ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

An Intelligent Visualization and Analysis Method for Employment Trends of College Graduates in Big Data Environment

Hongmei Wang[1] *and Xijie Sun*[2,*]

[1] School of Management, Suzhou Polytechnic University, Suzhou, Jiangsu, 215104, China
[2] Physical Education Department, Suzhou University of Science and Technology, Suzhou, Jiangsu, 215009, China

SUMMARY: *In the era of big data, the importance of college students' employment is self-evident, which is related to the development prospects of individuals, as well as the effective use of human resources and the level of high-quality economic and social development. In this paper, we design a visualization prediction system of college students' employment situation based on big data to address this issue. Firstly, we construct a comprehensive data collection model based on the on-campus student source data resource database, recruitment information released by network recruitment websites, enterprise talent database, and relevant data information in the statistical yearbooks of national and local governments. In addition, intelligent model construction on the basis of big data analysis can efficiently discover valuable information laws in the massive multi-source employment data and provide timely feedback to the relevant government departments, universities and college students themselves, thus helping to make correct judgments; with the help of data visualization tools, the information hidden in a large amount of data can be displayed in a more intuitive form, greatly improving the information conveying With the help of data visualization tools, the information hidden in a large amount of data can be presented in a more intuitive form, which greatly improves the communication of information and the cognitive understanding of users. The method adopted in this paper has been tested in practice, and it has improved the accuracy of prediction results and the speed of massive data analysis compared with the existing methods, which can provide effective tools and technical references for the analysis of the employment situation of college graduates.*

KEYWORDS: *big data; intelligent visualization; employment trend; college graduates; machine learning*

1 Introduction

The quality of education is a key variable in promoting economic growth, in which employment has been highly valued as an important direction in the restructuring of higher education [1]. Higher education should provide high-quality human resources for high-quality economic development and promote high-quality and full employment of college graduates. College graduates are valuable human resources of the country, and as the supply side of the labor market, the number of college graduates continues to be high, and the employment situation is becoming increasingly severe. From the perspective of the demand side of the overall labor market, the Ministry of Education, in collaboration with multiple national government

*1948@mail.usts.edu.cn
https://doi.org/10.65102/is2026111
Published: 30 April, 2026

departments, has taken the lead in establishing and improving the employment promotion policy system, and has introduced a number of policies to improve the employment volume and structure of college graduates, including coordinating the improvement of the scale and structure of further education, leveraging the role of higher education as a "storage tank", promoting recruitment in government agencies, public institutions, and state-owned enterprises, and expanding job supply, etc. [2-5]. However, with the transformation of economic development and the impact of the epidemic, industrial restructuring has led to an imbalance in the matching of supply and demand, the employment situation of college graduates has been changing and the employment pressure has been increasing [6-8]. The study of employment trends of college graduates is of great significance for understanding the current situation of graduate employment, adjusting graduate employment policies, enhancing graduate employment ability, and improving the quality of graduate employment.

The most used survey methods for college graduates' employment are questionnaires and offline research, but the sample size of such surveys is insufficient to cover the trend of the whole employment market, and the timeliness and dynamic trends are not presented enough to correctly respond to the current employment trend [9-11]. And in the context of big data, graduate-related employment data are multidimensional, extensive, heterogeneous and multimodal, which become the main basis for analyzing employment trends. For this reason, based on big data and artificial intelligence, literature [12] constructed an employment analysis scheme for college students by effectively processing large-scale text data through fuzzy hierarchical clustering and semantic similarity association feature extraction algorithms, which provides an accurate classification basis for employment guidance. Literature [13] used big data technology to analyze the employment rate of colleges and universities, and found that high employment rate universities generally possessed a student-centered vision and focused on the development of distinctive professional departments that meet the needs of society. Literature [14] adopts a data-driven approach based on the construction of a recommendation platform, using tools such as the Pearson correlation coefficient and the Lorenz curve to assess the degree of match between the supply of talent in colleges and universities and the demand of enterprises, and to provide analytical support for the specialty setting and accurate employment. Literature [15] constructed an employment service platform based on big data analysis, which collects, stores and integrates data related to college graduates and enterprises, and can provide employment services for students after relevant analysis. Literature [16] takes Shandong Province in China as an example, analyzes the employment trend of college students in the post epidemic era through questionnaire surveys and big data mining, and establishes a standardized response system and guidance model to provide data-driven solutions for college graduates' employment response.

In addition, literature [17] designed a multi-dimensional visualization and analysis platform for employment, obtaining recruitment data through crawler, analyzing it with probabilistic theme model, and using Django and visualization technology to intuitively display multi-dimensional information such as education, salary, and so on, which provides references for job seekers and enterprises. Literature [18] constructed an employment analysis visualization platform integrating school, market and student information based on the safe shell framework, which provides efficient query and intuitive display through front and back-end modules, optimizes the allocation of employment resources, and improves the efficiency of graduates' access to employment information. Visualization technology converts these complex data into images, charts and other forms, directly presenting the employment trends represented by these data, which can clearly reveal the employment analysis of a certain stage and a certain major.

In this paper, under the background of big data, a visual intelligent analysis method of employment trend of college graduates based on gray correlation analysis is proposed. Firstly, the influencing factors of employment trend including employment rate, professional matching rate and salary level are selected by using grey correlation analysis; secondly, the K-means clustering algorithm and visualization technology are used to realize the accurate modeling of large-scale employment data, which is conducive to the further popularization and application of the model. The above quantitative conclusions are exploratory results for revealing the operation law of the current employment market, which can be used as important reference information for education and teaching management in colleges and universities.

2 Research methodology

2.1 Data Acquisition and Processing

The diversified and dynamic characteristics of the information on graduates' destinations force us to build a network of diversified information sources, including the school system, social system, administrative system and market survey, and to develop different data attributes and their functional characteristics in practice. We notice that although the internal management information system has authoritative, reliable, continuous and stable data resources, such as basic information in the school registration management system, school grades in the academic affairs management system, employment information in the employment guidance service department, and even information on the growth history of graduates in the alumni association, etc., it is also characterized by slower dynamic updating, and limited access to data. The latest data on recruitment information, salary level, employers' tendency and industry trend, etc. obtained from the API of external recruitment websites or crawled from the Internet can supplement the time lag of the internal information of the company, and enhance the sensitivity of data analysis to a certain extent. Relevant labor market statistics and industry research reports published regularly by the National Bureau of Statistics or local governments provide us with policy background references for our analysis, while in-depth research reports from third-party professional research institutes provide useful reference information from the perspective of specific segments.

Combined with the specific research on the differences in the characteristics of various types of data sources, a data cleaning method system based on the combination of rules and machine learning is proposed. In terms of data cleaning, data cleaning is mainly carried out from the aspects of column correspondences in relational databases, data format standardization, and data code specification, etc., and a common data dictionary and metadata management system are established to complete the consistency checking and data integration of heterogeneous data. For some semi-structured data such as Extensible Markup Language (XML) files and JavaScript object representation format data, automatic parsing methods based on pattern recognition are used together with manual proofreading to ensure the correctness of the parsing; while for unstructured text data, relevant technical means of natural language processing are used in a large number of ways, for example, Chinese word segmentation, entity recognition, sentiment analysis, etc. to transform text that cannot be directly computed into text that can be computed. For unstructured text data, we use a lot of natural language processing related technology, such as Chinese word segmentation, entity recognition, sentiment analysis, etc., to transform the text that cannot be directly calculated into structured feature vectors that can be used for calculation.

In the data quality control section, specific indicators of completeness, consistency, accuracy, timeliness, and processing rules are formulated; for outliers, box-and-line diagram method, score test method, and isolated forest method are used to determine data anomalies, and different processing methods are carried out according to different severity levels of the data; and missing values are filled in by different types of missing cases using average value filling, regression filling, and multiple interpolation method. We use mean filling, regression filling and multiple interpolation to fill the missing values according to different types of missing cases. In particular, the similar samples matching filling method we constructed can utilize the whole data with the same feature values to fill the gap estimation of key business indicators.

The entity alignment and conflict resolution in the process of multi-source data fusion identifies the same entity records in different data sources through entity linking algorithm, and establishes a conflict resolution mechanism based on the credibility weight to guarantee the consistency of the fused data, and the whole processing process is equipped with a quality monitoring and traceability mechanism through the organic combination of automated test scripts and manual sampling to continuously monitor the data quality status and deal with various problems in time to lay a solid foundation of high-quality, intelligent analysis work. Intelligent analysis work has laid a solid foundation of high-quality data.

2.2 Intelligent visualization and analysis model design

In view of the characteristics of the employment data analysis of college graduates, based on the establishment of an intelligent visual analysis model, and considering that traditional analysis means can hardly meet the demand for effective mining of high-dimensional time-series data, a data analysis framework based on the combination of deep learning and visual analysis is proposed in Fig. 1. Among them, the data cleaning subsystem preprocesses the collected raw employment data, adopts the grey correlation method Gray correlation method is used for dimensionality reduction, and attribute reduction based on feature selection is used to mine and extract the indicator factors that have an important influence on the final evaluation of employment, and to provide high-quality data support for the subsequent clustering and visual parsing. The second step is the clustering step, which uses K-means clustering to cluster the corresponding analysis tasks. The front-end system of the final presentation layer is developed and designed by using the current Internet technology, which has the characteristics of interactivity and updatability, and can more conveniently discover and summarize the relevant characteristics and changes of college students' employment information through graphical methods.

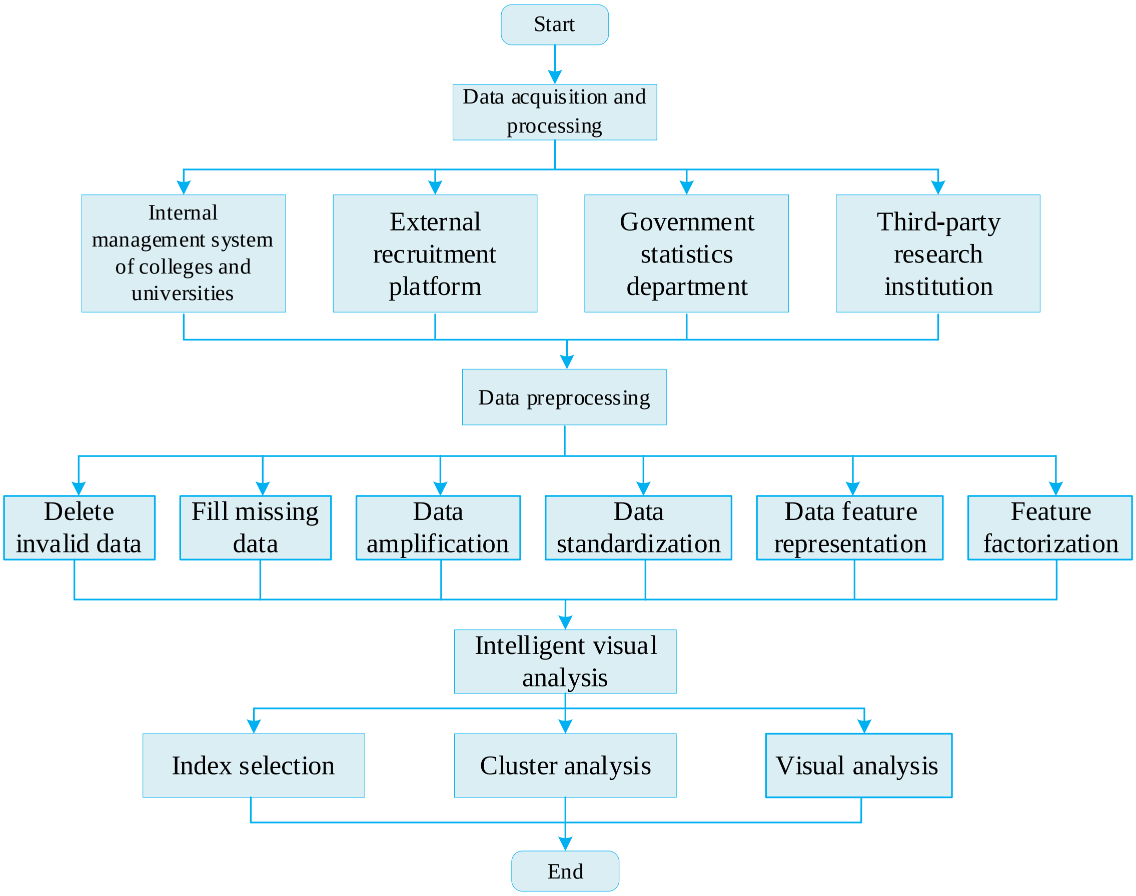

Figure 1: Comprehensive Analysis architecture

2.2.1 Selection of impact indicators for employment trends

There are many factors affecting the employment trend of college graduates, including relevant policies, laws and regulations, the level of economic development and the social and humanistic environment, as well as students' own comprehensive quality, their majors and other factors. In order to get more reasonable results of the employment trend of college graduates, it is necessary to integrate the relev ant indicators and use reasonable analytical means to deal with them. Gray scale correlation analysis is a common analysis method for the correlation between multiple indicators. By calculating the degree of correlation between the indicators, it can give the weight of the influence of each indicator on the trend, and use it to make predictions and analysis. Therefore, using the gray scale correlation method to make a reasonable prediction of the trend of employment of college graduates is a better choice. By measuring the gray correlation between the indicators affecting the employment trend of college graduates, we can judge the strength, size and order of the relationship between the indicators affecting the employment trend of college graduates. The basic process of the gray correlation analysis method is as follows:

(1) Collect data on the employment situation of college students, such as policy, economic, social, personal and other aspects of the relevant indicators of data; the collected indicators of data to expand and make explanations, the collection of different data to form a whole, and then transform the data into data that can be analyzed and visualized, so as to enable effective use of the data, to ensure that the data is reliable and valid.

(2) Cleaning, integrating, converting the expanded and interpreted data, removing the null values and outliers in the data, and carrying out data error correction and de-emphasis

processing to make the data comparable among the indicators.

(3) In data processing and analysis, in order to facilitate the comparison and analysis of data, it is necessary to select a reference series from the data of each indicator to form a matrix form with q data series, which is used as a standard of comparison for each other series. The matrix form is as follows:

$$C = \begin{bmatrix} c_1(1) & c_2(1) & \cdots & c_q(1) \\ c_1(2) & c_2(2) & \cdots & c_q(2) \\ \vdots & \vdots & \ddots & \vdots \\ c_1(\omega) & c_2(\omega) & \cdots & c_q(\omega) \end{bmatrix} \tag{1}$$

Where ω is the number of indicators related to the employment trend of college graduates.

(4) The data series to be analyzed by the above establishment includes a reference series $c_0(\omega)$ and a number of comparative series $c_i(\omega)$, $i = 1, 2, ..., q$. Several comparative series and a reference series are compared one by one, and the difference between them is calculated, and then the absolute value of the difference is calculated, which is expressed as:

$$z = \left| c_0(m) - c_i(m) \right| \tag{2}$$

$$i = 1, 2, ..., q;\ m = 1, 2, ..., \omega \tag{3}$$

Where z is the absolute difference; m is the degree of mutual information, which indicates the degree of association of the two variables; and i is the correlation coefficient. With the continuous increase of data volume and dimension, it is necessary to adjust the size of the values in the reference series at any time.

(5) Through the above steps, the maximum and minimum values of z are calculated and expressed as:

$$z_{\max} = \max_{i=1}^{q} \max_{m=1}^{\omega} \left| c_0(m) - c_i(m) \right| \tag{4}$$

$$z_{\min} = \min_{i=1}^{q} \min_{m=1}^{\omega} \left| c_0(m) - c_i(m) \right| \tag{5}$$

(6) On this basis, the correlation coefficients x; (m) are calculated for each of the impact indicators of the employment trend of the corresponding college graduates in the sub-series and the itemized series with the following formula:

$$x_i(m) = \frac{z_{\min} + v z_{\max}}{z + v z_{\max}}, m = 1, 2, ..., \omega \tag{6}$$

Where v is the discrimination coefficient.

(7) According to the correlation coefficient of each corresponding indicator of the employment trend of college graduates, the degree of correlation between the series as a whole is obtained, i.e., the average of the correlation coefficients of all the corresponding indicators of the employment trend of college graduates is calculated, and the formula is as follows:

$$b_{0i} = \frac{\sum_{i=1}^{\omega} x_i(m)}{\omega}, m = 1, 2, ..., \omega \tag{7}$$

Each college graduate employment trend impact indicator has a different degree of impact on college graduate employment trend, which can be regulated by the assignment method, expressed as:

$$b_{0i'} = \frac{\sum_{i=1}^{\omega} \varpi_m x_i(m)}{\omega}, m = 1, 2, ..., \omega \tag{8}$$

In the formula, ϖ is the weight of the influence indicators of the employment trend of graduates of each university.

(8) Sorting of correlation degree. According to the above steps, the degree of correlation can be calculated and sorted to obtain the influence indicators.

2.2.2 Cluster analysis session

Algorithm selection and improvement is a decision we made after fully considering the complexity of employment data. Based on the evaluation indexes of the constructed model, we used the cross-checking method and the grid search method to find the best parameter combinations of different algorithms. K-Means clustering is used to discover segments in the job market.The K-Means method is to minimize the sum of squared errors from each object to the point in the class to which it belongs, while the center of the cluster is determined by using the average of the data points in the cluster as the center of the cluster.

The main steps of K-means clustering are as follows:

(1) Determine the optimal number of clusters k value

In this paper, we use the contour coefficient method, the main parameter of which is the contour coefficient, and utilize the contour coefficient method to calculate the contour coefficient of each sample point in order to evaluate the clustering results, and then obtain the optimal number of clusters k. The contour coefficient value of a sample point X_i is defined as:

$$S = \frac{b - a}{\max(a, b)} \tag{9}$$

where a is the average distance between X_i and other samples in the same cluster, called cohesion; b is the average distance between X_i and all samples in the nearest cluster, called separation. And the nearest cluster is defined as:

$$C_j = \arg\min_{C_k} \frac{1}{n} \sum_{p \in C_k} |p - X_i|^2 \tag{10}$$

where p is a sample in a certain cluster C_k. In fact, it is the closest cluster to X_i that is chosen as the nearest cluster after using the average distance of all samples from X_i to a certain cluster as a measure of the distance from that point to that cluster. The average contour coefficient is obtained by averaging all the samples to obtain the average contour coefficient.

The average contour coefficient has a range of [-1, 1], and the closer the distance of the samples within the cluster and the farther the distance of the samples between the clusters, the larger the average contour coefficient and the better the clustering effect. Then, the k with the largest average profile coefficient is the optimal number of clusters.

(2) Clustering

1) Based on the determined k values, randomly select k research objects as the first set of clustering centers based on the data set consisting of n research objects;

2) Based on the determined k first group of clustering centers, calculate the distance between the remaining research objects and the k centers, and classify the corresponding research objects into categories based on the calculated value of the minimum distance to the nearest group of centers, so that the k categories are formed;

3) k a change in the intra-group mean value of the categories, through the correction of the intra-group mean value to re-determine the center position of each category;

4) The distance from the data points to the center of the clusters is calculated, and if the total cluster variance sum of squares is found to be decreasing, the center of the clusters has changed, and the data points need to be reassigned to the new clusters;

5) This process will continue until each category does not change, or until the number of iterations reaches a maximum, at which point the individual categories are the final clustering result. The formula is shown below:

$$E = \sum_{l=1}^{k} \sum_{i=1}^{n_k} (x_{il} - m_l) \tag{11}$$

where x_{il} represents a particular piece of data in C_l in the clustering, m_l represents the mean of the clustering C_l, and E is the sum of squared variances of the total distance.

2.2.3 Visualization and analysis session

The selection and setup of the visualization tool reflects our thinking process about the user experience and functionality, and after comparing the pros and cons of several visualization libraries, we chose a new visualization solution that is standardized on the web. Plotly Charting Engine has a powerful interactive capability and charting styles, and its built-in dynamic data-binding function allows the user to perform complex interactive operations such as chart zooming, moving, and filtering with the mouse. Users can use the mouse to zoom, move, filter and other complex interactive operations, and intuitively dig into the data.

D3.js provides a wealth of custom drawing modules, in the creation of complex charts such as mesh charts and Sankey diagrams have a unique advantage. ECharts, on the other hand, provides a complete Chinese documentation and optimization of mobile terminals. In terms of color selection, the color values are reasonably assigned according to the sensitivity of the human eye to color to ensure that all viewers can clearly read and understand the information expressed in the charts. The form of the charts is based on the data used and the results that want to be expressed, with scatter plots showing the correlation, heat plots showing the distribution of geographic locations, Sankey plots showing the direction of professional-industrial mobility, and line plots showing the change of the employment trend.

Interaction design focuses on ease of use, in the interface by dragging, clicking, hovering and other actions to complete the exploration of the data, multi-level drilling allows users to enter the micro view from the rough view, filtering and searching to facilitate the user to find the desired subset of the data, the export can be the results of the different file types for storage, so as to facilitate the post-production, reporting documents, and to assist in the leadership of the decision-making process.

3 Empirical research and analysis of results

3.1 Empirical studies

This part takes 12 colleges and universities of different levels and types located in the eastern, middle and western regions as the research object, mainly including comprehensive universities, polytechnics, teacher training universities and financial and professional schools, with the average annual number of graduates in each school ranging from 3,000 to 15,000, covering majors in 12 major disciplines such as science, engineering, literature, economics, management, education, art and so on. Relying on the long-term and close cooperation between the project team and the employment guidance department of each university, complete employment data information on students' personal basic information, academic performance, practice, job application and final employment status during the total of five years from 2019 to 2023 were obtained, with an effective data volume of more than 180,000, thus providing a large amount of data basis for the next simulation modeling. The employment data statistics of the graduates of the twelve universities are shown in Table 1, which shows that the average employment rate of the 173,200 samples counted reaches 90.7%, with the highest employment rate of graduates from polytechnic colleges reaching 94.5%. The average annual salary of graduates is 78,000 yuan, the rate of professional matching is 80.1%, and the employment area has a diversified distribution.

Table 1: Employment data statistics of college graduate

Type of institution	Sample size	Employment rate (%)	Average salary (ten thousand yuan)	Rate of major matching (%)	Regional flow direction
Comprehensive university	45,280	92.8	8.7	78.3	Mainly first-tier cities
Science and engineering colleges	38,650	94.5	9.2	85.6	First - and second-tier cities
Normal university	28,940	89.3	6.8	91.2	Uniform distribution among cities at all levels
Finance and economics colleges	32,180	91.7	8.9	82.4	Concentration in first-tier cities
Art colleges	15,720	87.2	7.1	73.8	A cultural industry cluster
Agricultural and forestry colleges	12,430	88.6	6.2	69.5	Small and medium-sized cities and rural areas
Total	173,200	90.7	7.8	80.1	Diversified distribution

This study uses stratified sampling to ensure that the proportion of students in each type of college, each type of discipline and region is basically the same, and then randomly assigns the collected data and divides it into two parts as experimental subjects. One part uses the visual analysis system proposed in this paper to predict and analyze the employment situation of graduates, while the other part is analyzed by traditional analysis methods, and the two parts use equal amounts of data. This paper compares the accuracy of the intelligent visual analytic model and traditional analytic methods in the analysis of employment trend of college graduates in five aspects, namely, employment rate prediction, salary level prediction, professional matching rate prediction, regional flow distribution and employment trend analysis, as shown in Figure 2. The results validate the systematic test of the model's effect by constructing a multi-dimensional evaluation system, and the experimental results show that the intelligent visualization analysis model has a prediction accuracy of 85.2% for the employment rate, which is 12.7 percentage points higher than the original algorithm (from 72.5% to 85.2%). The model also shows obvious performance advantages in the four core indicators of salary level, professional matching rate, regional flow distribution and employment trend analysis.

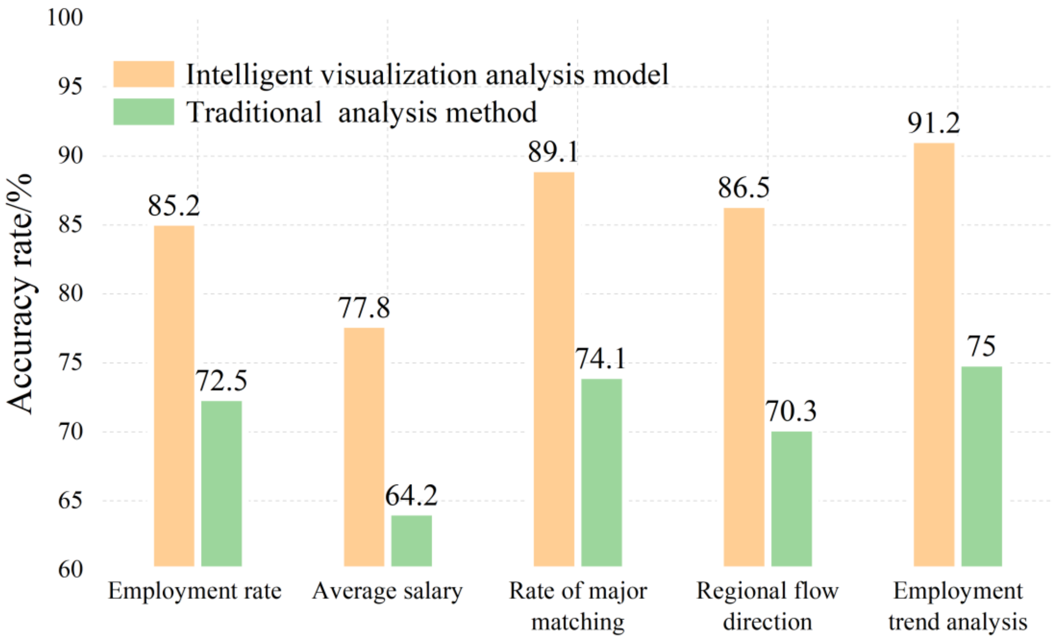

Figure 2: Comparison of accuracy rates between the two methods

In terms of timeliness, with the same order of magnitude of data volume, the time cost of the intelligent algorithm is about 76% lower than that of the traditional algorithm, which is of positive significance to the employment guidance departments of colleges and universities that are aware of employment trends in a timely manner. The user satisfaction test was conducted in the form of questionnaires and interviews, and the participating educational administrators and career guidance teachers found the system's visualization interface intuitive, friendly and easy to use, with an overall satisfaction rating of 4.6 (out of 5), and were satisfied with both the system's visual resolution and ease of use. The expert evaluation was conducted by relevant experts from the Student Employment Guidance Center of the Ministry of Education, the Institute of Higher Education and Human Resource Management to evaluate the research results, and the expert group fully affirmed that the model is technically innovative and practical and can be popularized, and that it has practical value and technical support in promoting the scientific and intelligent employment work of colleges and universities.

3.2 Results of data analysis

By intelligently visualizing and analyzing the employment data of twelve higher education institutions over a five-year period, the results of the analysis of the overall employment trend are obtained as shown in Table 2, and the employment trend of different types of colleges and universities from 2019 to 2023 is shown in Figure 3. After examining the employment situation of college graduates, it is not difficult to find that there is a multifaceted and dynamic process of change, which includes both the gradual increase in the employment rate of college students from 87.3% in 2019 to 92.1% in 2023, with an average employment rate of 89.2%. It also includes the situation in which there are obvious differences between different categories of colleges and universities, such as polytechnic colleges and universities are closer to the real needs of industrial development, and their employment rate basically stays at a higher level of 87% or more, and private colleges and universities of teachers' training, although it declined in 2020 due to the impact of the adjustment of the relevant policies in the field of education, it recovered rapidly in a short period of time and reached a historical high of 92.1% in 2023. The graduation implementation rate of finance and economics colleges and universities is greatly affected by the changes in the national economic environment but is generally on the rise, while the graduation implementation rate of arts and sports colleges and agriculture and forestry colleges and universities is relatively low due to the limitations of their majors but has a tendency to continue to increase.

Table 2: Employment trend analysis

Year	2019	2020	2021	2022	2023	Average value
Overall employment rate (%)	87.3	85.7	89.2	91.5	92.1	89.2
Average salary (ten thousand yuan)	6.8	7.1	7.6	8.2	8.9	7.7
Rate of major matching (%)	76.2	74.8	78.5	81.3	83.7	78.9
Entrepreneurship ratio (%)	2.1	2.3	2.7	3.2	3.8	2.8
Proportion of students advancing to higher education (%)	15.4	18.2	16.8	14.9	13.6	15.8
Flexible employment (%)	8.3	12.1	10.5	9.1	8.7	9.7

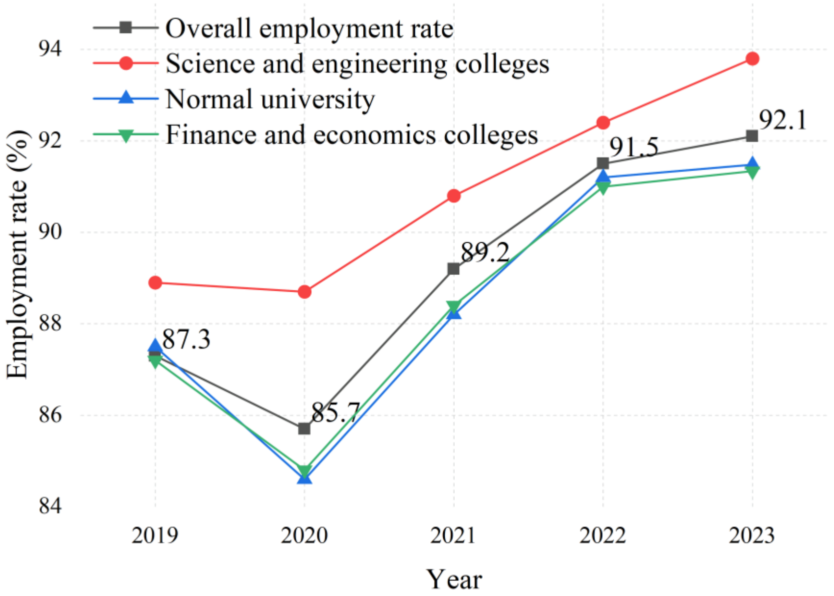

Figure 3: Employment trends of different types of universities from 2019 to 2023

At the same time, the specialty matching rate has increased from 76.2% in 2019 to 83.7% in 2023, further reflecting the improvement of the employment quality of college graduates. The overall counterpart rate of science and engineering majors is high, and the counterpart rate of old engineering majors such as mechanical engineering, electronic information engineering, civil engineering, etc. is over 90%, the counterpart rate of new cross-disciplines such as data science and big data technology, intelligent manufacturing engineering, etc. is over 85% due to the strong demand of the society despite the recent opening of these disciplines, and the counterpart rate of the majors of literature, history and philosophy is low, but the rate is steadily rising through the expansion of the scope of application and the increase of the practical link. Practical links show a steady upward trend, such as Chinese language and literature, history and other traditional arts and sciences disciplines, the major counterpart rate has increased by about 12% compared with that of five years ago.

The entrepreneurship rate increased from 2.1% in 2019 to 3.8% in 2023, which shows that the current college students' willingness to innovate and start businesses has gradually increased, and the types of college students' entrepreneurship are dominated by Internet technology, cultural and creative industries, educational services, and modern services, and the number of science and technology students' entrepreneurship accounted for more than 65% of the total number of entrepreneurs, and innovativeness has played a positive role in college students' entrepreneurial process.

The advancement rate returned to the norm after a sharp increase under the influence of the 2020 epidemic, reflecting the role played by China's postgraduate expansion policy on postgraduate training while graduates chose to go to graduate school to avoid pressure in the face of employment risks.

From the perspective of the inflow of college graduates, the traditional first-tier cities of Beijing, Shanghai, Guangzhou and Shenzhen are still the important inflow cities for college graduates, accounting for about 35.2%, but down 4.8 percentage points compared with five years ago, which shows that college graduates' employment destinations have become more diversified. Emerging first-tier cities, such as Hangzhou, Chengdu, Xi'an and Wuhan have become the choices of more people due to the many opportunities for development and the low cost of living, with the share of the total number of graduates rising from 28.6% to 33.4% in 2023. Secondly, the phenomenon of talent returning to the central and western regions brought about by the implementation of the national strategy of coordinated regional development and the strategy of industrial ladder transfer, especially in the areas of infrastructure, modernized agricultural development, and the development of cultural and tourism industries, provides a large number of opportunities suitable for the employment of local college graduates.

4 Conclusion

The intelligent visualization analysis method of college students' employment situation based on big data environment proposed in this paper has important application significance, and the data analysis of 180607 college students' employment information of 12 schools in the past five years has been carried out by using this method to prove that the method proposed in this paper is effective. From the results of the intelligent visual analysis model for employment rate prediction, salary level prediction, professional matching rate prediction, in terms of accuracy, in terms of prediction period, trend value, destination, regional mobility and other indicators, compared with the traditional statistical analysis model on average increased by more than 15%; computing speed increased by about three times. In the survey, it is found that the employment situation of college students has gradually improved, with the employment rate increasing from 87.3% to 92.1%, the employment matching degree gradually increasing to 83.7%, the average wage growth rate stabilizing at 7.2%, and the proportion of self-employment also increasing to 3.8%. The above data show the basic situation of college students' employment market, and provide a reference basis for school education and teaching management. The innovation of technical means is reflected in two aspects, this paper proposes an analysis framework based on deep learning and interactive visualization, combines the gray scale correlation analysis algorithm, K-means clustering algorithm and visualization tools to realize the effective modeling of the employment big data, and proposes a variety of visualization strategies for the analysis of the employment data, including scatterplot, heatmap, sanki chart, time series chart. The high degree of visualization largely improves the readability of the data as well as the reference value for decision-making, and the better human-computer interaction performance makes the model easy to promote its use.

The follow-up work includes considering the use of third-party data such as social networks and recruitment websites as supplements, constructing a visual human-computer interaction interface for data retrieval based on the NLP method, conducting joint analyses with other colleges and universities by adopting the federated learning method without disclosing student data, and designing an app to realize the push of relevant suggestions to college students and their parents at anytime and anywhere. The application and development of these technical means are used to promote the development of informationization of employment work in colleges and universities.

References

[1] Alpaydın, Y., & Kültür, K. (2022). Improving the transition from higher education to employment: A review of current policies. Education Policies in the 21st Century: Comparative Perspectives, 103-129.

[2] Hsieh, C. C. (2021). Improving the employment of college graduates in Taiwan: An analysis of policy design. Bulletin of Educational Research, 67(2), 41-79.

[3] Xiang, B., Wang, H., & Wang, H. (2023). Is There a Surplus of College Graduates in China? Exploring Strategies for Sustainable Employment of College Graduates. Sustainability, 15(21), 15540.

[4] Jin, R. (2021, January). Policy Insights: How Do Employment Policies Work to Benefit College Graduates in Zhejiang Province. In 2020 3rd International Seminar on Education Research and Social Science (ISERSS 2020) (pp. 496-500). Atlantis Press.

[5] Hu, H. (2022). RESEARCH ON THE EMPLOYMENT PRESSURE OF COLLEGE GRADUATES IN THE NEW ERA AND ITS COUNTERMEASURES. Psychiatria Danubina, 34(suppl 4), 398-398.

[6] Qian, T., Bian, J., & Chen, J. (2024). The status quo, causes, and countermeasures of employment difficulties faced by college graduates in China. Labor History, 65(4), 528-543.

[7] Wang, H., & Wang, C. (2024). Review of the impacts of COVID-19 pandemic on the employment of college graduates in China and countermeasures to it. Frontiers in Public Health, 12, 1390055.

[8] Green, F., & Henseke, G. (2021). Europe's evolving graduate labour markets: supply, demand, underemployment and pay. Journal for Labour Market Research, 55(1), 2.

[9] Shaokun, S., Xuwei, L., Wei, Z., & Wanting, Z. (2017). Investigation on employment quality and status of engineering graduates. INNOVATION AND MANAGEMENT, 50, 50.

[10] Nunley, J. M., Pugh, A., Romero, N., & Seals, R. A. (2017). The effects of unemployment and underemployment on employment opportunities: Results from a correspondence audit of the labor market for college graduates. Ilr review, 70(3), 642-669.

[11] Hunkerstorm, L., & Prescott, B. (2022). Where they work: Tracking the employment of college graduates by location. Change: The Magazine of Higher Learning, 54(5), 19-24.

[12] Yu, H., Zhang, R., & Kim, C. (2023). Intelligent analysis system of college students' employment and entrepreneurship situation: Big data and artificial intelligence-driven approach. Computers and Electrical Engineering, 110, 108823.

[13] Lim, H. W., & Kim, S. J. (2021). A study on ways to make employment improve through Big Data analysis of university information public. International Journal of Advanced Culture Technology, 9(3), 174-180.

[14] Li, L., & Tsai, S. B. (2022). An empirical study on the precise employment situation-oriented analysis of digital-driven talents with big data analysis. Mathematical Problems in Engineering, 2022(1), 8758898.

[15] Jiang, P., & Zhou, J. (2021, November). Application of big data analysis in college employment. In 2021 2nd International Conference on Information Science and Education (ICISE-IE) (pp. 797-800). IEEE.

[16] Fang, N. (2023, February). Research and Countermeasures on Changes in Employment Trends of College Students in Post-epidemic Era under Big Data Mining. In 2023 IEEE 6th Eurasian Conference on Educational Innovation (ECEI) (pp. 240-243). IEEE.

[17] Chen, Y., & Pan, R. (2022). Research on Data Analysis and Visualization of Recruitment Positions Based on Text Mining. Advances in Multimedia, 2022(1), 9047202.

[18] Li, Y., Zhang, J., Zhang, X., Luo, J., & Yin, J. (2025, June). Visualization Platform for Employment Analysis of College Graduates. In 2025 IEEE International Conference on Computation, Big-Data and Engineering (ICCBE) (pp. 990-993). IEEE.