ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

Research on Big Data Mining and Intelligent Reasoning Algorithms in College Students' Career Planning Education in Artificial Intelligence Environment

Kaiwen Deng[1] *and Fan Yu*[1,*]

[1] School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai, Guangdong, 519000, China

SUMMARY: *The rapid development of data-driven artificial intelligence promotes the precision and personalization of college students' career planning education. This paper analyzes career trends and job market demand through big data mining and machine learning technology, etc., to provide data support for college students' career planning. The career planning decision-making module is analyzed in terms of both goals and processes. Strong association rules are extracted from data such as graduates' employment situation to find the key factors affecting career planning and actual employment. Construct an accurate employment prediction model through feature engineering optimization and Stacking integrated model design to provide reference for final career planning. The method of this paper is applied to the association rule mining of actual college students' employment data, and the accuracy of the prediction results is compared to analyze the specific correlation between academic feature attributes and career planning and actual employment. The results show that mining yields three classes of comprehensive attribute classification rules. The AUC value of this paper's algorithm Stacking is 0.843, and the prediction results are better than the comparison algorithm. The most influential attribute on career planning and actual employment, "employment intention", is significantly correlated at the 0.01 level. The next most influential attributes were specialty, intention to work in the region, awards, and working hours.*

KEYWORDS: *career planning education; career decision-making module; association rules; feature engineering; stacking model*

1 Introduction

With the arrival of the fourth industrial revolution characterized by digitalization and intelligence, today's society is experiencing a paradigm shift from the "digital age" to the "digital age", and its impact has permeated every aspect of education, including higher education. College students' career planning education is an important part of higher education, and it is also a strategic high ground for colleges and universities to practice the mission of "educating people for the Party and educating talents for the country" [1]. With the arrival of the era of digital intelligence, the object characteristics, parenting goals, methods and paths of college students' career planning education have gradually changed [2]. Therefore, it is urgent to systematically sort out the opportunities and challenges faced by college students' career planning education and explore the countermeasures and suggestions for the reform of college students' career planning education in the new era [3, 4].

*yufan5313@jnu.edu.cn
https://doi.org/10.65102/is20261140 Published: 10 June, 2026

Artificial intelligence, as a product of the era of digital intelligence, has reshaped traditional career planning education in the process of application, and has brought new opportunities and challenges for career planning education in colleges and universities. On the one hand, colleges and universities should continue to use self-media and new media technology to make the employment guidance and career planning work of colleges and universities live up based on big data, promote the traditional advantages of college students' career planning education and information technology highly integrated to enhance the sense of the times, the sense of participation and attraction [5-8]. The use of big data mining technology summary to better personalized guidance for students' career planning education, so as to provide better preparation and realization opportunities for college students to have core competitiveness in employment [9-11]. On the other hand, colleges and universities can solve the problems in college students' career planning education with the help of intelligent algorithms to gain insight into the obstacles to students' career education [12, 13]. The logical reasoning theory is added to the ambiguous expression of college students' problems consulting, and the intelligent algorithm is used to analyze the real degree of the problem, which provides help in reasoning the problems of students [14, 15]. In conclusion, the application of artificial intelligence represented by big data mining and intelligent reasoning algorithms in career planning education plays a useful role in information integration, pattern optimization, etc., and provides important help for college students to improve their personal abilities and form employment competitiveness.

Using techniques such as data mining and intelligent reasoning algorithms to analyze the influencing factors of career planning and actual employment can help assist college students in developing a scientific career path. This paper focuses on the application of big data in college students' career planning and analyzes the importance of multiple types of data such as career data and job market forecasts. Combining the career decision module and intelligent reasoning algorithm, etc., it mines the association rules of academic characteristic attributes with career planning and actual employment, and gives the prediction results of students' employment unit intention. The accuracy of the prediction results is compared to prove the prediction validity of this paper's method. The importance and relevance of feature attributes are analyzed to explore the micro relationship between feature attributes and career planning and actual employment.

2 Big Data Mining and Intelligent Reasoning Algorithm Description

Based on big data mining and intelligent reasoning algorithms, this chapter constructs the theoretical framework and technical path of college students' career planning education, and provides scientific and systematic technical solutions for college students' career planning education.

2.1 Application of big data in college students' career planning

2.1.1 Analysis of career data

The core of career data analytics lies in the collection and analysis of large-scale data. Data sources include job advertisements, social media, industry reports, corporate surveys, etc. These data sources cover information from a wide range of industries, from salary levels to skill needs to career growth trends. The application of big data analytics methods is critical to career data analysis. Data scientists and analysts use techniques such as data mining, machine learning, and natural language processing to extract key information from massive amounts of data in order to identify trends in career development, such as which career fields have potential for growth in the future, and which skills will become more popular in the future. In practical terms, career data analytics can provide important insights for college students. Students can learn about the job outlook for their field of study, including salary ranges, employment opportunities, and industry trends, enabling them to better understand whether their career choices match market demand. For example, if data analysis shows that job opportunities in a particular field are growing rapidly, students may consider specializing in that field.

2.1.2 Job market projections

Job market forecasting relies on multiple data sources, including job advertisements, job boards, government data, and business surveys. These data sources provide information about future job demand, including industry, geographic location, skill requirements, and more. The key to predicting future job demand is building accurate predictive models. Big data technologies can be used to build machine learning algorithms that analyze past hiring data and predict future trends, and these models can take into account a variety of factors, such as economic conditions, technological advances, and societal trends. Job market forecasts are important for college students' career planning; by understanding future job demands, students can better choose a career path that suits their interests and skills. For example, if projections show that there will be a shortage of jobs in a particular field in the future, students may consider pursuing further education in that field.

2.2 Career Decision Making Module Objectives and Process Analysis

2.2.1 Objective analysis

The purpose of the Career Decision Making module is to obtain information about students before enrollment and analyze whether there are any strong correlation rules between the information and the information after enrollment. The purpose of data mining is firstly to clarify the purpose of data mining, and then to clarify how to transform the data mining in order to solve the related problems. On the basis of research with secondary school users and university workers, the needs of the career decision-making module are more clearly defined, such as:

1) The relationship between different employment categories and individual students;

2) The relationship between the type of advancement to a particular major and gender, and the relationship between employment status and place of origin;

3) The relationship between enrollment grades and college course grades;

4) The relationship between enrollment grades and graduation-related information.

The later extraction and organization of related information can provide reference for decision makers to ensure the accuracy of decision making.

2.2.2 Process analysis

1) The first step is to select the appropriate set of prescribed data items, and then the main sentence is organized so as to match the system requirements, and the data must be refined according to the demand.

2) Give the minimum support, because the amount of data is too large, the minimum support is generally in accordance with the requirements to make changes corresponding to it, in accordance with the minimum support given to obtain the frequent item set, that is, a large number of coexisting student information.

3) Minimum confidence is given. Based on this, the association rules in the frequent item set can be obtained.

2.3 Association rules

Association rule mining is currently one of the most researched methods in the field of research, which focuses on mining connections between things, and it was first proposed for analyzing the connections between product information in supermarket transaction databases.

2.3.1 Basic concepts of association rules

The first definition: if $I = \{i1, i2, \cdots, im\}$, belongs to a set of m items, each of which is different, ik is called that is, an item. The set of items I is also the itemset, where the number of elements is the length of the itemset, and the itemset whose length is k is called the $k-$ itemset.

The second definition: a transaction carried out in a transaction T belongs to a subset of the item set I, the relevant transactions will exist with the corresponding identification of the transaction number, which has the uniqueness, called TID. The transaction contains all the data that form the transaction database D, $|D|$ that is, the number of transactions contained in the D.

The third definition: in terms of the itemset X, if $count(X \subseteq T)$ belongs to the number of X transactions covered in the transaction set D, then the support of the itemset X is denoted as:

$$\sup port(X) = count(X \subseteq T)/|D| \tag{1}$$

The fourth definition: the item set minimum support threshold is actually the minimum support, SUPmin, which indicates the minimum importance of the association rules of the set where the user is located, and the item set with support exceeding SUPmin is a frequent set, and a frequent set of length k is called a k-frequent set.

Fifth definition: association rules are actually represented as:

$$R: X \Rightarrow Y \tag{2}$$

In this equation, $X \subset I$, $Y \subset I$, while $X \cap Y = \varnothing$. That is, the itemset X arises in a transaction, then Y occurs according to the corresponding frequency, and the association rules, which are valued by the user, are measured by support, confidence.

Sixth definition: the support of an association rule R actually belongs to the set of transactions and covers the ratio between the number of X, Y transactions and $|D|$. That is:

$$\sup port(X \Rightarrow Y) = count(X \cup Y)/|D| \tag{3}$$

The support can show the probability that X and Y are generated together, and the association rule support is the frequent set support.

Seventh definition: In the case of an association rule R, its confidence level covers the number of X, Y transactions and the ratio of the number of X transactions. That is:

$$confidence(X \Rightarrow Y) = \sup port(X \Rightarrow Y)/\sup port(X) \tag{4}$$

Confidence can show the probability that a transaction produces Y if X is covered in the transaction. Usually, the association rule between support and execution is the most critical for users.

Eighth definition: an association rule is strong if its minimum support is SUPmin and its minimum confidence is CONFmin. Rule R has greater support than SUPmin and its confidence is higher than CONFmin. The main purpose of association rule mining is to find strong association rules in order to help in decision making.

2.3.2 Nature of association rules

The more common forms in association rule mining include 3:

The first property: if X is a term set while $X \subseteq Y \subseteq I$ and $X \neq \varnothing$, then X support is higher than Y support:

$$\sup port(X) \geq \sup port(Y) \tag{5}$$

The second property: if X is a itemset while $X \subseteq Y \subseteq I$ and $X \neq \varnothing$, then the explicit data transaction set D, the minimum support minsup, is that if the itemset X belongs to the infrequent itemset, then Y is consistent with it. That is, there is no infrequent itemset superset, the same belongs to the infrequent itemset.

The third property: if X is a itemset, and at the same time $X \subseteq I$, $X \neq \varnothing$, $Y \subseteq X$, for the given then explicit data transaction set D, the minimum support $\min\sup$ that is: if the itemset X belongs to the frequent itemset, accordingly, Y also belongs to the frequent itemset. That is, any subset of frequent itemsets must be a frequent itemset.

By the first property and the second property, 2 corollaries can be obtained:

Corollary 1: If a K-dimensional itemset X is frequent, then all the $k-1$-dimensional subsets it contains must be frequent.

Corollary 2: If any k-dimensional subset of a K-dimensional itemset X is infrequent, then the K-dimensional itemset itself must not be frequent.

2.4 Feature engineering and modeling

2.4.1 Feature selection

In the feature engineering stage, features that have a significant impact on employment forecasting should be selected according to the actual situation. Feature selection can be carried out using methods such as statistical analysis, correlation analysis and the use of expert knowledge. In addition, feature selection algorithms, such as chi-square test, mutual information and L1 regularization, can be used to automatically select features with the most predictive power.

1) Cardinality validation

The chi-square test is used to assess the correlation between 2 categorical variables. In feature selection, it is able to use the chi-square test to calculate the correlation between each feature and the target variable as shown in Equation (6).

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} \tag{6}$$

where: f_0 is the observed frequency, reflecting the actual observed data distribution; f_e is the expected frequency, which is the predicted data distribution under the independence assumption.

By calculating the chi-square value, the degree of correlation between the feature and the target variable can be assessed, and the larger the chi-square value indicates the stronger the correlation.

2) Mutual Information

Mutual information is used to assess the degree of correlation and interdependence between 2 random variables. In feature selection, mutual information can be used to calculate the correlation between each feature and the target variable, as shown in equation (7).

$$I(x;y)=\sum\sum p(x,y)\cdot\log(p(x,y)/p(x)\cdot p(y)) \tag{7}$$

where: x and y are 2 random variables respectively; $p(x,y)$ is the probability of x and y occurring at the same time, and $p(x)$ and $p(y)$ are the probability of x and y occurring individually respectively. The degree of correlation between the feature and the target variable can be assessed by calculating the value of mutual information, the larger the value of mutual information, the stronger the correlation.

3) L1 regularization L_1 regularization is a technique used to reduce model complexity and feature selection. In feature selection, L_1 regularization can be applied to push the model to reset the weights of some features to 0. This method can realize the automatic selection of features as shown in Equation (8).

$$J(\theta)=Loss(\theta)+\lambda\cdot\sum|\theta| \tag{8}$$

where: $J(\theta)$ is the loss function after adding L_1 regularization; $Loss(\theta)$ is the original loss function; λ is the regularization parameter; and $|\theta|$ is the L_1 paradigm of the parameter vector. By adjusting the regularization parameter λ, the sparsity of the feature weights can be controlled to achieve feature selection.

2.4.2 Feature scaling

Since the range of values of different features may vary greatly, scaling of these features is usually required to ensure the stability of the model and to improve its convergence speed. Common feature scaling methods include normalization (making the distribution of features to have mean 0 and variance 1), normalization (scaling features to $[0,1]$ and so on.

Normalization is shown in equation (9).

$$x'=\frac{x-\bar{x}}{\sigma} \tag{9}$$

where: x is the original eigenvalue; $\bar{x}$ is the mean of the eigenvalue; σ is the standard deviation of the eigenvalue.

The normalization is shown in equation (10).

$$x'=\frac{x-\min(x)}{\max(x)-\min(x)} \tag{10}$$

where: x is the original eigenvalue; $\min(x)$ is the minimum value of the eigenvalue; $\max(x)$ is the maximum value of the eigenvalue.

2.4.3 Feature Combination and Interaction

In addition to the individual features themselves, combinations and interactions between features may also have an important impact on employment projections. For example, feature engineering methods can be utilized to construct new features, including proportions, differences, polynomial features, and so on. In addition, considering the correlation between different features, feature crossover can be performed, such as calculating the product, sum, and difference of features.

2.4.4 Model selection and design

In graduate employment forecasting, commonly used models include logistic regression, decision tree, random forest, support vector machine and neural network. The choice of model should be weighed according to factors such as the characteristics of the data, the complexity of the problem and the computational resources. Also, integrated learning methods, such as Bagging, Boosting, and Stacking, can be utilized to combine multiple models as a way to improve the prediction capability.

Figure 1 shows the computational approach of Stacking used in this paper.Stacking achieves integration by constructing multi-layer models, where each layer of the model uses the prediction results of the previous layer of the model as input features and is trained to predict the final result.

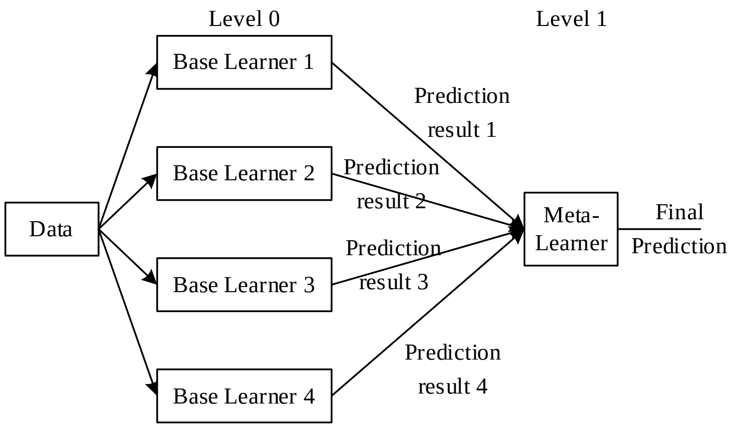

Figure 1: Stacking calculation method

3 Association rule mining and predictive application analysis

This chapter explores the potential association of academic characteristic attributes with college students' career planning and actual employment through association rule mining and multidimensional analysis of historical data. Combined with the algorithm of this paper, employment prediction is carried out and the accuracy of the prediction results of the same type of algorithm is compared. Further feature attribute importance assessment is conducted to reveal the specific relationship between feature attributes and college students' career planning and actual employment.

3.1 Association rule mining and result analysis

In this section multidimensional association rule mining technique is utilized for data mining of graduated student data. If we comprehensively consider the management data of graduated students, it includes various attributes such as outstanding students' grades in various subjects, awards, whether they are student cadres, whether they are members of the league, and so on. Moreover, the large number of graduated students over the years will inevitably result in a large amount of data to be processed. This paper aims to test the use of association rules to mine such data. In the selection of data attribute fields are selected is the specialty, awards, employment intentions, intention to work in the region, work time 5 categories. Graduation data discrete processing before the use of association rules need to do discrete data, that is, the different attribute fields continue to classify, in order to reduce the workload of data processing, but also to facilitate the analysis.

Professional attribute discretization process. Graduating students' work specialization was divided into two segments: A1 [this specialization] and A2 [interspecialization]. Award discretization processing. Integrate the awards won by graduated students during their school years into three parts: B1 [county level and below], B2 [provincial and municipal level], B3 [national level and above]. Discrete treatment of work area. The graduating students' intention to work is categorized as C1 [local], C2 [province], C3 [nationwide], and C4 [abroad]. Employment discretization processing. The employment intention of graduated students is divided into four categories: D1 [further education], D2 [school assignment], D3 [self-employment], D4 [entrepreneurship]. Work time discretization. The time for vocational school students to participate in work is divided into three time periods each year: E1 [November] (the end of the academic level examination), E2 [April] (the end of the transfer examination), E3 [June] (the end of the semester).

Table 1 shows the data samples (partially) obtained after discretization of the graduation data. From the data samples in Table 1, it can be seen that: among the majors, the graduated students are mainly in their own majors; among the awards received, the county level and below are the main ones; among the regions of work intention, the local area is the main one; among the employment intentions, the self-employment is the main one; among the time periods of participating in the work, the month of November is the main one; and the number of people of each correlation rule ranges from 4 to 10.

Table 1: Sample set of graduation data (parts)

ID	Profession	Awards	Work area intention	Employment intention	Working time	Number of people
1	College major	County level and below	Locality	Independent employment	November	10
2	College major	County level and below	Locality	Independent employment	June	8
3	College major	County level and below	Locality	Enter a higher school	April	7
4	College major	County level and below	Locality	School allocation	November	6
5	College major	County level and below	Locality	Independent employment	November	5
6	Cross-college major	Province and city level	Locality	Independent employment	November	5
7	Cross-college major	Province and city level	Province	Independent employment	November	5
8	Cross-college major	Province and city level	Province	Independent employment	November	4
9	Cross-college major	Province and city level	Whole country	Enter a higher school	June	4
10	Cross-college major	County level and below	Province	Enter a higher school	April	4

In order to mine useful rules from the above data, the minimum support for frequent mining is set at 15% and the minimum confidence level is set at 75%. If the minimum support is set too high or too low, too few or too many association rules are formed, which is unfavorable for the analysis results. According to the minimum support of 15%, Table 2 shows the frequent itemsets obtained from the first mining.1 After the first mining, the number of frequent items counted between each attribute is 6, and the number of each item count varies from 8-13.

Table 2: Frequent item set 1 from the first mining

ID	Stats 1	Stats 2	Stats 3	Frequent count
1	County level and below	Cross-college major	Independent employment	13
2	County level and below	College major	Enter a higher school	12
3	County level and below	College major	Independent employment	9
4	College major	Locality	Enter a higher school	8
5	College major	Province	Independent employment	8
6	Province and city level	April	Independent employment	8

Table 3 shows the set of frequent items obtained from the 2nd mining.2 After the 2nd mining, the count of the number of frequent items among the attributes is 3 and the number of items varies from 13-28.

Table 3: Frequent item set 2 from the second mining

ID	Stats 1	Stats 2	Stats 3	Stats 4	Frequent count
1	County level and below	Locality	November	Independent employment	28
2	Cross-college major	County level and below	June	Enter a higher school	17
3	College major	County level and below	Locality	Independent employment	13

A lot of valuable information can be derived from data mining of graduated students. The classification rules of 3 types of comprehensive attributes of majors, awards, employment intention, work area intention, and work time are used to determine their influence on career planning. Table 4 shows the results derived from the analysis of the method of this paper. Analyzing Table 4, it can be found that students in Rule 1 who won awards at county level and below, preferred local employment and self-employment, and were employed after the end of the November academic level examination had a support level of 50%, a confidence level of 80%, and an implementation level of 95% for taking excellent state-owned enterprises as their ultimate goal of career planning. Similarly, the support for students in Rule 2 to take excellent colleges and universities as the final goal of their career planning is 57%, the confidence level is 85%, and the execution level is 97%. The support of students in Rule 3 to take excellent private companies as the final goal of career planning is 55%, the confidence level is 88%, and the implementation level is 93%. Therefore, as a career planning education and management department, it can provide targeted supportive guidance training for students' employment.

Table 4: Influence of comprehensive attributes on employment unit planning

Synthetic attribute	Occupational unit planning forecast	Support degree(%)	Confidence degree(%)	Execution degree(%)
1	Outstanding state-owned enterprise	50	80	95
2	Excellent school	57	85	97
3	Outstanding private enterprise	55	88	93

3.2 Application Analysis of Employment Forecasting Algorithms

3.2.1 Comparison of algorithmic classification prediction results

In order to judge the advantage of analytical prediction effect of this paper's algorithm, three same type of classification prediction algorithms are selected: integrated learning algorithm improved based on gradient boosting decision tree (XGB), gradient boosting regression algorithm (GBR), and Random Forest algorithm (RF), to compare the area under the surface situation (AUC) under the same dataset, and to analyze the accuracy of prediction effect. Figure 2 shows the AUC values of the four algorithms. As can be seen from Figure 2, the AUC value of this paper's algorithm Stacking is 0.843, which is higher than that of the comparison algorithm's 0.824, 0.829, and 0.834. It shows that the dependence of the classifiers in this paper's algorithm is not strong, which does not lead to the problem of constant accumulation of error, and the prediction results will not be affected by the subclassifiers with larger errors, so the prediction accuracy is higher compared to the comparison algorithm.

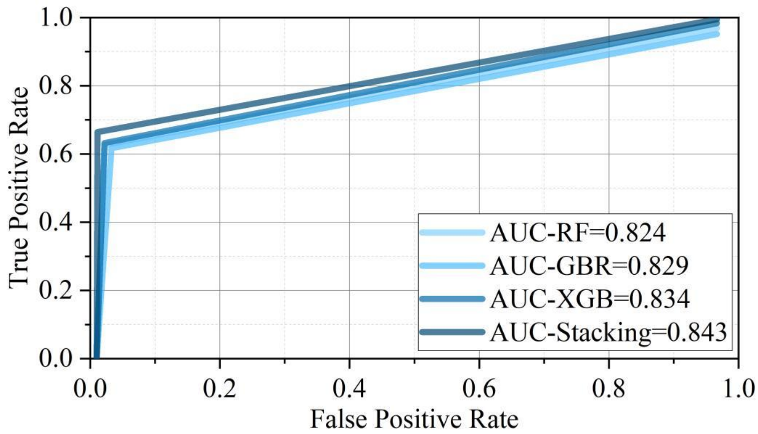

Figure 2: AUC values of the four algorithms

3.2.2 Importance Analysis of Characteristic Attributes

In addition to the accuracy of the final prediction of employment units, etc., in order to further analyze the key factors affecting the career planning and actual employment of college students, the filtered feature selection SelectKBest function is used, and the data processed in the previous section is imported into SelectKBest using sklearn.feature_selection to analyze the importance of relevant feature attributes. Figure 3 shows the results of feature attribute importance analysis. The horizontal coordinate indicates the name of the attribute column of the actual training data, and the vertical coordinate indicates the score value of the attribute, the higher the score value, the more important the attribute is. Analyzing Figure 3, the results of the score size ranking of feature attributes are as follows: employment intention (127), specialty (96), work area intention (95), awards (85), and work time (80). According to the score ranking results, it can be judged that employment intention has the greatest influence on college students' career planning and actual employment among the characteristic attributes. In the education of college students' career planning, relevant administrators should guide students to dig out their real employment ideas, and correlate employment intention with job market demand and so on, to explore the direction they want to go deeper in the future.

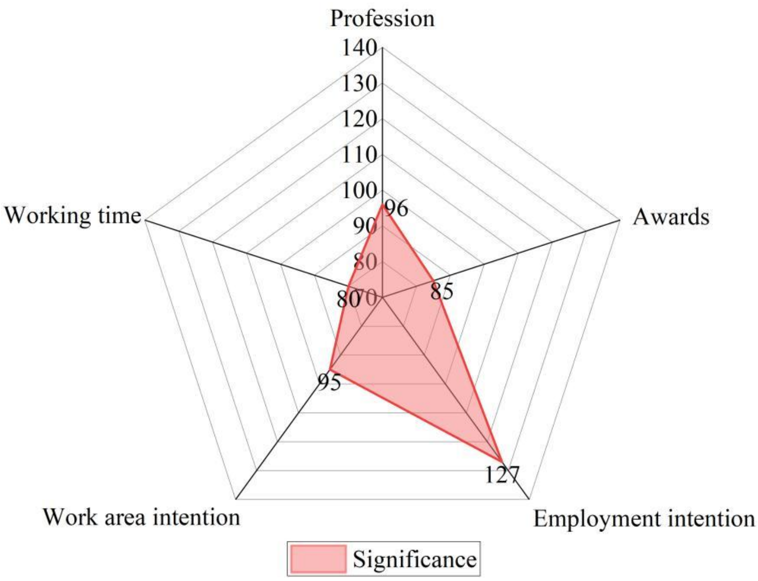

Figure 3: Feature importance analysis results

In order to further illustrate the specific relationship of each characteristic attribute with career planning and actual employment, correlation analysis was used for the study. Correlation analysis is a statistical method to study the direction and strength of correlation between variables. The linear relationship between two variables is quantitatively described by the correlation coefficient, which takes values between [-1, 1], and a correlation coefficient greater than 0 is a positive correlation, while less than 0 is a negative correlation. Pearson's correlation coefficient, Kendall's correlation coefficient and Spearman's correlation coefficient were used to describe the correlation in this study. Using SPSS Statistics software as a data analysis tool, we analyzed the correlation between five attributes, namely, employment intention, major, work area intention, award-winning situation, and working time, and “career planning and actual employment”, and calculated the correlation coefficient and significance between the two.

Table 5 shows the results of the correlation analysis. In Table 5, ** means 0.01 level, the correlation is significant. ** denotes 0.05 level, significant correlation. It is obvious from Table 5 that the three values of correlation coefficient and significance are different, but the same conclusion is obtained. Characteristic attributes with significant correlation at the 0.01 level are employment intention, P=0.001; and those with significant correlation at the 0.05 level are specialty, work area intention, awards, and working hours. So the key factors that are judged to influence students' career planning and actual employment are in order: employment intention, major, work area intention, awards, and working hours, which is consistent with the results of SelectKBest.

Table 5: Results of correlation analysis

Statistics	Pearson correlation coefficient		Kendall correlation coefficient		Spearman correlation coefficient	
Profession	Correlation coefficient	0.155*	Correlation coefficient	0.155*	Correlation coefficient	0.155*
	Significance	0.002	Significance	0.002	Significance	0.002
Awards	Correlation coefficient	0.147*	Correlation coefficient	0.253*	Correlation coefficient	0.267*
	Significance	0.005	Significance	0.005	Significance	0.005
Employment intention	Correlation coefficient	0.007**	Correlation coefficient	0.016**	Correlation coefficient	0.017**
	Significance	0.001	Significance	0.001	Significance	0.001
Work area intention	Correlation coefficient	0.021*	Correlation coefficient	0.008*	Correlation coefficient	0.009*
	Significance	0.004	Significance	0.004	Significance	0.004
Working time	Correlation coefficient	0.166*	Correlation coefficient	0.164*	Correlation coefficient	0.165*
	Significance	0.005	Significance	0.005	Significance	0.005

4 Conclusion

This paper constructs the analysis and prediction model of college students' career planning education through big data mining and intelligent reasoning algorithm. Three classes of comprehensive attribute classification rules are obtained through association rule mining of employment data of college graduates and so on. Each class of rules has different combinations of feature attributes such as major, award, employment intention, work area intention, work time, etc., and the final prediction results of employment unit intention are also different. Comparing the four different prediction algorithms, the AUC value of this paper's algorithm is 0.843, and the prediction error of each classifier is small and the prediction accuracy is high. In the importance analysis of feature attributes, the importance values of employment intention, specialty, work area intention, award, and work time reached 127, 96, 95, 85, and 80, respectively. Further analyzing the specific related situations, it is found that employment intention is significantly correlated with career planning and actual employment at the 0.01 level, and specialty, work area intention, award situation, and working time are correlated with career planning and actual employment at the 0.05 level.

In the future, real-time feedback data from enterprises can be further introduced into the model to improve the timeliness of the model, help students quickly understand the job market demand and other situations, and realize intelligent career planning education services.

References

[1] Serbes, M., & Albay, M. (2017). Importance of career planning and development in education. International Journal of Social Sciences & Educational Studies, 4(2), 149-154.

[2] Peng, H., & Lin, W. C. (2019). Evaluation a career planning course with case-based teaching model on college students’ career decisions and learning satisfaction. International Journal of Psychological Studies, 11(4), 102-114.

[3] Gu, X. (2022). Effect of deep learning on college students' career planning. Mathematical Problems in Engineering, 2022(1), 1573635.

[4] Mobley, C., Sharp, J. L., Hammond, C., Withington, C., & Stipanovic, N. (2017). The influence of career-focused education on student career planning and development: A comparison of CTE and non-CTE students. Career and Technical Education Research, 42(1), 57-75.

[5] Ning, X. (2021, December). Data mining analysis of career planning based on artificial intelligence technology. In 2021 International Conference on Aviation Safety and Information Technology (pp. 91-94).

[6] Arafath, M. Y., Saifuzzaman, M., Ahmed, S., & Hossain, S. A. (2018, September). Predicting career using data mining. In 2018 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 889-894). IEEE.

[7] Han, C. (2023). A study on career planning and development decisions of university students based on improved association rule algorithm. International Journal of Knowledge-Based Development, 13(2-4), 379-393.

[8] Mahboob, K., Asif, R., & Haider, N. G. (2024). Career planning matters: Intelligence-based career path predictions using data mining models-A longitudinal study. Mehran University Research Journal of Engineering and Technology, 43(4), 192-213.

[9] Kurmasheva, L., Kurmashev, I., Kulikov, V., Kulikova, V., & Tajigitov, A. (2024). The Use of Data Mining in the Management of the Career Guidance Work of the University. Annals of Data Science, 1-18.

[10] Zhang, Y. (2023). Application of data mining based on improved ant colony algorithm in college students' employment and entrepreneurship education. Soft Computing, 1-10.

[11] Guleria, P., & Sood, M. (2023). Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling. Education and Information Technologies, 28(1), 1081-1116.

[12] Huang, Y. (2025). Individualized career planning intelligent teaching system for college students based on adaptive artificial intelligence. Journal of Computational Methods in Sciences and Engineering, 14727978251315731.

[13] Shi, Z., Wu, Z., Zhang, Z., Chen, Y., & Liu, X. (2022). Learning path planning algorithm based on career goals and artificial intelligence. International Journal of Emerging Technologies in Learning (iJET), 17(10), 256-272.

[14] Tarasova, O. Y., & Doroshko, V. S. (2025). Methodological foundations of teaching the basics of artificial intelligence to lyceum students. In CEUR Workshop Proceedings (pp. 240-249).

[15] Ezenkwu, C. P., Johnson, E. H., & Jerome, O. B. (2015). Automated career guidance expert system using casebased reasoning. Computing, Information Systems, Development Informatics & Allied Research Journal.