World Scientific Research Journal Volume 10 Issue 11, 2024
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

Design and Application of Education Management Decision Support System Based on Big Data and Natural Language Processing

Hao Li[1], Fengzhen Wang[2], Rui Ye[3, 4], Jianfeng Li[5, 6, 7], Yongze Chen[8], Yunting Fan[8, *]

[1]Graduate School, University of the East, Manila1008, The Philippines

[2]Jiangxi New Energy Technology Institute, Xinyu338000, China

[3]Cavite State University, Indang 4100, The Philippines

[4]Hainan Jingerxin Information Technology Co., Ltd, Danzhou 571700, China

[5]XING YUN CHEN (Hongkong) Technology Co., Ltd, Hongkong 999077, China

[6]Hainan Cloud Spacetime Information Technology Co., Ltd, Sanya 572025, China

[7]College of Civil Engineering, Fuzhou University, Fuzhou350108, China

[8]Wesleyan College of Manila, Manila 0900, The Philippines

[*]Corresponding Author

Abstract

The rapid development of information technology promotes the transformation of education field to data-driven. In order to improve the quality and efficiency of educational administration, this paper proposes an educational administration decision support system based on big data and natural language processing by using advanced data analysis technology. Specifically, this paper first pays attention to the importance of data collection and preprocessing to ensure the quality and consistency of data. Then, the application of natural language processing technology in educational data analysis is discussed, including emotion analysis, semantic role analysis and text conversion. The system design uses distributed file system and database technology to ensure high data availability and high system performance. The experimental results show that the system has reached the high performance standard in response speed and stability, and the highest user satisfaction is 9.9, which shows that teachers highly recognize the system. After the system is put into operation, the employment rate of students and the coverage rate of curriculum resources have been significantly improved (99.8% and 97.3% respectively), which shows that the system has great potential to improve the allocation of educational resources and improve the utilization rate of resources.

Keywords

Big Data; Natural Language Processing; Decision Support Systems; Curriculum Resource Coverage.

1. INTRODUCTION

The integration of big data and artificial intelligence technology provides a new perspective and tool for education management. Nowadays, educational administrators can access a large amount of data about education, such as students' performance, course feedback and the use of

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

resources in teaching. However, how to use these data and develop them into valuable decision support has become a difficult problem in education management.

In order to improve the quality and efficiency of educational management decision-making, this paper designs and implements an educational management decision support system based on big data and natural language processing. Through these advanced data processing methods and natural language processing algorithms, the educational data are analyzed and reasoned, providing real decision-making suggestions for managers. This study not only pays attention to the theoretical application of big data and natural language processing technology in education management, but also verifies the effectiveness of the system in practice.

Firstly, this paper introduces the research background, problems and article structure. The second part explains the research status of decision support system and education management and the progress of natural language processing technology. The third section describes the methodology of the system design, including data collection and pre-processing, the application of natural language processing techniques, and the design of the decision support system. The Results and Discussion section shows the effect analysis of the system implementation, the evaluation of the decision support effect, and the problems and challenges encountered. Finally, the main contributions of the system and future research directions are summarised.

2. RELATED WORK

In today's rapidly evolving information technology era, the convergence of Artificial Intelligence and Decision support system (DSS) is revolutionising various industries. Zhou Zhimin elaborated on the application value of artificial intelligence technology in university logistics management. Secondly, based on artificial intelligence technology, he constructed a new decision support system platform for university logistics management, effectively improving the level of information construction in decision-making deployment, support, and other aspects of university logistics management, thereby achieving the goal of building a smart campus[1]. An Jianliang elaborated on the characteristics of data collection and digital technology, analyzed the application of data collection and digital technology in education management, and the application of decision support systems, including data collection methods and tools, system architecture and design principles, data analysis and decision support functions[2]. He Huan selected 66 patients who underwent knee arthroscopic meniscus surgery admitted to Lu'an Hospital affiliated with Anhui Medical University from December 2021 to April 2022 to explore the application effect of a knowledge-based clinical decision support system on the discharge readiness of patients undergoing knee arthroscopic meniscus surgery[3]. Zhao Yanhua aimed to improve the efficiency of medical staff in managing chronic diseases and accelerate the construction of smart healthcare. By continuously updating clinical data, iterating continuously, and optimizing plans, she had formed an intelligent comprehensive management support system for hypertension[4]. Based on the analysis of the current situation and business needs of shipyard planning management, Zhu Anqing designed the architecture of the shipyard planning management decision support system, constructed a shipyard planning management data warehouse, analyzed and summarized the themes of monthly plan assessment, monthly plan tracking, engineering health, and overall progress, and built a logical model of the data warehouse[5].

Gupta S introduced the synergy, differences, and overlap between artificial intelligence, DSS, and the operating room. In addition, the literature on methods based on developing DSS was clarified, and basic theories were proposed [6]. Rani P proposed a hybrid decision support system that can assist in early detection of heart disease based on the clinical parameters of patients [7]. Pagano A established a DSS for emergency management personnel to determine and prioritize the most appropriate emergency water supply measures [8]. Fox N reviewed the

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

most advanced gamified DSS for urban planning participation and climate education, promoting community participation in green space planning and providing information for design and management strategies [9]. Althubiti S proposed an innovative quantum black widow optimization algorithm and combined it with machine learning technology to develop a new type of decision support system [10]. Although existing research has achieved remarkable results, there is still room for expansion in the breadth and depth of data. Based on the limitations of the above research, by analyzing the big data in the field of education, this study uses natural language processing technology to design and implement an efficient, accurate and user-friendly educational management decision support system. The system will be able to process and analyze complex educational data and provide accurate decision-making suggestions, thus improving the quality and efficiency of educational management.

3. METHODS

3.1. Data Collection and Pre-processing

Data collection is the first step to build an educational management decision support system. The collection of students' performance data is extracted from the school's performance management system, including students' test scores, homework scores and classroom participation. The specific data are shown in Table 1:

Table 1. Student performance data

Student Unique Identifier	Major	Exam Score	Assignment Score	Classroom Participation	Attendance Rate	Average Score	Semester
SUID-001	Computer Science and Technology	85	90	High	98%	87.50	Spring 2024
SUID-002	Physics	78	82	Medium	95%	80.00	Spring 2024
SUID-003	Chemistry	92	88	Low	90%	90.00	Spring 2024
SUID-004	English Language and Literature	88	92	Medium	96%	90.00	Spring 2024
SUID-005	Biotechnology	76	80	High	92%	78.33	Spring 2024
SUID-006	History	82	85	Low	88%	83.50	Spring 2024
SUID-007	Geographic Information Science	90	95	High	97%	92.50	Spring 2024

Table 1 shows the specific information of students' learning situation, which is an important reference for educational management decision support system. These data in the table are helpful to analyze students' learning effects, adjust teaching strategies, optimize courses and improve the quality of education. Teacher evaluation data provides the same type of evaluation as student evaluation data from the perspective of students and teachers. Curriculum resource data focuses on the use of teaching materials and network resources, including the version of teaching materials, frequency of use and online resource access records. Course resource data are extracted from different online platforms, and the invalid and expired resource data are removed after data cleaning. Data cleaning aims to clean data, handle abnormal values and delete duplicate data by identifying and processing missing data to ensure accurate data for analysis.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

3.2. Natural Language Processing Technology

The application of natural language processing technology is closely related to educational management decision support system [11-12]. First of all, the system uses data mining technology to analyze educational data, such as student interaction and teacher feedback, and discovers potential patterns and trends in the data through mining:

$$P(y \mid X) = \frac{1}{1 + e^{-(WX+b)}} \tag{1}$$

Where $P(y \mid X)$ is the probability of target category y given input feature X. W is the weight matrix, b is the bias term, and e is the base of the natural logarithm.

The system uses sentiment analysis techniques to analyse educational documents. Sentiment analysis can be mathematically modelled by the following equation:

$$S = f(W_1 \cdot x_1 + W_2 \cdot x_2 + \ldots + W_n \cdot x_n) \tag{2}$$

S is the sentiment score of the text, W_n is the weight vector, x_n is the words or phrases in the text, and f is the activation function used to convert the linear combination into the predicted sentiment categories.

Speech recognition technology converts classroom recordings into text and this process is achieved through Hidden Markov Modelling:

$$P(O \mid H) = \frac{P(H \mid O)P(O)}{P(H)} \tag{3}$$

Where $P(O \mid H)$ is the probability of observing sequence O given state sequence H, $P(H \mid O)$ is the forward probability, $P(O)$ is the initial probability of observing sequence, and $P(H)$ is the initial probability of state sequence.

Natural language understanding techniques enable the system to understand the user's query intent, which involves semantic role labelling and can be achieved through conditional random field models [13-14]:

$$P(y_1, y_2, \ldots, y_n \mid x_1, x_2, \ldots, x_n) = \frac{1}{Z} \prod_{i=1}^{n} P(y_i \mid y_{i-1}, x_i) \tag{4}$$

Here, P is the joint probability of labelling sequence $y_1, y_2, \ldots, y_n$ given observation sequence $x_1, x_2, \ldots, x_n$, y_i is the predicted labels given context x_i, n is the total number of labels, i is the label index, and Z is a normalisation factor that ensures that the probability distributions sum to one.

Natural language generation techniques translate the results of the analysis into natural language descriptions through sequence-to-sequence modelling:

$$P(y_1, y_2, \ldots, y_m \mid x_1, x_2, \ldots, x_n) = \prod_{j=1}^{m} P(y_j \mid y_1, y_{j-1}, x_1, x_n) \tag{5}$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

Here, P is the joint probability of the labelled sequence $y_1, y_2, \ldots, y_m$ given the observation sequence $x_1, x_2, \ldots, x_n$, m is the length of the labelled sequence, n is the length of the observation sequence, and j is the index in the labelled sequence.

3.3. Decision Support System Design

The decision support system for educational management based on big data and natural language processing contains a multi-tier data architecture, with efficient data storage and indexing mechanisms in the data tier to support large-scale educational data storage and fast retrieval, and the use of distributed file systems and database technologies to ensure high availability and scalability of data [15-16]. The processing layer is responsible for cleaning, transforming and loading the data, as well as applying data mining algorithms to extract valuable information and patterns.

The model layer integrates sentiment analysis and text conversion for natural language processing to process and analyse educational text data to identify key topics and sentiment tendencies [17]. The service layer includes a set of RESTful API that enable front-end applications to call back-end services. These interfaces are designed to allow users to interact with the system to perform data queries, view visual analysis results and generate reports. The interface considers different users, and has personalized views, the ability to customize reports and the ability to respond (to adapt to different devices and different screen sizes). The functional modules of the system are: 1. data management module-user management of import/export and batch data; 2. analysis module-including advanced data analysis tools, which helps users to make forecast analysis or trend analysis; 3. report module–users in the module generate their own customized reports; 4. decision support module-provides decision trees and optimization algorithms to assist decision-making.

4. RESULTS AND DISCUSSION

4.1. Analysis of the Effectiveness of the System Implementation

In order to verify the effectiveness and practicability of the decision-making system constructed in this paper, a series of experiments are carried out to evaluate the performance of the system. These experiments are aimed at understanding the response speed, stability and user interaction experience of the system in dealing with educational management decisions, ensuring that the system can not only provide accurate data analysis results, but also meet the needs of users in practical applications, showing efficient and reliable decision support capabilities. The results of the response speed experiment are shown in Figure 1:

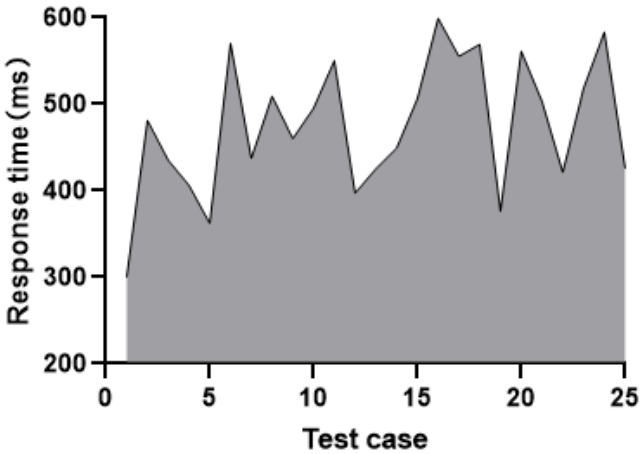

Figure 1. Response time

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

According to the data obtained from the experiments in Figure 1, the decision support system constructed in this paper has a good response speed, with a minimum response time of only 300ms, a result that shows that the system is able to quickly process user requests and return results. Comprehensively, the decision support system constructed in this paper achieves the standard of high performance in terms of response speed, which is of great significance for improving user experience and meeting the real-time requirements of educational management decisions.

Figure 2 shows the results of the system stability test:

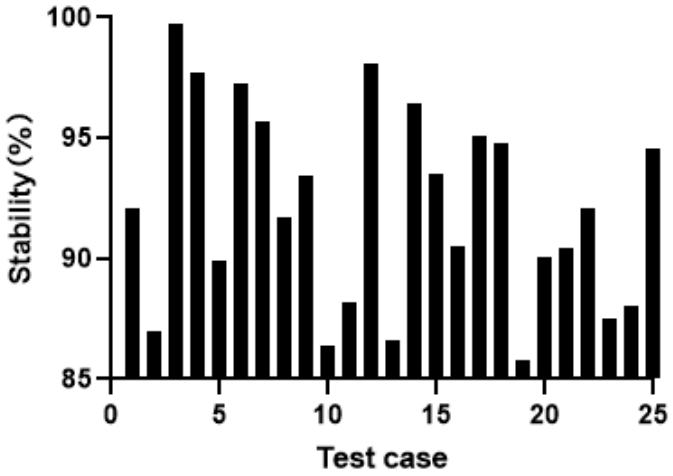

Figure 2. Stability

Analyzing the data in Figure 2, it can be found that the decision support system constructed in this paper has excellent stability, which is generally at a high level, up to 99.7%, and this remarkable stability index means that the system has demonstrated extremely high reliability during long periods of operation, and has been able to maintain virtually uninterrupted service, even in the face of sustained high loads.

In addition, in order to understand the evaluation of the system among the users, this paper collects the user satisfaction of 30 teachers about the system after using it as shown in Figure 3.

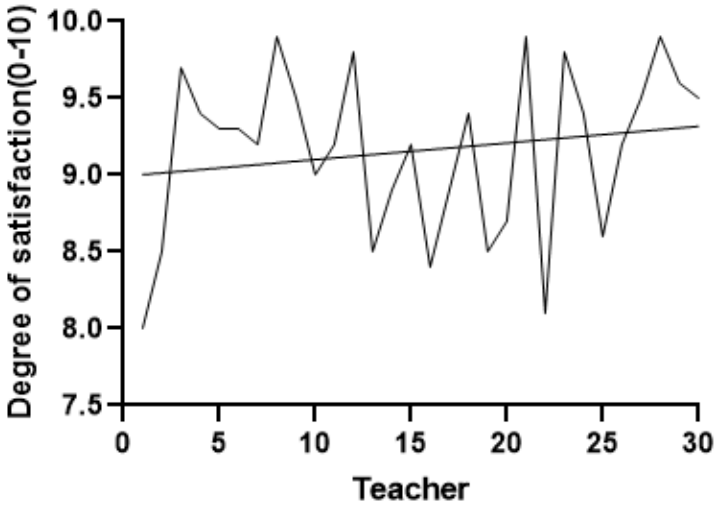

Figure 3. User satisfaction

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

According to the user evaluation data in Figure 3, all 30 teachers gave high satisfaction ratings to the system, with the highest being 9.9, which reflects that teachers generally believe that the decision support system plays a positive role in assisting teaching and educational management. Such high ratings not only indicate that the system's user interface is friendly and easy to use, but also show that the functions provided by the system are able to meet the practical needs of teachers in their daily teaching and student guidance.

4.2. Evaluation of Decision Support Effectiveness

In order to visually present the differences between before and after the implementation of the system, this paper compares the student employment rate and the coverage of course resources before and after the application of the system, in order to reveal the significant impact of the educational management decision support system on improving the quality of education and the effectiveness of decision-making, and to provide data support for the further optimisation of educational strategies and resource allocation. The comparison results are shown in Figures 4 and 5, respectively.

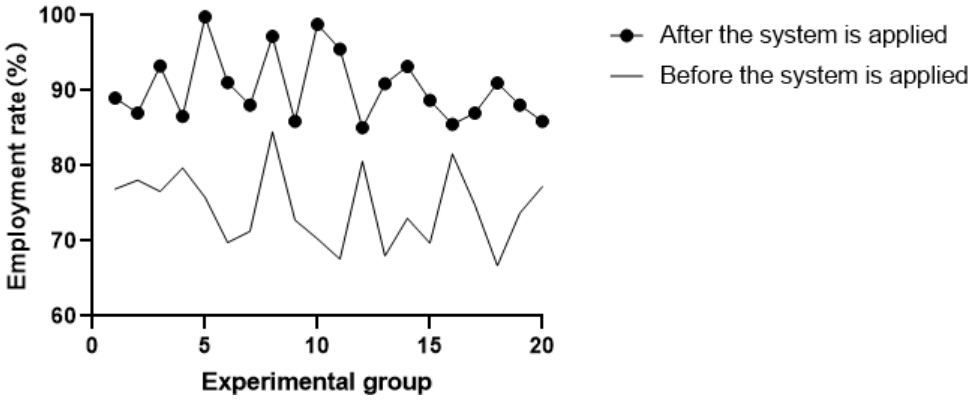

Figure 4. Employment rate

As shown in Figure 4, the employment rate before the implementation of the Education Management Decision Support System (EMDSS) was at a low level, with a maximum of 84.5%, but after the implementation of the system it reached a maximum of 99.8%, a significant increase that not only proves the great potential of the decision support system to improve educational outcomes, but also reflects the key role that the system has played in optimising career guidance and resource allocation.

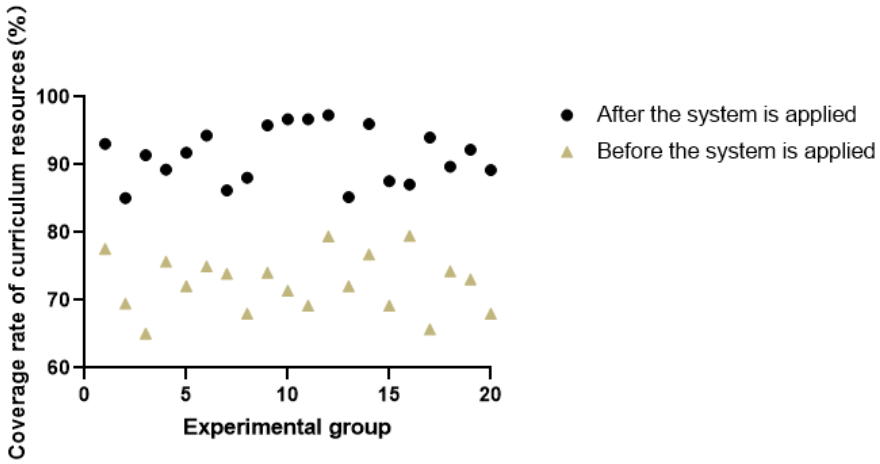

Figure 5. Coverage of curriculum resources

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

As shown in the data in Figure 5, the coverage of curriculum resources after the implementation of the system is significantly higher than that before the implementation of the system, ranging from 85.1% to 97.3%, which indicates that the decision support system plays a significant role in optimizing the allocation of educational resources and improving the utilization rate of resources. This enhancement reflects not only the system's ability to ensure that the wider student population has access to the necessary learning materials, but also the system's efficiency in accurately matching teaching and learning resources with student needs.

4.3. Problems and Challenges Encountered

In designing and implementing a decision support system for educational management based on big data and natural language processing, the challenge is to ensure the accuracy and completeness of the data, which is the basis for the system's analyses and decision-making. In addition, the integration of technology requires that the new system can properly interact with the existing educational management platform, and the natural language processing model needs to be optimized to enhance its ability to analyze educational texts. Real-time data processing and multilingual cultural adaptability are two other aspects that should be paid attention to in system design. The system needs to be able to respond to changes in the educational environment in real time, and to process data in multiple languages and cultural backgrounds to ensure that all users are treated fairly. In order to enhance the trust of users, the system needs to consider the interpretability of the algorithm when designing the algorithm. The rapid iteration of technology requires that the system can be continuously updated, and an effective evaluation and feedback mechanism needs to be established to continuously improve and optimize the performance of the system.

5. CONCLUSION

Through the above theoretical research and experimental verification, this paper realizes the educational management decision support system based on big data and natural language processing, and solves the challenge of data-driven decision-making in the field of education. The decision support system constructed in this paper uses data mining technology and natural language processing algorithm to analyze educational data, which provides accurate decision-making suggestions for educational managers and improves the quality and efficiency of educational management. Through the experimental verification, the system shows the characteristics of fast response and high stability, and the user satisfaction is high, which proves its effectiveness and reliability in practical application. In addition, after the implementation of the system, the employment rate of students and the coverage rate of curriculum resources have been significantly improved, which further confirms the potential of the system in optimizing the allocation of educational resources and improving the quality of education.

However, there are still some limitations in this study. Firstly, the system constructed in this paper currently focuses on specific aspects of educational management, and its ability to handle a wider range of educational scenarios and complex problems needs to be further expanded. Secondly, with the rapid development of technology and the changing needs of education, the system needs to be constantly updated and iterated to adapt to new challenges. In the future, the scope of the system application should be expanded to explore more application possibilities in educational scenarios; the adaptability and flexibility of the system algorithm should be strengthened to improve its performance in dealing with complex and changing educational problems.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202411_10(11).0003

REFERENCES

[1] Zhou Zhimin. Research on Logistics Management Decision Support System Based on Artificial Intelligence [J]. Wireless Internet Technology, 2023,20 (15): 45-47+51.

[2] An Jianliang. Analysis of Educational Decision Support System Based on Data Collection and Digital Technology [J]. Integrated Circuit Application, 2023,40 (8): 206-207.

[3] He Huan, Li Li, He Cuihuan, Fang Lingling, Ma Li, Liu Awei, Li Congyang. The application of clinical decision support systems in the preparation for discharge of patients undergoing knee arthroscopic meniscus surgery [J]. Anhui Medical Journal, 2022, 43 (9): 1092-1095.

[4] Zhao Yanhua, Zhang Jiachen, Ma Changxiao. Design and practice of a comprehensive management support system for hypertension based on national guidelines [J]. Chinese Journal of Health Information Management, 2022, 19 (3): 365-371.

[5] Zhu Anqing, Hao Dejing, Chen Ting. Research on Decision Support Systems for Shipyard Planning Management [J]. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2021,35 (6): 78-83.

[6] Gupta S, Modgil S, Bhattacharyya S, et al. Artificial intelligence for decision support systems in the field of operations research: review and future scope of research[J]. Annals of Operations Research, 2022, 308(1): 215-274.

[7] Rani P, Kumar R, Ahmed N M O S, et al. A decision support system for heart disease prediction based upon machine learning[J]. Journal of Reliable Intelligent Environments, 2021, 7(3): 263-275.

[8] Pagano A, Giordano R, Vurro M. A decision support system based on AHP for ranking strategies to manage emergencies on drinking water supply systems[J]. Water Resources Management, 2021, 35(2): 613-628.

[9] Fox N, Campbell-Arvai V, Lindquist M, et al. Gamifying decision support systems to promote inclusive and engaged urban resilience planning[J]. Urban Planning, 2022, 7(2): 239-252.

[10] Althubiti S, Escorcia-Gutierrez J, Gamarra M, et al. Improved metaheuristics with machine learning enabled medical decision support system[J]. Computers, Materials and Continua, 2022, 73(2): 2423-2439.

[11] Khurana D, Koli A, Khatter K, et al. Natural language processing: state of the art, current trends and challenges[J]. Multimedia tools and applications, 2023, 82(3): 3713-3744.

[12] Guellil I, Saâdane H, Azouaou F, et al. Arabic natural language processing: An overview[J]. Journal of King Saud University-Computer and Information Sciences, 2021, 33(5): 497-507.

[13] Tsujii J. Natural language processing and computational linguistics[J]. Computational Linguistics, 2021, 47(4): 707-727.

[14] Maulud D H, Zeebaree S R M, Jacksi K, et al. State of art for semantic analysis of natural language processing[J]. Qubahan academic journal, 2021, 1(2): 21-28.

[15] Min B, Ross H, Sulem E, et al. Recent advances in natural language processing via large pre-trained language models: A survey[J]. ACM Computing Surveys, 2023, 56(2): 1-40.

[16] Gu Y, Tinn R, Cheng H, et al. Domain-specific language model pretraining for biomedical natural language processing[J]. ACM Transactions on Computing for Healthcare (HEALTH), 2021, 3(1): 1-23.

[17] Zhang T, Schoene A M, Ji S, et al. Natural language processing applied to mental illness detection: a narrative review[J]. NPJ digital medicine, 2022, 5(1): 1-13.