Eng Int Syst (2025) 4: 445–455

Engineering
Intelligent Systems

Mining of Potential Fields and Structure Optimization of ESI Discipline in Universities Based on Artificial Intelligence Algorithm

Fang Zhao[1] **and Yuan Tu**[2]*

[1]*Southwest Jiaotong University Hope College, Chengdu 610400, Sichuan, China*
[2]*Jiangxi University of Technolog, Nanchang 213001, Jiangxi, China*

This study examines the application of artificial intelligence (AI) algorithms in optimizing discipline layout and identifying potential fields in higher education. The objective is to enhance the development potential and resource allocation efficiency of essential science indicators (ESI) disciplines in universities through technical means. The paper explores the theoretical foundations and ESI discipline evaluation criteria based on artificial intelligence technology, providing an in-depth analysis of these technologies' specific applications in discipline layout optimization. By means of comprehensive data collection and preprocessing, this study establishes a framework for algorithm design and performance evaluation, enabling the precise identification and analysis of potential fields for discipline development. Furthermore, the practical utility of the algorithm in optimizing discipline layout strategies is examined and discussed, offering suggestions for algorithm improvement and future development directions. This work provides a scientific basis for optimal decision-making regarding courses layout in colleges and universities, supporting the optimal allocation of higher education resources and the effective formulation of discipline development strategies.

Keywords: Artificial Intelligence algorithm, ESI subject evaluation, Optimization of discipline layout, Potential field mining

1. INTRODUCTION

In today's global knowledge economy, higher education institutions are being offered opportunities as well as facing unprecedented challenges. With the continuous expansion and deepening of the field of scientific research, optimizing discipline layout has become a crucial strategy for colleges and universities to enhance their competitiveness, promote academic innovation, and cultivate talent. As an important tool for evaluating the performance and influence of academic research, ESI discipline evaluation provides quantitative indicators for discipline layout. It helps higher education institutions identify strengths and potential areas and optimize resource allocation. Concurrently, the rapid development of artificial intelligence algorithms offers a new technical means for processing large-scale academic data, mining deep academic insights, and optimizing discipline layout. By applying these advanced algorithms, we can more accurately identify discipline development trends and evaluate the interaction between disciplines and its impact on the overall academic ecosystem. Therefore, exploring the application of artificial intelligence algorithms in optimizing ESI discipline layout and potential field mining in colleges and universities will help promote the improvement of academic research and education quality, providing a new perspective and tool for higher education management and policy formulation.

As evident in the review of research literature on potential field mining and layout optimization of ESI disciplines

---

*Corresponding author e-mail: tuyuan163@hotmail.com

in colleges and universities based on artificial intelligence algorithms, many studies provide theoretical foundations and methodological guidance. Zhang et al. emphasized the role of discipline evaluation in promoting the development of disciplines in colleges and universities through ESI data evaluation [1]. Additionally, Zhang et al. demonstrated the method of analyzing the distribution characteristics of ESI disciplines in China, the United States, and the United Kingdom using sub-disciplines and text content [2]. These studies show that subdividing disciplines and in-depth content analysis are critical for understanding the potential of disciplines and their global distribution patterns.

Yuan et al. discussed the logic and risks inherent in the construction of double first-class universities, and pointed out the key factors to be considered in university development planning [3]. This provides an important strategic perspective for optimizing discipline layout. In regard to the application of AI, Haq et al. discussed the prospect of applying this technology to ECG analysis, highlighting the potential and application value of AI technology in specific fields [4]. Kim focused on the legal status, recent trends, and legal implications of AI algorithms, discussing AI bias in particular [5], which provides an important reference for the ethical consideration of applying AI algorithms in discipline evaluation and layout optimization.

Furthermore, Zheng studied the architecture and evaluation of high-performance computers dealing with AI applications [6], while Hwang et al. proposed a simulation and optimization framework ASimOV [7] for embedded AI accelerators. These studies illustrate the possibility and necessity of optimizing AI algorithms at technical and architectural levels. Xia and Cai et al. discussed the management system framework of ideological and political education in colleges and universities from the perspective of big data, respectively [8], and the correlation between higher education level and college students' public mental health driven by AI [9], showcasing the application prospect of AI in higher education management and public health analysis.

In summary, the current literature provides a solid theoretical and methodological foundation for this study, from the evaluation of disciplines in colleges and universities to the application of AI technology, and the discussion on the strategy of higher education management and discipline layout optimization. Future research will continue to investigate the application of AI in the exploration and layout optimization of potential academic areas in colleges and universities and address the challenges and ethical issues that may arise during the technology implementation process.

From a theoretical perspective, this study enriches the cross-disciplinary research on artificial intelligence and educational management, providing a new theoretical viewpoint for optimizing the discipline layout of higher education. By applying AI algorithms to discipline evaluation and resource allocation, the effectiveness of data-driven decision-making in higher education is demonstrated, further promoting the integration and innovation of educational technology and management science. Through the evaluation and optimization of algorithm performance, the research also offers improvement directions and methodological guidance for specific AI application scenarios.

In terms of practical application, this study provides a tool for higher education institutions to optimize discipline layout based on data and algorithms, assisting school management to make more accurate and forward-looking decisions according to scientific research trends and internal resources. Especially when there are limited resources, by effectively identifying and tapping into the potential fields of disciplines, the rational distribution and utilization of educational resources can be facilitated, enhancing the comprehensive strength and competitive advantage of disciplines.

The purpose of this study is to explore the application and effect of AI algorithms in the potential field mining and layout optimization of ESI disciplines in colleges and universities. The research covers various aspects from theoretical basis to practical application. First, the research will systematically give an overview of AI technology, especially the algorithms that are widely used for data analysis and pattern recognition, providing theoretical support for subsequent analysis of discipline potential fields. Then, the paper will discuss the evaluation standard of ESI disciplines, and how this standard is applied in global higher education to measure the research performance and influence of disciplines, providing a quantitative evaluation basis for discipline layout. Following this, the study will conduct data acquisition and preprocessing, detailing the data sources, acquisition technology, and preprocessing methods to ensure data accuracy and reliability, establishing the basis for the algorithm's effective operation. Then, the selection and design of the algorithm is undertaken; this includes the design criteria, selection rationale, specific implementation details, and optimization strategies. Additionally, the method used to evaluate the performance of the algorithm is explained, as this evaluation ensures that the algorithm can be effectively applied to mining ESI potential fields and optimizing discipline layout.

Finally, the study will discuss the specific application of the algorithm in optimizing discipline layout, analyzing how the algorithm influences the formation of optimization strategies, and the potential benefits of these strategies for discipline development and resource allocation in colleges and universities. Through this series of research activities, this study aims to provide a scientific and effective method for the optimization of discipline layout in higher education institutions, supporting the rational allocation of educational resources and the strategic planning of discipline development.

2. THEORETICAL BASIS

2.1 Overview of Artificial Intelligence Technology

Artificial intelligence (AI), a branch of computer science, is applied to create intelligent systems capable of performing complex tasks. These tasks typically require human intelligence, including learning, reasoning, problem-solving, perception, and natural language understanding. As technology has advanced, AI has evolved from simple pattern recognition and logical reasoning to sophisticated capabilities

Table 1 Technical characteristics of artificial intelligence.

Characteristic	Describe
Adaptive learning	Ability to learn from data and improve performance without explicit programming.
Automatic reasoning	Using logical reasoning to solve problems and simulate human decision-making process.
natural language processing	Understand, interpret and generate human language and realize human-computer interaction.
perception	Through image and speech recognition, we can deal with complex perceptual tasks.

Table 2 Main indicators of ESI evaluation.

Index	Description
Citation times	The total number of citations of papers by other studies reflects the academic influence.
Highly cited papers	Papers cited in the top 1% of the world indicate the quality and influence of research.
Research hot papers	Papers cited rapidly in recent years reflect the current research focus.

like handling complex decisions, deep learning, and natural language processing [10]. AI applications are extensive, ranging from intelligent assistants in daily life to academic research in higher education, and extending to complex tasks in medical, financial, and manufacturing fields, all of which rely heavily on AI technology [11, 12].

Among the core technologies of AI, machine learning and deep learning are pivotal in driving the field's advancement. Machine learning enables computer systems to enhance their performance through experience, while deep learning, a subset of machine learning, processes and analyzes large volumes of data, learning complex patterns and features by mimicking the neural network structure of the human brain. With the advancement of these technologies, AI can learn and adapt autonomously without explicit programming, thereby providing more accurate and efficient decision support across various scenarios [13]. Its main technical features are shown in Table 1.

As shown in Table 1, artificial intelligence technology is evolving rapidly, and its progress in theory and application provides new possibilities for solving complex problems that traditionally require high intelligence. Especially in data-intensive fields, such as the optimization of discipline layout in higher education, artificial intelligence technology provides a scientific basis for decision-making through accurate data analysis and pattern recognition, which suggests the wide application potential of artificial intelligence technology and marks its important role in promoting scientific research and the innovation of practical applications.

2.2 ESI discipline Evaluation Criteria

ESI is a vital tool for measuring the influence of academic achievements of global scientific research institutions and their researchers. By analyzing the frequency of scientific research paper citations, ESI can determine the influence and academic value of scientific research achievements on a global scale [14]. ESI covers 22 major disciplines comprising the natural and social sciences, providing quantitative indicators for comparing and evaluating academic research.

The core indicators of ESI discipline evaluation include the number of citations, the number of highly-cited papers, and the number of hot research papers, which collectively reflect the academic quality and influence of research results within a discipline. The number of citations indicates the degree of recognition that specific research results receive from the academic community, while highly-cited papers and research hotspot papers highlight academic achievements with significant influence and important research trends in the field [15]. Its main indicators are shown in Table 2.

ESI discipline evaluation standards provide scientific research institutions and scholars with tools to measure the influence of their own research results, and also provide a scientific basis for higher education institutions to formulate discipline development strategies and optimize discipline layout. Through the in-depth analysis of ESI data, colleges and universities can identify their own position in global scientific research, find potentially advantageous disciplines, and then formulate corresponding development strategies, optimize resource allocation, and enhance the overall competitiveness and influence of disciplines.

2.3 The Application of Artificial Intelligence in the Optimization of Discipline Layout

The application of artificial intelligence technology for the optimization of the discipline layout in colleges and universities signifies that higher education management has entered a new era of data-driven decision-making. Utilizing AI algorithms such as deep learning and data mining, universities can analyze extensive academic and educational data, thereby identifying potential areas for discipline development, optimizing resource allocation, and enhancing the quality of education and research [16]. This process relies on the

Table 3 Application of artificial intelligence in the optimization of discipline layout.

Application area	Technology	Function
Analysis of discipline potential	Deep learning, data mining	Identify development potential disciplines and provide strategic planning basis.
the distribution of resources	Optimization algorithm and prediction model	Realize the rational distribution and optimization of educational resources.
Trend forecast	Time series analysis, machine learning	Predict the development trend of disciplines and guide the adjustment of discipline layout

highly complex and adaptive capability of the algorithms and requires accurate processing and an in-depth understanding of educational data.

Artificial intelligence algorithms can facilitate the optimization of discipline layout through the following aspects:

(1) Analysis of Discipline Potential: Algorithms can analyze historical data and current trends to identify disciplines with development potential, providing a basis for strategic planning in universities.

(2) Optimization of Resource Allocation: By evaluating multidimensional data such as research outcomes, and the performance of teachers and students, algorithms help managers to allocate educational resources efficiently.

(3) Trend Prediction: AI technology can forecast the development trends and future demands of disciplines, guiding universities to make forward-looking adjustments to discipline layout.

The basic application of AI in discipline layout is shown in Table 3.

Through the above applications, artificial intelligence technology provides strong technical support for the optimization of discipline layout in colleges and universities. It is helpful for colleges and universities to strengthen the competitiveness and influence of disciplines, and it can also promote the effective use of educational resources and the continuous innovation of academic research. Therefore, the application of artificial intelligence for the optimization of discipline layout reflects the deep integration and application of emerging technologies in the field of higher education, and shows the development direction of educational management and decision-making science.

3. DATA ACQUISITION AND PRETREATMENT

3.1 Data Source and Acquisition Technology

3.1.1 Introduction of Data Sources

Choosing the right data source is crucial for accurate analysis and effective decision-making. This study primarily relies on the following data sources:

(1) Academic Databases: Databases such as Web of Science and Scopus provide extensive academic publications and citation data, covering research achievements across various disciplines and recording paper citations. These databases offer a crucial foundation for analyzing the influence and academic contribution of different fields.

(2) ESI Database: As a key tool for measuring the academic impact of research institutions and their researchers, the ESI database provides direct evaluation data, including highly cited papers and hot research papers. This information is instrumental in identifying the development potential of disciplines and optimizing their layout.

(3) Statistical Data of Higher Education: Released by education departments or relevant institutions, these statistics include information on resource allocation, faculty strength, and research investment. Such data help to understand the current state of discipline layout in universities and evaluate the rationality of resource allocation.

(4) Network and Social Media Data: Interactions between academic communities and researchers on social media platforms reflect the latest trends and research interests. Analyzing this data can help discover emerging disciplines and research hotspots.

3.1.2 Acquisition Technology and Methods

(1) API Calls: Many academic databases and social media platforms offer application programming interfaces (APIs), allowing researchers to obtain data directly from these platforms. By writing specific scripts or using these APIs, one can efficiently collect the necessary academic publications, citation data, and social media interactions.

(2) Web Crawlers: For data sources without APIs or with limited API access, web crawler technology can be employed. Web crawlers simulate a browser to access websites, automatically extracting useful information. When using web crawlers, it is essential to observe the site's usage agreement and data policies to avoid copyright infringement or other legal issues.

(3) Open Dataset Downloads: Some academic institutions, government departments, and international organizations provide open datasets for download on their websites. These datasets are usually preprocessed, ensuring high data quality and availability.

By applying these technical methods, researchers can obtain rich and accurate data from multiple channels and

Table 4 Partial statistics of data set after pretreatment.

Subject name	Total number of papers	Average citation times (after standardization)
physics	3752	0.88
chemistry	4190	0.92
biology	3127	0.97
computer science	2893	1.03

platforms, providing a solid foundation for subsequent data analysis and discipline layout optimization research.

3.2 Data Preprocessing

In this study, the collected data undergo meticulous preprocessing to ensure accuracy and efficacy in the analysis. The key steps in preprocessing include:

(1) Discipline and Subdivision Labeling: Initially, all collected data are labeled according to their ESI disciplines. For data that can be divided into specific subdivisions, more detailed labeling is conducted.

(2) Handling Missing Values: Given that datasets may contain missing values, the mean filling method is used to address any gaps in the data.

By carefully processing the data, the study ensures robust and reliable results, facilitating an effective optimization of the discipline layout. For numerical data, such as the citation times of papers, if a record is missing, it will be filled with the average value of other records in the same discipline, as shown in (1).

$$X_{\text{filled}} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{1}$$

here, X_{filled} is the filled value, X_i represents the number of non-missing citations in the same discipline, and n is the total number of non-missing citations in the discipline.

(3) Outlier processing: In order to eliminate the interference of outliers on data analysis, the method based on *IQR* (quartile distance) is adopted to identify and process outliers. Specifically, if a data point is outside the *IQR* calculated by the first quartile (Q1) and the third quartile (Q3), it is regarded as an abnormal value and excluded. *IQR* calculation formula is shown in (2).

$$IQR = Q3 - Q1 \tag{2}$$

The recognition range of outliers is defined as: $\mathit{Lower\ Bound} = Q1 - 1.5 \times IQR$ and $\mathit{Upper\ Bound} = Q3 + 1.5 \times IQR$.

(4) Data standardization: In order to process data of different orders of magnitude, the Z-score standardization method is adopted, as shown in (3).

$$Z = \frac{(X - \mu)}{\sigma} \tag{3}$$

where X represents the original data point, and μ and σ are the mean and standard deviation of that data set, respectively.

After the above pretreatment steps, graduate students become a cleaned and standardized data set for subsequent analysis. The following is an example of some data after processing, as shown in Table 4 below.

Through these preprocessing steps, the research ensures the quality and consistency of data and provides a solid foundation for the subsequent analysis of artificial intelligence algorithms.

4. ALGORITHM DEVELOPMENT AND PERFORMANCE EVALUATION

4.1 Algorithm design and selection

4.1.1 Design Objectives

The specific objectives of the algorithm design are:

(1) Potential Field Identification: The primary objective of the algorithm is to identify the ESI disciplines in colleges and universities that possess untapped research potential. This involves analyzing historical data of various disciplines to pinpoint those with rapid growth and increasing research interest, even if their current citation rates or publication numbers are not the highest.

(2) Suggestions for Layout Optimization: Based on the identification of potential areas, the second goal is to provide specific recommendations for optimizing discipline layout. This includes determining the disciplines that should receive more resources, the ones that could be considered for merging or resource reduction, and suggesting directions for the establishment of emerging disciplines in the future.

(3) Efficiency Analysis of Resource Allocation: The algorithm is designed to analyze and optimize the efficiency of resource allocation within colleges and universities. Quantitative analysis seeks to determine the relationship between resource allocation and the development potential of disciplines, proposing optimal resource allocation schemes to enhance the overall competitiveness of university disciplines.

By achieving these design goals, the algorithm developed in this study will offer a scientific basis and effective tool for optimizing discipline layout and identifying potential areas within colleges and universities. This will assist higher education institutions to make more accurate and efficient decisions in a complex and dynamic academic environment.

4.1.2 Algorithm Selection

The following algorithms are selected for development and application:

(1) Clustering algorithm (K-means): In view of the need to identify the fields of ESI disciplines with development potential in colleges and universities, clustering algorithm can effectively group disciplines according to their respective characteristics and identify potential fields of AI. The K-means algorithm performs clustering by minimizing intra-cluster distances while maximizing inter-cluster separation. Its objective function is defined as follows (4).

$$J = \sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)} - c_j\right\|^2 \tag{4}$$

where J is the objective function used to minimize the sum of squares of the distance from the point in each cluster to the cluster center c_j, k is the number of clusters, n is the number of points in each cluster, and $x_i^{(j)}$ is the i-th point in the JTH cluster.

(2) Decision tree algorithm: In order to propose specific ways of optimizing the discipline layout, the decision tree algorithm is selected to analyze the discipline data and identify key decision-making nodes. A decision tree is suitable for dealing with classification and regression tasks by constructing a model and assigning examples to predefined categories. The decision tree is constructed based on feature space division, with the aim of maximizing information gain (IG), as shown in (5).

$$IG(D_p, f) = I(D_p) - \sum_{j=1}^{m}\frac{N_j}{N_p}I(D_j) \tag{5}$$

where $IG(D_p, f)$is the information gain of the feature f on the parent setD_p, I is the measure of impurity, $N_p$is the number of instances in the parent set, $N_j$is the number of instances in the J-TH segmented subset, and m is the number of splits.

(3) Neural network algorithm: Considering that the algorithm needs to have the ability to process large-scale data sets and highly complex pattern recognition, a neural network algorithm, specifically a deep learning model, is selected as the analysis tool. The deep learning model efficiently extracts and recognizes the features of complex data through multi-layer nonlinear transformation. Its general form is shown in (6) below.

$$\begin{aligned} f(x) = {} & W_n\sigma(W_{n-1}\sigma\,(\ldots\sigma\,(W_1x + b_1)\ldots) \\ & + b_{n-1}) + b_n \end{aligned} \tag{6}$$

where $f(x)$ is the output of the deep learning model, W_i and bi are the weights and biases of the i-th layer, respectively, and σ is the activation function.

The above algorithms are chosen in order to use their respective advantages to meet the challenges of discipline layout optimization and potential field mining, and to ensure that the research can achieve the set goals accurately and efficiently.

4.2 Implementation Details and Optimization

4.2.1 Algorithm Implementation Process

(1) Implementation of clustering algorithm: Taking K-means as an example, the implementation process begins with selecting an appropriate number of clustering centers, then iteratively calculating the distance from each data point to the clustering center, and allocating data points to the nearest clustering center according to the principle of minimum distance. The iterative process continues until the change of cluster center is below the preset threshold or reaches the maximum number of iterations. This process can be expressed by the following iterative formula, as shown in (7).

$$C_i^{(t+1)} = \frac{1}{|S_i^{(t)}|}\sum_{x\in S_i^{(t)}} x \tag{7}$$

where $C_i^{(t+1)}$ is the center of cluster i after the $(t+1)$ iteration, $S_i^{(t)}$ is the set of data points belonging to cluster i in the t iteration, and x is the data point in the set.

(2) Decision tree construction: Next, each cluster is analyzed using a decision tree algorithm, and a decision tree is constructed to identify key decision nodes. The construction of a decision tree is based on information gain or other similar indicators, and the features that can most increase the classification accuracy are selected as the basis for node splitting.

(3) Neural network training: Finally, the neural network model is used to further analyze the data of various disciplines, especially the potential disciplines pointed out by clustering and decision tree analysis. The training process of a neural network includes calculating the loss by forward propagation, then adjusting the model parameters by back propagation algorithm, and finally obtaining the model that can be accurately predicted or classified.

The overall implementation flow of the algorithm is shown in Figure 1.

After completing the above implementation process, comprehensive testing and evaluation are carried out, and the algorithm is adjusted and optimized according to the test results to ensure its effectiveness and accuracy in practical application.

4.2.2 Algorithm Optimization Strategy

For the development and performance evaluation of the algorithm, this study adopted a series of algorithm optimization strategies to improve the accuracy and efficiency of the model.

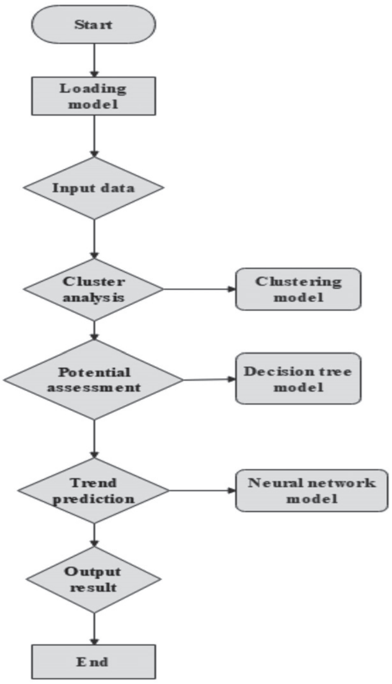

Figure 1 Overall implementation flow of the algorithm.

(1) Hyper-parameter tuning: For the neural network model, the selection of hyper-parameters has a great influence on the model performance. In this study, Grid Search strategy is used to optimize the key hyperparameters, including learning rate, batch size, number of network layers and number of neurons. By training and verifying a series of preset parameter combinations, the best parameter combination is selected fromthe verification set.

(2) Application of regularization technology: In order to prevent the over-fitting of the model, L2 regularization technology is introduced into the neural network. L2 regularization reduces the large value of the model weight by adding a regularization term to the loss function, as shown in (8).

$$L_{\text{new}} = L_{\text{original}} + \lambda \sum_{w} w^2 \tag{8}$$

where L_{new} is the new loss function after adding the regular term, L_{original} is the original loss function, λ is the regularization parameter, and w represents the weight of the model.

(3) Cross-validation: In order to evaluate the generalization ability of the model, this study uses the K-fold cross-validation method. The data set is divided into k subsets on average, one subset is taken as the test set and the other $K-1$ subsets as the training set one at a time, repeat this process for k times; finally, the average of k test results are taken as the performance index of the model. Part of the training process is shown in Figure 2.

To sum up, this study has significantly improved the performance of the algorithm by optimizing the system parameters, introducing regularization technology and applying cross-validation, thereby ensuring the effectiveness and accuracy of the algorithm in practical application. The implementation of these strategies provides strong technical support for the accurate mining and layout optimization of ESI potential fields in colleges and universities.

In the process of optimizing the ESI discipline layout in universities, artificial intelligence algorithms play a crucial role. By employing clustering algorithms like K-means, decision tree models, and neural network algorithms, the research effectively identifies potential disciplines with growth opportunities. The data sources include academic databases, ESI databases, higher education statistical data, and social media interactions, ensuring comprehensive coverage. Data preprocessing involves discipline labeling, handling missing values, outlier processing, and data standardization, resulting in a clean dataset for analysis. The algorithm is designed to identify potential fields, suggest layout optimization, and analyze the efficiency of resource allocation. Hyper-parameter tuning, regularization techniques, and cross-validation enhance the model's performance. The results indicate high accuracy in predicting potential disciplines, with notable performance in biology, physics, and chemistry. The application of the algorithm leads to resource redistribution, curriculum adjustments, and interdisciplinary cooperation, promoting data-driven decision-making in higher education management and supporting strategic planning and resource allocation.

4.3 Performance evaluation method

4.3.1 Evaluation Indicators

In order to comprehensively evaluate the performance of the algorithm developed in this study in the field of ESI potential

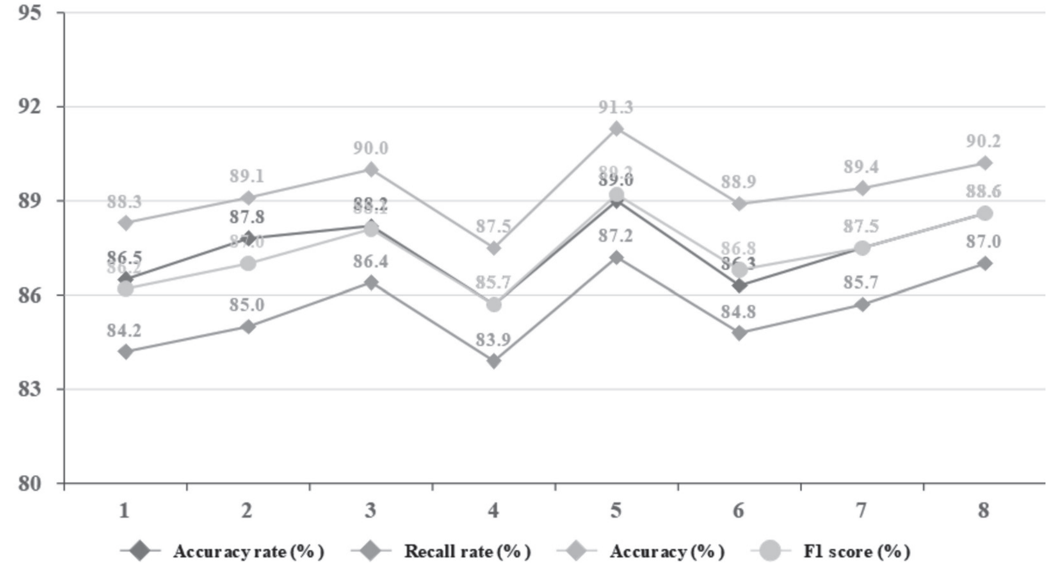

Figure 2 Part of the training process.

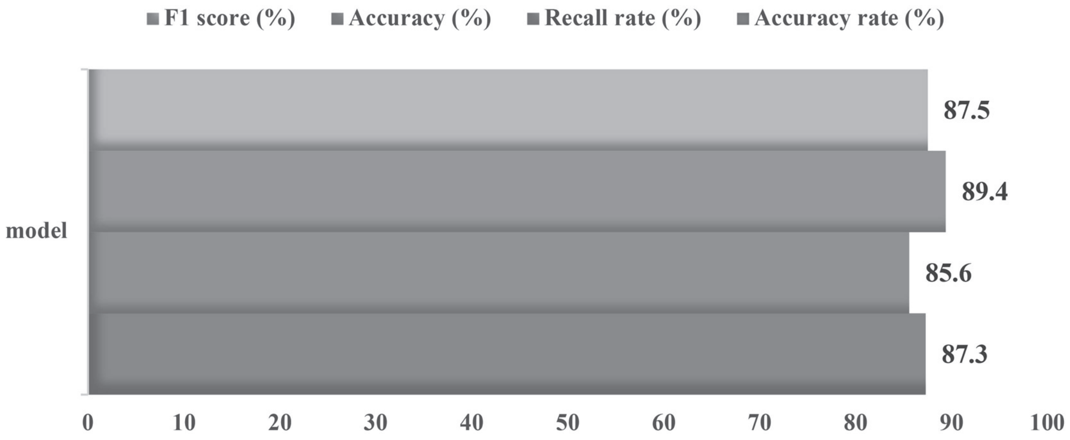

Figure 3 Performance of the algorithm across the whole discipline.

mining and layout optimization in colleges and universities, the following key indicators are adopted:

(1) Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

where, TP is a real case (correctly identified potential discipline), TN is a true negative case (correctly identified non-potential discipline), FP is a false positive case, and FN is a false negative case. In order to judge the accuracy of these classifications, the research refers to objective standards such as expert evaluation, historical data analysis, discipline ranking and evaluation indicators, and discipline funding and incentives. By comparing these objective standards, we can effectively verify the authenticity of the prediction results of the algorithm, thus ensuring the reliability and validity of the research results.

(2) Recall rate: The recall rate is the ratio of potential disciplines identified by the model to all potential disciplines, which is especially suitable for evaluating the model's ability to identify a few categories (potential disciplines), as shown in (10).

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

(3) Precision: The calculation formula is:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{11}$$

(4) $F1$ Score: the calculation formula is:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

By measuring the accuracy, recall, precision and $F1$ score of the model, the effectiveness and reliability of the model can be comprehensively evaluated.

4.3.2 Evaluation Results

As shown in Figure 3 and Figure 4 above, the overall model has a high-performance index, and the algorithm has shown even better performance in specific disciplines, such as biology, physics and chemistry. This shows that the algorithm can be effectively adapted and optimized according to the characteristics of different disciplines.

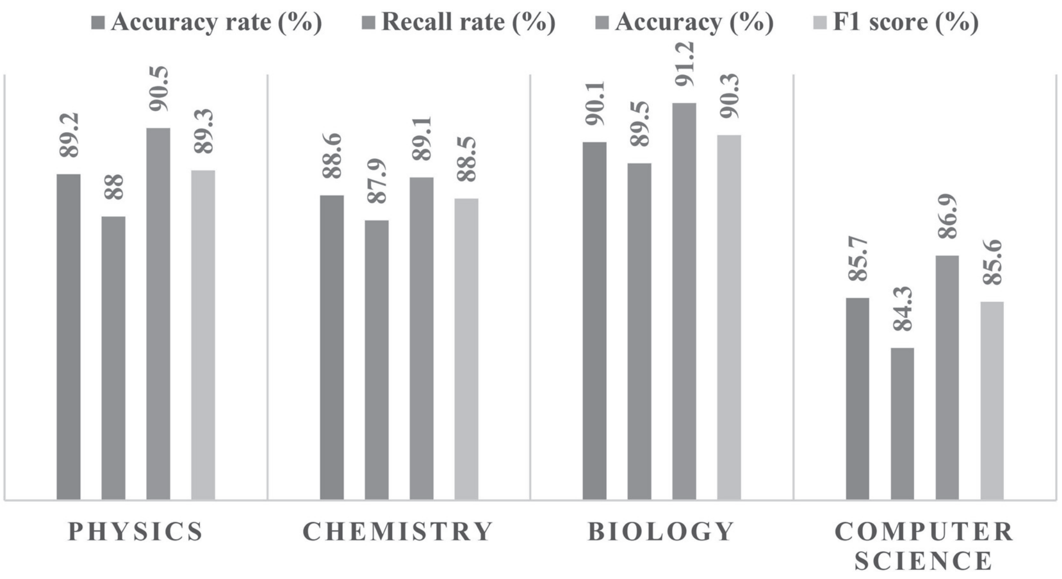

Figure 4 Performance of the algorithm in different disciplines.

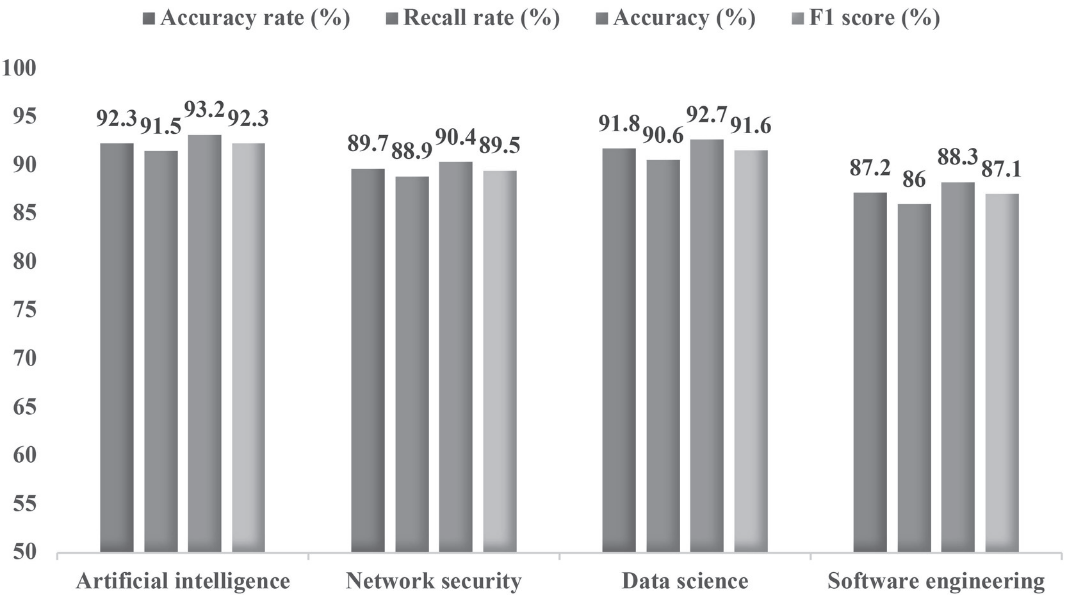

Figure 5 Performance of computer science department.

Furthermore, this study discusses in depth the application effect of the algorithm in one department of the large computer science faculty. The results are shown in the following Figure 5.

Performance examples of segmentation areas:

The evaluation results of the computer science subdivision presented in Figure 5 above indicate that the algorithm performs particularly well in the fields of artificial intelligence and data science. This may be attributed to the high consistency between the data characteristics of these fields and the training characteristics of the algorithm.

Through the above evaluation, this study shows that the developed algorithm shows good performance as a whole, and can also achieve a high degree of accurate prediction and classification in specific disciplines and subdivisions. This result proves the practical application value of the algorithm in the field of mining and optimizing the potential of ESI disciplines in colleges and universities, especially in providing data support for the discipline development strategy and resource allocation. Through in-depth analysis of a specific department, this study confirms the flexibility and applicability of the algorithm, which lays the foundation for future research directions and practical implementation.

5. ALGORITHM APPLICATION AND LAYOUT OPTIMIZATION STRATEGY

5.1 Application of Algorithm in the Optimization of Discipline Layout

(1) Identification of Potential Disciplines: First, using clustering algorithms and decision tree models, the existing discipline data of colleges and universities—such as citation counts, published papers, and the frequency of interdisciplinary cooperation—are analyzed to identify

disciplines or sub-sectors with growth potential. These identified potential disciplines will become the focus of resource allocation and development.

(2) Resource Redistribution: Based on the potential disciplines identified by the algorithm, resources are reallocated. This includes increasing investment in these disciplines, improving the staffing of researchers, and enhancing laboratory facilities to further develop these disciplines.

(3) Adjustment of Curriculum and Training Plans: According to the results of algorithm analysis, adjustments are made to the curriculum and postgraduate training plans of related disciplines. This involves introducing or strengthening courses and research topics related to potential disciplines, and attracting and cultivating talent in these fields.

(4) Promoting Interdisciplinary Cooperation: The algorithm analyzes potential opportunities for interdisciplinary cooperation, encouraging and promoting interdisciplinary research projects and teamwork to innovate research and application at the intersections of different fields.

(5) Performance Tracking and Feedback: After implementing the optimization strategies recommended by the algorithm, the development performance of the discipline—such as changes in discipline influence, the quality and quantity of research output, and the efficiency of resource allocation—is regularly tracked and evaluated. Based on the evaluation results, strategies are adjusted to form a closed-loop management system.

By implementing the above steps, the algorithm provided in this study can identify and recommend potential subject areas and assist universities to scientifically optimize the layout of disciplines and resource allocation. This process leverages data-driven decision support, making the optimization of discipline layout more objective, accurate, and efficient.

5.2 Algorithm Optimization and Future Development Directions

(1) Refinement of the Algorithm: Continuously improve the accuracy and generalization ability of the algorithm, particularly by introducing more dimensions of data analysis, such as indicators of teacher quality, student quality, and employment outcomes, to evaluate the potential of disciplines more comprehensively.

(2) Application of Deep Learning Models: Considering the potential of deep learning in handling complex data sets, future research will explore more deep learning-based models to uncover deeper relationships and patterns within discipline data.

(3) Interactive Decision Support System: Develop an interactive decision support system for discipline layout optimization, capable of reflecting the analysis results of the algorithm in real time and allowing users to adjust parameters or preferences according to the actual situation to obtain customized optimization suggestions.

(4) Dynamic Discipline Development Tracking: Implement dynamic tracking and prediction of discipline development trends, regularly updating data and models to capture the latest trends in discipline development and provide timely layout adjustment suggestions for universities.

(5) Multidisciplinary Integration Research: Encourage and support interdisciplinary integration research, using algorithms to analyze and identify potential opportunities for interdisciplinary cooperation, and promote resource sharing and knowledge exchange among disciplines.

By exploring the above optimization and development directions, this study aims to continuously improve the performance and application scope of the algorithm, better serving the scientific optimization and strategic decision-making of discipline layout in colleges and universities. These efforts are expected to promote the effective allocation of higher education resources and foster the development and innovation of disciplines.

6. CONCLUSION

In this study, an algorithm framework is developed and evaluated to identify potential ESI disciplines in universities and propose corresponding layout optimization strategies. Through an in-depth analysis of discipline data, including key indicators such as citation counts and the number of papers published, the algorithm successfully identified several subject areas with growth potential. Furthermore, this study puts forward specific optimization strategies based on the results of the algorithm analysis, such as the redistribution of discipline resources, adjustment of curriculum and training plans, and promotion of interdisciplinary cooperation. These strategies provide data support and decision-making references for the optimization of discipline layout in colleges and universities.

This study introduces advanced artificial intelligence technologies, particularly clustering algorithms, decision trees, and neural network models, to process and analyze large-scale discipline data. Through these methods, the research improves the accuracy and efficiency of identifying potential fields of disciplines and enhances the scientific and systematic decision-making process for the optimization of discipline layout.

However, this study also has several limitations. Firstly, although the algorithm shows good performance in many disciplines, its adaptability and accuracy for some specific disciplines still need improvement. Secondly, the performance of the algorithm depends heavily on the quality and completeness of the data, and challenges in data collection and preprocessing may affect the accuracy of the analysis results. Finally, the application scenario of this study mainly focuses on the field of ESI in colleges and universities, and its applicability to other types of academic or research institutions requires further exploration.

Future research directions will focus on several key areas. Firstly, the research will explore more advanced artificial intelligence technologies and models to further improve the performance and adaptability of the algorithm. Efforts will be made to collect and integrate more extensive and detailed data resources to enhance the model's comprehensiveness and prediction accuracy. Additionally, future research will consider applying the algorithm in different types of educational and research institutions to explore its feasibility and effectiveness in a wider range of scenarios. It is anticipated that these efforts will provide more scientific and effective tools and methods for the optimization of discipline layout in higher education, thereby supporting the development and innovation of disciplines.

REFERENCES

1. Zhang W, Wang X, Yu X. A research on evaluation of disciplines in universities based on ESI data. Science Research Management. 2020; 41(2): 60–72. doi:10.2995/JYELWS.2020.41.2.60.
2. Zhang W, Yao H, Wang X. Analyzing characteristics of ESI discipline distribution in China, U.S. and U.K. with sub-disciplines and text contents. Data Analysis and Knowledge Discovery. 2022; 6(9): 14–26. doi:10.3467/DAKD.2022.6.9.14.
3. Yuan X, Dai Q, Song G. An analysis of the inner logic and risks in planning construction of Double First-Class universities. Science Research Management. 2019; 40(11): 12–21. doi:10.2995/SYLDXJ.2019.40.11.12.
4. Haq KT, Howell SJ, Tereshchenko LG. Applying artificial intelligence to ECG analysis: promise of a better future. Circulation: Arrhythmia and Electrophysiology. 2020; 13(8): e009111. doi:10.1161/CIRCEP.120.009111.
5. Kim J. Artificial intelligence algorithm regulation status, recent trends, and legal implications - focusing on the issue of artificial intelligence bias. The Digital Ethics. 2020; 4(2): 27–42.
6. Zheng W. Architecture and evaluation of high-performance computers for processing artificial intelligence applications. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition). 2021; 33(2): 171–175.
7. Hwang DH, Han CY, Oh HW, Lee SE. ASimOV: A framework for simulation and optimization of an embedded AI accelerator. Micromachines. 2021; 12(7): 838. doi:10.3390/mi12070838.
8. Xia Y. Big data-based research on the management system framework of ideological and political education in colleges and universities. Journal of Intelligent & Fuzzy Systems. 2021; 40(2): 3777–3786. doi:10.3233/JIFS-189411.
9. Cai Y, Tang L. Correlation analysis between higher education level and college students' public mental health driven by AI. Computational Intelligence and Neuroscience. 2022; 2022: 4204500. doi:10.1155/2022/4204500.
10. Kosciejew MRH. Algorithms of education: How Datafication and artificial intelligence shape policy. Journal of Education Policy. 2024; 39(1): 166–168. doi:10.1080/02680939.2023.2194150.
11. Gao B. Research and implementation of intelligent evaluation system of teaching quality in universities based on artificial intelligence neural network model. Mathematical Problems in Engineering. 2022; 2022: 8224184. doi:10.1155/2022/8224184.
12. Ji M, Deng Q, Liu Y. Integration, extension, and darning: A review of MIT cambridge's centennial construction process for scientific research spaces. South Architecture. 2023; (1): 81–89.
13. Zeng D, Yu Y, Wen J, Wang Y. Characteristics and evolution of the trend of interdisciplinarity in China: An analysis based on web of science classifications. Journal of the China Society for Scientific and Technical Information. 2020; 39(8): 872–884.
14. Zhang C. Development trends of science discipline: Five-Dimension analysis method. Journal of the China Society for Scientific and Technical Information. 2020; 39(11): 1204–1213.
15. Wang Z, Du P. The logical connotation of discipline layout and Chinese practice. Science and Technology Review. 2021; 39(3): 123–129.
16. Srivastava M, Jain K. Tracing Evolution of management of technology discipline. IEEE Transactions on Engineering Management. 2023. doi:10.1109/TEM.2023.3260861.