fields related to educational technology evaluation, such as teaching effectiveness evaluation, technology application evaluation, teaching resource evaluation, policy and standard evaluation, and educational innovation evaluation. It proves the effectiveness and practicality of artificial intelligence learning in improving educational quality, optimizing educational resource allocation, promoting educational policy reform and innovation, as well as promoting the development of artificial intelligence and enhancing the learning ability of college students.

## References

[1] Han Yongli, Research on the Development of Education and Teaching in the Context of "Internet Plus" [J]. Ningxia People's Education Press, 2021,163

[2] Jiang Huaibin and Zhang Bin, Research Methods in Psychology [J] Xiamen University Press, 2019,446

[3] Wang Jing, Ba Anne, and Wu Bao-suo, goal, mechanism and architecture of intelligent teaching feedback [J].Research on modern distance education, 2023.

[4] Zhao Qin equal, virtual and real fusion network space security summary [J].Chinese Science: Information Science, 2024.54(04): 817-852.

[5] Zhang Ruilin, Jin Lijie, and Wang Zhiwen Li, the level evaluation and influencing factors of the integrated development of the Chinese skiing industry and tourism industry [J].Journal of Capital Institute of Physical Education, 2024. 36(01) .

[6] Yang Baojun, on the law of news [J] . Chinese People's University Press, 2024,597.

[7] Xu Bin, Ma Jun, Chen Qing, etc.. Study on Comprehensive Evaluation Index System and Investment Strategy of distribution network in Economic Development Zone based on improved AHP-TOPSIS Method [J] . Power system protection and control, 2019,47(22): 35 -44.

[8] Stanford University Report: Artificial Intelligence and Life in 2030[R].Stanford University, 2016.

[9] GE Ting, Chen Yanbo, Wang Yingxiang.Based on triangular fuzzy multi-attribute decision making. Investment benefit evaluation of distribution network [J].Grid and clean energy, 2018,34(10): 12-20.

[10] Zhu Yangbao. Multi-factor stock selection scheme based on XGBoost and LIGHTGBM algorithm [D]. Nanjing University, 2020.

[11] Python data analysis of data collation [J] . Cheung Heung Wai. Electronic world. 2021(11).