Big Data Analysis and Modeling Method of College Student Employment Management Based on Deep Learning Model

Yibei Yin, Anyang Institute of Technology, China*

ABSTRACT

In order to study the big data of college students’ employment, this paper takes the big data of college students’ employment as the premise, analyzes the current employment data by establishing a DBN model, and puts forward relevant management measures, aiming to provide scientific basis for the management of graduates’ employment data. The results are as follows: By comparing the application evaluation of linear regression method, BP neural network and DBN model, this paper finds that DBN model has better accuracy and lower error and has better advantages in the application of college students’ employment data management characteristics. In addition, the development of social economy and the number of college graduates are the key factors for the employment rate of college students. Therefore, this paper suggests that through the use of big data technology, college will build a data platform for college students’ employment management and provide a carrier for college students to obtain professional information and employment management information.

KEYWORDS

Big Data, BP Neural Network, Deep Belief Networks, Deep Learning Model, Employment Management

INTRODUCTION

College students’ employability represents an important cornerstone to ensure the sustainable development of national education and is related to national economic development and social construction. Due to the expansion of university enrollment, the number of university students has continued to grow. In 2020, there will be as many as 8.74 million fresh graduates of domestic universities, while in 2021, the number will be 9.09 million—an increase of 350,000. In addition, affected by the COVID-19 pandemic, the demand for employment of enterprises has decreased, some posts are saturated, and the employment situation of college students is extremely pessimistic. At the same time, in order to improve their competitiveness, more and more college graduates have joined the ranks of the postgraduate entrance examination. Data shows that in just the ten years from 2009 to 2019, the number of graduate students admitted to colleges and universities has grown from 440,000

DOI: 10.4018/IJWLTT.330245

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.
to 720,000. On February 28, 2020, the Ministry of Education in China plans to increase the enrollment of master’s students by 189,000. The continuous increase of the number of graduates, the employment of graduates is facing a complex and severe situation, and the difficulty of graduate employment is also increasing, representing a complex and serious situation (Dong, 2021). The state attaches great importance to employment, especially for college graduates. In the new era, colleges and universities should combine their own reality and carry out more targeted employment guidance services, formulate corresponding employment opportunities according to the student condition promote the innovation of an employment work system for college graduates, and optimize employment guidance.

With the popularization of higher education, the employment environment and situation of college students have become complicated, which hinders the rational and scientific employment of college students in the future. In this context, colleges and universities should strengthen the employment management of college students, give full play to the role of the employment management of college students, strengthen the employment guidance of college students so that college students know the direction of future development, and promote the good employment of college students in the future. Based on this, in the practice of college education, it is necessary to focus not only on the development of college students’ employment management, but also on the innovation of college students’ employment management, so as to improve the quality of college students, thereby improving the quality of college personnel training and realizing the practical significance of college students’ employment management. Although college employment data management methods have achieved good development in recent decades, there are still shortcomings. Deep learning is a novel machine learning method proposed in the field of artificial intelligence in recent years. Deep learning can mine and capture the deep connections between big data by training big data and can improve the accuracy of classification and prediction (Cheng, 2022). It is an effective big data processing method. In addition, the training of deep learning models is faster, and with the increase of training samples, it can show better performance growth than general methods. The college student employment data management model based on deep learning can overcome the shortcomings of existing methods. The reasons are as follows:

- Sufficient data can ensure the accuracy of the model.
- The model can deeply mine the data relationship and establish an accurate proxy model between employment indicators and impact factors.
- The deep learning model can avoid the defects and uncertainties of the single employment indicator model to a certain extent (Gugnani & Singh, 2022).

In order to study the big data of college students’ employment, this paper takes the big data of college students’ employment as the premise, analyzes the current employment data by establishing a deep belief network (DBN) model, and proposes relevant management measures, aiming to provide a scientific basis for the management of graduates’ employment data.

LITERATURE REVIEW

Research on the Employment Data Management of College Students

The use of information and data in the employment management system of various colleges and universities is limited to ordinary general queries, and the utilization rate is very low. Most of the data are only stored on hard disk and have not exerted their value. The use of traditional working methods lacks the relevance, process, and integrity analysis of data information, which leads to the inability to objectively and accurately obtain dynamic employment information and analyze the changing information of students in the employment process, thus leading to the phenomena of “forced
employment” and “false employment.” In May 2019, the General Office of the Chinese Ministry of Education issued the “Notice on Further Strengthening the Statistical Verification of the Employment Status of College Graduates,” which pointed out that it is necessary to further improve the three-level verification mechanism of the employment statistics of graduates in local and national colleges and universities and check the employment of everyone individually (Chang et al., 2022). It can be seen that China’s requirements for employment statistics are not only limited to the employment rate, but also impose stricter requirements on the quality of employment and the authenticity of employment data. How should this problem be solved? With the emergence and development of various analysis models, combined with the reality that the number of college graduates continues to grow and the employment situation is extremely severe, scientific analysis model technology is applied to the employment system to create real-time analysis and prediction of employment information. Effectively guiding the employment of college students and relieving the pressure of college students’ employment has become a topic worthy of research at the current stage (Kong, 2020).

Many scholars have made various attempts to predict the employment of college graduates, and there are many effective methods for doing so (Kenny et al., 2021). Some scholars have proposed a method for predicting the employment of college graduates based on gray theory. This method regards the employment problem of college graduates as a gray problem and predicts their employment by fitting the relationship between input and output (Arboretti et al., 2009). It is simple and has high modeling efficiency, but because it simplifies the employment prediction problem of college graduates, the results are unreliable. Subsequently, a neural network-based employment prediction method for college graduates emerged, which was modeled by simulating the working principle of the human brain’s neural network (Bhagavan et al., 2021). However, the artificial neural network requires prior knowledge of the problem, and the defects are obvious. For example, it is too complicated and there is a large error in the employment prediction of college graduates. In recent years, the employment method of college graduates based on support vector machine has appeared (HakemZadeh et al., 2021). Although this method has made some progress in improving the effectiveness of employment prediction for college graduates, its modeling time is too long and extremely inefficient to meet the current requirements of employment demand for college graduates (Mitchell, 2020).

The Current Situation and Problems of Big Data Management of College Students’ Employment

The Employment Management of College Students Lacks Comprehensiveness

The employment management of college students in colleges and universities should be implemented into the whole process of college students’ development and give targeted and effective guidance to college students at every stage of their development and growth, so as to enhance their employability and promote their future employment (Wen et al., 2020). However, the current practice of college students’ employment management in China lacks comprehensive implementation work, which, in turn, neglects the importance of employment management in the process of college students’ development.

According to the practice survey, some colleges and universities still regard college students’ employment guidance as an important part of professional teaching (Li et al., 2020). However, this employment management model lacks the comprehensiveness of employment guidance for college students. At the same time, in the development of college students’ employment guidance, there is still the problem of attaching importance to theory and ignoring practice. The existence of this problem affects college students’ comprehensive cognition of employment guidance information, is not conducive to the improvement and cultivation of college students’ employability, and causes college students to be unemployed. In addition, there is still the problem of insufficient connection between college students’ employment management and professional education, which affects the development of college students’ professionalization (Wu et al., 2022).
**The Employment Management of College Students Is too Isolated**

At present, from the perspective of the current situation of college students’ employment management in China, there is a problem that college students’ employment management is too isolated. In the process of management, it is impossible to deeply examine and understand the advantages and disadvantages of college students’ employment, the advantages and disadvantages of college students’ employment, including what difficulties college students will encounter in employment, and how college students can solve the actual problems of their jobs, which is not conducive to the cultivation of college students’ employability. At the same time, the isolated problem of college students’ employment management method is reflected in the lack of informatization practice of college students’ employment management and the failure to lead college students to conduct market analysis and market research with advanced technology, which hinders the improvement of the quality of college students’ employment management (Warner et al., 2022).

**The Employment Management of College Students Lacks In-Depth Analysis, Application, and Tracking of Data**

Currently, although many colleges and universities carry out graduation surveys, they only focus on the statistics of graduate employment data, especially the employment rate statistics (Wu & Tsai, 2021). These data can only reflect part of the information, and these statistics have not been deeply analyzed and applied. It is impossible to guide colleges and universities to make important management decisions at the college level only by solving problems at the “operational level.” In addition, many colleges and universities in China do not pay attention to the data tracking of graduates after employment, and students cut off any contact with the school after graduation. In fact, the career development of many college graduates is not stable in the short term after graduation. The data survey and tracking of graduates will play an important role in guiding the discipline and specialty construction of colleges and universities and will also help colleges and universities master the needs of different sectors of society for talents in talent training. However, from the current point of view, many colleges and universities have relatively little data tracking of graduates after employment, and it is difficult to obtain medium and long-term career development data of graduates (He et al., 2020).

**METHODOLOGY**

Deep learning is a deep neural network, which originates from research on neural networks. Through its deep features, it can solve complex and high-dimensional data problems (Wang & Jiang, 2019). The currently recognized basic models of deep learning include deep belief network (DBN) based on restricted Boltzmann machine (RBM), stacked automatic based on AutoEncoder (AE) encoders (Stacked AutoEncoders, SAE), convolutional neural networks (CNN), and recurrent neural networks (RNN) (Xie et al., 2022). This paper will discuss the establishment and application of a DBN model in college student employment management.

**Model Introduction**

The deep belief network (DBN) is composed of a multi-layer restricted Boltzmann machine (RBM) network and a one-layer backpropagation (BP) network (Cortes et al., 2020). Each node of the model is subject to Bernoulli distribution. The bottom layer is the observation variable layer, and then the first layer and the second layer in order. Each layer represents the weight, as shown in Figure 1.

In the deep belief network, the feature output obtained by the previous layer of the RBM network is used as the input of the next layer, so that each layer can better abstract the features of the previous layer and extract data features layer by layer (Park et al., 2022). The top-level BP network takes the features extracted by the RBM network as input for classification or prediction. The input layer and hidden layer of the deep belief network can be expressed as formula (1) and formula (2), where $m$ and
$n$ respectively represent the number of input layer and hidden layer nodes. The restricted Boltzmann machine parameter vector is expressed as formula (3), where $w$ represents the weight matrix connecting the input layer and the hidden layer, and then formula (4) can be obtained:

\[ v = \{v_1, v_2, \ldots, v_m\} \]  

\[ h = \{h_1, h_2, \ldots, h_n\} \]  

\[ \theta = \{w, a, b\} \]  

\[ E\{v, h; \theta\} = -\sum_{i=1}^{m} a_i v_i - \sum_{j=1}^{n} b_j h_j - \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij} v_i h_j \]  

In formula (4), $a$ and $b$ represent the thresholds of the input layer and the hidden layer, respectively. According to formula (4), the joint probability distribution of the input layer and the hidden layer can be established, as follows:

\[ P\{v; \theta\} = \frac{1}{Z(\theta)} \sum_{h} e^{-E(v, h)} \]
\[ P\{h; \theta\} = \frac{1}{Z(\theta)} \sum_v e^{-E_v(v,h)} \]  

In the formulas, \(Z(\theta)\) is the normalization function, which is defined as follows:

\[ Z\{\theta\} = \sum_v \sum_h e^{-E_v(v,h,\theta)} \]

The same principle can be used to obtain the conditional probability distribution of the input layer and the hidden layer, as follows:

\[ P\{v|h; \theta\} = \prod_i P(v_i|h) \]

\[ P\{h|v; \theta\} = \prod_j P(h_j|h) \]

This paper takes the employment rate as the characteristic factor of the model and the number of graduates and the economy as the influencing factors, so the input elements are the number of college students and the economic situation. From formula (8) and formula (9), it can be determined that the input data is firstly input into the deep belief network for learning, it is mapped to the hidden layer according to formula (4), and the data is reconstructed according to formula (9). Then the error between the heavy data and the input data is calculated, and the gradient descent algorithm is used to adjust the connection weights and thresholds. When the reconstruction error reaches the minimum, it is terminated.

Although the theoretical and experimental research on deep belief network has lasted for many years, the experimental results show its good performance and have fully realized that it has a huge application prospect in the field of target classification and recognition. However, it is undeniable that the deep belief network is still at the primary stage of its discipline, and its development and design also focus on experiments, which can be improved through additional studies. Therefore, there are still problems to be further solved, such as how to adopt more appropriate indicators to control the training process of RBM, how to speed up the training and improve the recognition rate at the same time, and how to effectively solve the generalization ability in the small sample problem.

**System Model Construction**

The system model is constructed according to the following steps:

1. Analyze the influencing factors affecting employment data management and collect data on the influencing factors through simulation experiments.
2. For different influencing factor values, different employment data management characteristics are obtained, and the employment rate is selected to represent the employment data management characteristics.
3. Due to the inconsistency of units in the factors that affect the characteristics of employment data management, the so-called dimension inconsistency problem, factor data cannot be directly input into the deep learning algorithm for mathematical modeling, and dimensionless operations must
be performed first. The extreme value method is used to process the factor data of the management characteristics of employment data—that is:

\[
x_i(h) = \frac{x_i(h) - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}
\]

where \(x_i(h)\) represents the hth factor of the mathematical modeling sample of the h employment data management characteristic.

4. Set the parameters of the deep learning network, weights, initial values of thresholds, learning accuracy, etc.
5. Input the training samples of the mathematical modeling of employment data management characteristics into the deep learning network for learning, and continuously update the parameters. When searching for the optimal parameters, the optimal mathematical model of employment data management characteristics is established.
6. Use test samples to analyze the performance of the optimal mathematical model of employment data management characteristics and output the results.
7. Perform mathematical analysis on the output results and make corresponding predictions.

The mathematical modeling of the employment data management characteristics of deep learning is shown in Figure 2.

Figure 2. Mathematical modeling process of employment data management characteristics of deep learning
Simulation Experiment and Model Evaluation

The original data set of college students’ employment is divided into a conditional attribute set and a target attribute set, and each employment data attribute is generalized to divide the value interval of the continuous attribute of the employment data in the original data set of college students’ employment into many small areas. There is a discrete symbol in each interval. In this way, a decision system can be obtained, and the employment data without differences in the decision system can be merged to obtain the decision system used to support the classification model.

In order to make the results of the mathematical modeling experiment of employment data management characteristics more convincing, five simulation experiments were carried out, and a mathematical modeling sample of employment data was collected each time for different years. In this paper, the employment of students and the economic situation of China’s colleges and universities in the past 20 years are selected; and the sample numbers used to carry out mathematical modeling of employment data management include 3 years, 5 years, 10 years, 15 years, and 20 years, as shown in Figure 3.

For the training sample data set of mathematical modeling of employment data management characteristics, three methods are used to establish mathematical models: linear regression, BP neural network, and deep learning; then the test samples are analyzed and the employment statistics of each experiment are calculated by formula (11). The average precision of mathematical modeling of data management characteristics is shown in Figure 4. The mathematical modeling error of employment data management characteristics is calculated by formula (12), and the results are shown in Figure 5:

\[
Corr = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{R_i - \bar{R}}{\sigma_R} \times \frac{P_i - \bar{P}}{\sigma_p} \right)
\]  

(11)

Among them, \( n \) is the number of predicted sample points; \( R_i \) and \( P_i \) are the actual mean and predicted mean of the test sample points, respectively; \( \bar{R} \) and \( \sigma_R \) are the mean and standard deviation of \( R_i \), respectively; \( \bar{P} \) and \( \sigma_p \) are the mean and standard deviation of \( P_i \), respectively.

Figure 3. Sample data for mathematical modeling of employment data management characteristics
The formula for calculating the mean absolute error is as follows:

\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |R_i - P_i| \]  

(12)

It can be seen from Figure 4 and Figure 5 that:

- The mathematical modeling accuracy and average error of the employment data management characteristics of the linear regression method are 81.99% and 18.01%, respectively, and the mathematical modeling error is much higher than the requirements of practical applications, which is mainly due to the linear regression. Due to the limitations of the method, the modeling results have no practical application value.

- The mathematical modeling accuracy and error mean of employment data management characteristics of BP neural network are 89.55% and 10.45%, respectively. Compared with the linear regression method, the mathematical modeling error of BP neural network is reduced by 7.56%. This is because the BP neural network has good nonlinear learning ability; however, because the performance of the BP neural network is not stable, there are many sample points with large errors, and the overall mathematical modeling accuracy is low.

- The mathematical modeling accuracy and average error of the employment data management feature of in-depth learning are 94.36% and 5.64%, respectively. Compared with the linear regression method and the BP neural network, the average mathematical modeling error of deep learning is reduced by about 12.37% and 4.81%. The accuracy of mathematical modeling has been improved. This is because the deep learning algorithm has established a better mathematical model, which accurately describes the changing relationship between the characteristics of employment data management and influencing factors. The comparison results verify the advantages of deep learning mathematical modeling methods.
To summarize, whether it is the accuracy or error of modeling, the deep learning mathematical modeling method has good advantages.

**Application Examples**

This study uses the employment situation and related data of college graduates in China from 2001 to 2020 as the modeling sample and the related data from 2016 to 2020 as the test sample. Since there are many factors affecting the employment situation of college students, socioeconomic status and the number of graduates are the most critical factors, and their data are shown in Figure 6 and Figure 7. Since 2001, China’s total GDP has been developing rapidly, from US$1.2 trillion to US$14.8 trillion, with an average growth rate of 5%. In the future, China’s economy will also maintain a growth rate of 5%. In the past 20 years, the number of Chinese college graduates has increased nearly eightfold; in 2001, there were 1.15 million graduates nationwide, and in 2020, there were 9.74 million graduates. However, in recent years, the growth rate of college graduates has continued to decline, and the number of graduates has remained stable, mainly due to the Covid-19 pandemic and the
global economic depression. The prediction model was built according to the experimental method described above, and the prediction result of the model is shown in Figure 8. It can be seen from Figure 8 that the prediction results based on the deep belief network are more accurate, the prediction accuracy is relatively high in the case of a small number of samples, and the average prediction error is only within 5%. Combined with the characteristics of deep belief networks, as the amount of data increases, the key information in the data can be more accurately extracted during model training, which will make the prediction results of the model more accurate.

ANALYSIS AND DISCUSSION

The management of employment data of college students has always been an important issue. However, it is also a thorny issue because there are many factors that affect employment, including students’ own knowledge, employment awareness, socioeconomic status, social stability, and job shortage.
(Guo & Qi, 2022). These random changes greatly increase the difficulty of student employment data management. According to the evaluation of the linear regression method, the BP neural network, and the DBN, the mathematical modeling accuracy and error mean of the employment data management characteristics of the linear regression method are far higher than the requirements of practical application, but the modeling results do not have any practical application value (Jiang et al., 2022). Compared with the linear regression method, the average value of mathematical modeling error of the employment data management characteristics of BP neural network is reduced, but its performance is unstable, there are many sample points with large errors, and the overall mathematical modeling accuracy is low. Compared with linear regression method and BP neural network, DBN optimizes the mathematical modeling accuracy and average error of the in-depth learning employment data management feature, improves the accuracy of the model, reduces the average error of the model, and accurately describes the relationship between the employment data management feature and the influencing factors (Ma & Liang, 2022). Therefore, the DBN model has good advantages in the application of the characteristics of college students’ employment data management, and its ability to fit deep data mining and nonlinear problems is stronger than linear regression method and the BP neural network model. In addition, a comparative analysis of the data management results of this model and the support vector machine model on college students’ employment shows that there are differences in the tracking results of the two models on college students’ employment (Steele et al., 2020). The recall rate of the tracking results of the model in this paper is better than that of the support vector machine model, the miss rate and error rate are low, and the tracking performance is good.

Through the application analysis of the DBN model in employment data management, in the past 20 years, the employment rate of Chinese college students has been mainly influenced by the number of graduates and social and economic conditions (Li et al., 2020). When the social economy is developing rapidly, the employment rate is significantly higher. When the economy develops steadily, the student employment rate also eases. In the next few years, due to the pandemic, social and economic development will be greatly impacted, the overall economy will decline, the national CDP growth rate will be less than 5%, the social industry will be seriously affected, employment opportunities will be greatly reduced, and the employment rate of college students will also be lower. Therefore, in order to increase the employment rate of college students, rapid economic development should be resumed, the transformation and change of social industries should be accelerated, and more employment opportunities should be provided for the unemployed.

At present, the management of student employment data in Chinese colleges and universities has great deficiencies, for which more attention should be paid to employment management. First, there is a need to update the concept of employment management for college students. In the practice of college students’ employment management, colleges and universities should enhance their awareness of employment management work and clarify the importance of employment management work. At the same time, aiming at the shortage of college students’ employment management practice, colleges and universities should promote the integration of new management concepts and methods and optimize the college students’ employment management concept, so as to promote the sound development of college employment management. Second, colleges and universities should strengthen the publicity of employment management, increase the attention of college students to the integration of employment guidance, promote the integration of college students’ employment management and professional education, and rely on college students’ professional education to promote the development of employment management, thus helping college students’ professional development. Third, there is a need to build an employment management information platform for college students. Under the background of education informatization, the practice of college students’ employment management should establish network education thinking, build the college students’ employment management informatization platform, and promote the construction of college students’ employment management informatization. In addition, the development of college students’ employment management uses the internet as a platform to carry out the second class of college students’ employment management, promote
the integration of online teaching and classroom teaching, and extend the platform of employment
guidance and employment management. Fourth, according to the characteristics of college students’
career development, colleges and universities should promote the online implementation of employment
management, integrate employment guidance and employment management information, and use big
data technology to build a data platform for college students’ employment management, providing a
means for college students to obtain career information and employment management information.
Finally, it is necessary to strengthen the in-depth analysis, application, and tracking of college students’
employment data. The characteristics of college students’ employment data are difficult to collect and
analyze, mainly because there are too many influencing factors and there is no inevitable connection
between them. In this regard, data analysis needs to establish some reasonable mathematical models
and, through the establishment of models, carry out multi-dimensional analysis of various data and
achieve certain predictions (Lijuan & Changliang, 2017). In addition, employment management also
needs to track and analyze the employment information of graduates, obtain a new employment chain,
provide good data support for subsequent college graduates, and provide employment data for colleges
and universities, government, and other departments to establish scientific management processes.

CONCLUSION

This paper mainly adopts the relevant algorithm combining big data and deep learning for research.
Based on the big data of Chinese college students’ employment, this paper analyzes the current
employment figures by establishing a DBN model and proposes corresponding management measures.
The main conclusions are as follows. According to the comparative evaluation of the application of
linear regression, BP neural network, and DBN models, it is found that the DBN model has better
accuracy, a lower rate of error, and better advantages in the application of college students’ employment
data management characteristics. The development of social economy and the number of college
graduates are the key factors affecting the employment rate of college students. The better the social
and economic development, the faster the industrial structure is updated, and the more social jobs, the
higher the employment rate will be. To improve the employment rate of college students, it is necessary
for colleges and universities to update the concept of employment management of college students
and enhance their awareness of employment management. Through the use of big data technology, a
data platform for college students’ employment management is built to provide a means for college
students to obtain professional information and employment management information. In addition,
we should strengthen the in-depth analysis, application, and tracking of college students’ employment
data to provide reliable support for subsequent employment data management.

CONFLICT OF INTERESTS

The author declares that they have no known competing financial interests or personal relationships
that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study are available from the corresponding author upon request.

FUNDING

This work was supported by National Natural Science Foundation of China: Management Research
Innovation Method Based on Portfolio Innovation Theory (No.71801195); Research and Practice
Project of Higher Education Teaching Reform in Henan Province: Research on Talent Training Model
of Local Applied Undergraduate Universities (No.2014SJGLX069).
REFERENCES


He, Y., Zhang, W., Xu, W., & Sui, X. (2022). Exploring the employment quality evaluation model of application-oriented university graduates by deep learning. *Computational Intelligence and Neuroscience*, 2022, 1–10. Advance online publication. doi:10.1155/2022/2823614 PMID:35502350


Yibei Yin is a professor in the School of Electronic Information and Electrical Engineering at Anyang Institute of Technology and specializes in deep learning.