

## Research Article

# Research on the Evaluation Model of Educational Management Theory Based on Data Mining from the Perspective of Neural Network

Die Zheng 

Public Affairs Department, Beijing Technology and Business University, Beijing 100048, China

Correspondence should be addressed to Die Zheng; zhengdie0518@cumt.edu.cn

Received 19 October 2021; Accepted 18 November 2021; Published 20 December 2021

Academic Editor: Deepak Kumar Jain

Copyright © 2021 Die Zheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The improvement of the theoretical quality of education management is an indispensable part of a country's education modernization. However, the existing research on the evaluation of educational management theory is still relatively small, and there is a lack of scientific educational management theory evaluation model. Designing a comprehensive and accurate educational management theory evaluation model has important theoretical value and practical significance. It is possible to process a lot of information in parallel using the artificial neural network method. By optimizing the artificial neural network, data mining of characteristic information data can be realized. Therefore, this paper uses neural network to conduct data mining on education management theory and conduct a comprehensive system evaluation of education management theory. At the same time, the traditional BP algorithm is improved. To train a neural network with large amounts of data, the BP algorithm uses a lot of gradient calculation, which takes a long time and often results in training going to extremes in the local area. BP neural networks are trained using the particle swarm optimization algorithm, and the backward propagation process in the BP algorithm is replaced with particle swarm iteration. To improve algorithm execution efficiency and speed up neural network training, a large number of gradient operations can be avoided. This can help overcome the limitations of the BP algorithm when dealing with large amounts of data. The improved BP algorithm is applied to the evaluation system of education management theory, and the quality evaluation prediction of management education theory is realized.

## 1. Introduction

Education been around since 1999, and the goal has been to raise the cultural quality of everyone and meet the need for new workers as the domestic economy grows rapidly. As a result of the government's education reforms, the number of students enrolled in many colleges and universities has increased significantly for the last four years. So, we can say that higher education is entering a new phase, one that we can call "popular education." This change has both advantages and disadvantages, just like previous reforms. After some colleges and universities expand their enrollment, due to the influence of many factors (such as the lack of teaching resources and teachers and the quality of student sources), the quality of school education management theory has shown a downward trend. Education balance has gotten

a lot of attention recently, especially with the rapid development of "Internet+" education and the full implementation of the strategic goal of revitalizing the country through science and education. High-quality education has now become a prerequisite for people's livelihood and a better life. High-quality education does more than just instill moral values and cultivate citizens. It is also used as a yardstick for gauging the effectiveness of global education and social justice initiatives. Creating a workable school education management theory evaluation system is a pressing issue in ensuring that education continues to develop in a healthy manner. School administrators try every means to improve the quality of school management education, but they must ensure that they achieve their goals without adding additional manpower and material resources. There is no doubt that this will be a huge challenge [1–7].

Research on school education management theory evaluation has tended to focus on identifying problems and making recommendations, rather than conducting quantitative assessments. The educational management theory evaluation system's evaluation indicators and methods are still in the exploratory stage in these quantitative studies. In addition, there are issues such as incomplete indicators and incomplete methods of evaluation. Students, teachers, teaching space, sports fields, teaching equipment value, computers, textbooks, and teachers with higher education degrees are all considered in the current evaluation process. Other indicators include the number of full-time teachers and the number of teachers with advanced degrees. The above indicators evaluate the school's teaching management from multiple perspectives such as students, teachers, building area, and teaching equipment. It can reflect the actual level of school education resources to a certain extent, but the ability to reflect the theory of education management is slightly lacking. To enrich the evaluation system for education management theory, this article is reconstructing the corresponding evaluation system and researching the evaluation indicators that can fully reflect the theory of school education management [8–15].

School management education theory can be used for scientific and systematic evaluation in today's society, where informatization is constantly developing. Information technology can be used to establish a more complete system of management education theory evaluation. The development of the contemporary information society will inevitably lead to the popularization of information technology in all sectors. The use of information technology to establish an educational management theory evaluation system is more than just a sign of progress in social and technological terms. At the same time, it can improve teaching quality better and faster and establish an informatized teaching system. The research results of this article can provide data support and decision support for education and other related departments and can effectively verify the scientific nature of education and related departments' plans. And to further improve the efficiency of education work and service quality is significant to further promote the balanced development.

The contribution of this work can be concluded as following content: to train a neural network with large amounts of data, the BP algorithm uses a lot of gradient calculation, which takes a long time and often results in training going to extremes in the local area. BP neural networks are trained using the particle swarm optimization algorithm, and the backward propagation process in the BP algorithm is replaced with particle swarm iteration. To improve algorithm execution efficiency and speed up neural network training, a large number of gradient operations can be avoided. This can help overcome the limitations of the BP algorithm when dealing with large amounts of data.

## 2. Related Work

The theory of educational management is an important part of college and university teaching, and it is one of the most important things keeping colleges and universities alive

and growing in the present day. It is possible to gauge the quality of educational management theory by looking at various indicators. These indicators are influenced by a wide range of factors. Education management theory evaluation is a nonlinear abstract problem that is difficult to model mathematically or analyze with a mathematical formula. Because of its strong nonlinear processing capabilities, the neural network model can achieve mutual mapping in any dimension. An effective measure to address the issue of measuring teaching quality in higher education institutions is to build a neural network model. To improve teaching quality, the neural network model must be applied to educational management theory evaluations. This promotes continuous improvement of teaching goals and enhances scientific decision-making in educational institutions. The development of intelligent and standardized teaching management in colleges and universities will also be greatly aided by this research. To solve problem of complex and abstract nonlinear university teaching quality evaluation, the author consulted dozens of research documents related to artificial neural networks and university teaching management quality evaluation at home and abroad and analyzed the research status.

Since the neuron mathematical model was proposed in the 1940s, many scholars had been interested in the research of artificial neural networks. Literature [16] proposed the perceptron model, which was called the world's first real artificial neural network model. In the 1970s, because it was impossible to prove theoretically that the multi-layer network's perceptron model was meaningful, the research on artificial neural networks entered a low point. Literature [17] absorbed and summarized the research results of artificial neural network structure and algorithm, proposed the Hopfield model, and proved that under certain conditions, and the network could reach a stable state. In 1986, the error back propagation algorithm was proposed in the literature [18]. Since then, the BP neural network had been widely used in different fields. At the same time, BP had problems such as slow convergence, easy to fall a local location, and fixed learning rate. Scholars at home and abroad have begun to improve in many aspects such as activation function selection, network structure design, acceleration of learning rate, and learning methods. Document [19] described the relationship between the decision tree algorithm of machine learning and the hidden layer of the neural network to determine the structure of the neural network. Literature [20] proposed the elastic BP algorithm to evaluate the influence of the gradient on the network convergence speed. Literature [21] used the function variable step size search method to improve the algorithm to improve the learning speed of the network. Literature [22] proposed an algorithm that can self-adjust the network learning rate, which has a higher learning speed and reduces oscillation. Literature [23] added a term proportional to the error on the basis of the momentum term and improved the BP algorithm from the aspects of the excitation function and the error function.

Literature [24] first nonlinearly normalized the original data, then proposed memory-type initial weights and thresholds, and optimized the parameters, which ultimately

improved the prediction accuracy and convergence speed of the BP neural network. Literature [25] proposed a method to modify the learning factor, which could adjust the size of learning factor during weight adjustment process. Literature [26] introduced the principle of compressed mapping to the BP, which accelerated the convergence speed and made up for insufficiency of the selection of the initial value of the weight. Literature [27] proposed an improvement method for the overall variable learning rate based on standard BP, which significantly improved the calculation accuracy and learning speed. Literature [28] proposed improved BP weights and thresholds to make up for its slow convergence speed and other shortcomings. Literature [29] proposed the BP based on conjugate gradient method, and literature [30] improved the Newton method and proposed the LM algorithm. Literature [31] proposed to combine BP neural network with genetic algorithm and ant colony algorithm.

### 3. Method

**3.1. BP Algorithm.** Artificial neural network is a mathematical model that can simulate the nervous system of the human brain to process complex information after understanding and sampling the human brain structure and external stimulus response mechanism. The model has the characteristics of intelligence, high fault tolerance, and autonomous learning. It can realize nonlinear relationships while performing complex logic operations and has a wide application for pattern recognition and image processing. An artificial neural network is composed of highly connected neuron nodes. Among the various artificial neural network models currently known, multilayer feedforward network is the most widely used. Its basic structure is shown in Figure 1.

The network is divided into three layers. When information comes in, it goes into the input layer, which processes it before sending it to the artificial neural network to be processed further. The input layer sends data to the hidden layer, which receives it and processes it. The user has no idea what is going on behind the scenes. In the output layer, the neurons in each layer form a fully interconnected connection, and the neurons in each layer are not connected.

Artificial neural networks have the ability to approximate unknown functions with arbitrary precision and some derived functions, making neural networks more and more popular. It has been widely used in various fields. The optimization is mainly the network weights and network topology, and the most important is the optimization of neural network weights. The commonly used algorithm for training neural networks is the BP algorithm.

In order to run the artificial neural network, it must go through three stages: (1) training and learning stages. It repeatedly provides a series of input and output modes to the artificial neural network and continuously adjusts the interconnection weights between nodes until the specific input produces the desired output. There are many algorithms for training artificial neural networks, among which the most widely used is the back propagation algorithm (BP algorithm). (2) Reverberation stage: input a series of

input patterns that have been used in the training phase to the artificial neuron network and adjust the system to make it more reliable. (3) Prediction stage: input a new model to the neural network to make predictions on unknown samples.

The main steps of backpropagation are as follows:

*Step 1.* First carry out forward calculation and calculate the activation value of each layer separately.

*Step 2.* For each neuron  $i$  in the output layer, calculate the residual of each node according to the following formula:

$$\delta_i^{(l)} = \frac{\partial}{\partial z_i^{(l)}} \frac{1}{2} \|y - hw, b(x)\|^2 = -(yi - a_i^{(l)})f'(z_i^{(l)}). \quad (1)$$

Among them,  $z_i^{(l)}$  and  $a_i^{(l)}$  are input sum and activation value.

*Step 3.* For other layers except the output layer, the residual is calculated as follows:

$$\delta_i^{(l)} = \left( \sum_{j=1}^{sl+1} w_{ji}^{(l)} \delta_j^{(l+1)} \right) f'(z_i^{(l)}). \quad (2)$$

*Step 4.* Calculate  $\partial E/\partial w$  and  $\partial E/\partial b$  based on the residual:

$$\frac{\partial E}{\partial w_{ij}^l} = a_i^{(l)} \delta_i^{(l+1)} \quad (3)$$

$$\frac{\partial E}{\partial b_i^l} = \delta_i^{(l+1)}. \quad (4)$$

*Step 5.* Update weights and biases:

$$w' = w - \alpha \frac{\partial E}{\partial w} \quad (5)$$

$$b' = b - \alpha \frac{\partial E}{\partial b}, \quad (6)$$

where  $\alpha$  is the learning rate.

According to the above formula, the key of the gradient descent method is calculating the loss function. This led to the Back Propagation (BP) algorithm, which is an efficient algorithm for calculating partial derivatives.

One of the most widely used neural network models is the BP neural network. The BP algorithm, on which it is based, is well established and widely used. Even the most complex nonlinear relations are fully approximated by it. Because the information is dispersed and stored in the network's neurons, it has high fault tolerance. Due to the parallel processing, the calculation is lightning fast. Because neural networks are self-learning and adaptive, they can deal with uncertain or unknown systems. When it comes to processing quantitative and qualitative information at the same

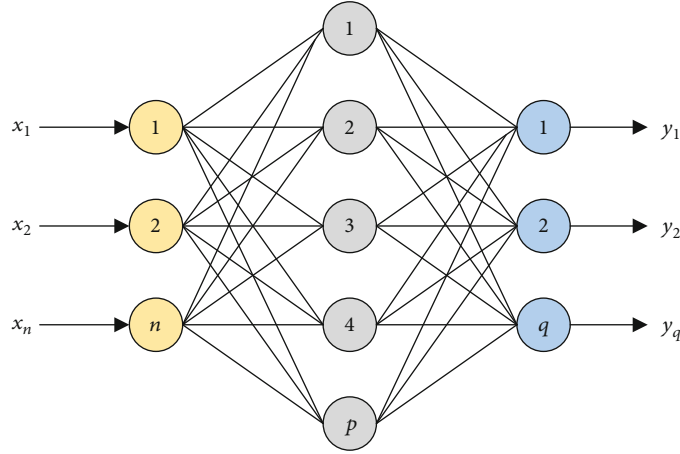


FIGURE 1: BP structure.

time, this system excels. It can coordinate a wide range of input information relationships, making it ideal for fusion and multimedia applications.

Although BP was the main factor in the successful application of neural networks in the past, its performance is increasingly unable to meet the requirements of practical applications. The BP is a feed-forward network. This kind of network only makes the network have complex nonlinear mapping through the compound action of many neurons with simple processing capabilities. The traditional BP learning algorithm is based on the essence of gradient descent, which inevitably brings the following problems: (1) the generalization of the network, whether a large number of input vectors that have not been learned can be processed correctly, and whether the network has a certain predictive ability. (2) The error surface of the network based on the BP algorithm has many global minimum solutions. There are many local minimum points. Under certain initial value conditions, the result of the algorithm will fall into the local minimum. When learning rate is set large, oscillation may occur. (3) The amount of calculation is large, the convergence speed of the learning process is slow, and the training time is long. (4) A lack of robustness means the network's performance is more dependent on the network's initial settings. (5) The number of layers and layer nodes lacks a unified and complete theoretical guidance.

**3.2. PSO Algorithm.** The idea of particle swarm optimization (PSO) is derived from the bird predation behavior. When a flock of birds is searching for food at random, the simplest and most effective strategy is to search the area around the bird closest to the food if there is only one piece of food there. An optimization algorithm known as the PSO algorithm is based on this model. In addition, one of the fundamental concepts of PSO is that people frequently base their decisions on their own and other people's experiences in making those decisions.

When comparing particle swarm algorithms with others, it is easy to see how they are related: it also relocates the members of the group to a better location based on how well they can adapt to the surrounding environment. Because of

this, some believe it is a form of algorithm that evolves over time. It differs from other evolutionary algorithms in that it does not use individual evolutionary operators. As a result, each person in the search space is treated as if they were a massless particle traveling at a constant speed. This speed can be dynamically adjusted based on an indepth analysis of individual and collective flight experience.

Suppose the position in the group is  $X_i$ , the best position is  $p$ . The best experienced position is  $pg$ , and the velocity of particle  $i$  is denoted by  $V_i$ . For each iteration, the movement of particle  $i$  in space follows the following equation:

$$v_i^{k+1} = v_i^k + \alpha_1 \text{Rand}() (p^k - x_i^k) + \alpha_2 \text{Rand}() (pg^k - x_i^k), \quad (7)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1}, \quad (8)$$

where  $\alpha_1$  and  $\alpha_2$  are the acceleration constants. They make each particle accelerate to the position  $p$  and  $pg$ , and  $\text{Rand}()$  is a random number varying in the range of  $[0, 1]$ .

Particles trust their current state of motion; so, they move inertia-free according to their own speed, which is represented by the particle's previous speed. The cognitive part, on the other hand, represents how the particle thinks. To put it another way, the likelihood of increased randomness in the future will rise. Cognition is the behavior being discussed, and it is posited that acquiring correct knowledge is made easier as a result, leading to an improved learning experience. The social component, which represents the sharing of information and cooperation among particles, completes the process. These psychological assumptions of particle swarm optimization are uncontroversial. In the process of seeking consistent cognition, individuals often remember their beliefs. At the same time, considering colleagues' beliefs, when individuals perceive beliefs are better, it will make adaptive adjustments.

Unlike most optimization algorithms based on gradient information, particle swarm optimization relies on probabilistic search algorithms. Although probabilistic search algorithms use more functions, their benefits are still efficient compared with traditional algorithms. (1) there is no



centralized control constraint, and the solution of the entire problem will not be affected by the failure of individual individuals, which ensures that the system has stronger robustness. (2) The scalability of the system is ensured by indirect information exchange. (3) It is suitable for parallel distributed algorithm processing and can make full use of multiprocessors. It can be seen from the working process of the particle swarm algorithm that parallel operations can be performed on the search space. At the same time, the search process of each individual can also be carried out in parallel, and the advantages of parallel computing can greatly reduce the time to solve the problem in the case of solving a complex-scale optimization problem with a very large amount of calculation. (4) There are no special requirements for the continuity of the problem definition. This is a very good property of particle swarm optimization that distinguishes it from traditional optimization algorithms, which expands the scope of its application. (5) There are few parameters that need to be adjusted, so that specific optimization algorithms can be configured by adjusting fewer parameters for specific problems in engineering. (6) Because of its simplicity, the algorithm can be quickly and easily implemented. Simple mathematical operations are all that the algorithm requires. This method only needs to know the objective function's output value, not its gradient information, because the data processing process does not require a lot of CPU or memory. According to the current state of particle swarm algorithm and application method research, the particle swarm optimization approach is a novel approach that can successfully address the majority of global optimization challenges. More importantly, the potential parallelism and distributed characteristics of the particle swarm algorithm provide technical guarantee for processing a large amount of data in the form of a database. Whether it is analyzed from the perspective of theoretical research or application research, particle swarm algorithm and application research have important academic significance and practical value.

**3.3. Improved BP Based on PSO.** Convergence is quick, operation is simple, and implementation is simple for the particle swarm optimization algorithm. The genetic algorithm does not include any complicated operations like encoding and decoding, hybridization, or mutation. However, there are drawbacks to using the particle swarm optimization algorithm. There is a directionality problem here. The algorithm moves towards an optimal solution based on all of the particles and its own search history. In the later stage of evolution, there is a slowdown and premature phenomenon.

The PSO algorithm initially randomly generates a set of solutions of the objective function, and each solution is called a particle. Each particle searches for the optimal solution follows the method of the current optimal particle. Adjust the parameter form of the expression to obtain a new formula to update the particle:

$$v(i+1) = w * v(i) + \alpha_1 \beta_1 (p - x(i)) + \alpha_2 \beta_2 (g - x(i)), \quad (9)$$

$$x(i+1) = x(i) + v(i+1), \quad (10)$$

TABLE 1: The division of dataset.

Training	Validation	Testing
157	37	120

TABLE 2: The experimental environment.

Item	Parameter
CPU	Intel Core i7
Operating system	Windows 10
Software	Matlab

TABLE 3: Grade-score correspondence table.

Grade	Parameter
Excellent	[90, 100]
Good	[80, 90]
Average	[70, 80]
Qualified	[60, 70]
Poor	[0, 60]

where  $p$  represents the optimal solution, and  $g$  represents the optimal solution generated by PSO, that is, the global optimum found until the current generation.  $w$  is called the inertia factor, which is used to adjust the movement speed of the particles during this iteration according to the speed obtained in the previous iteration.  $\alpha$  is the acceleration coefficient. If it is small, the particle may be far away from the target, and if it is large, it leads to flying over the target.  $\beta$  is the random numbers between  $(0, 1)$ . The fitness value calculated according to the objective function is utilized to measure which position is better. The fitness value drives each particle toward the individual optimum in the search space and the global optimum found so far.

It can be seen that when the current position of a particle is close to but has not reached the global optimal position, if the movement speed and inertia weight of the particle in the previous iteration are nonzero, the particle will gradually move away from the position. In addition, once all particles catch up with the current global optimal particle, these particles will stop moving. At this time, it may be premature or the optimal solution has been found; so, it is extremely necessary to judge whether it is premature.

In order to avoid premature maturity, consider introducing some interference factors in the algorithm search process. There are two necessary conditions for the occurrence of premature convergence of the algorithm: each particle catches up with the global optimal particle, and the speed of the particle rapidly decreases to close to zero.

When the inertia weight is close to zero, it also affects the speed. If the inertia weight is set closer to zero, the effect can be ignored. For a random search process, the first condition is unavoidable. Therefore, to avoid the premature PSO, it is important to consider possibility of destroying the second

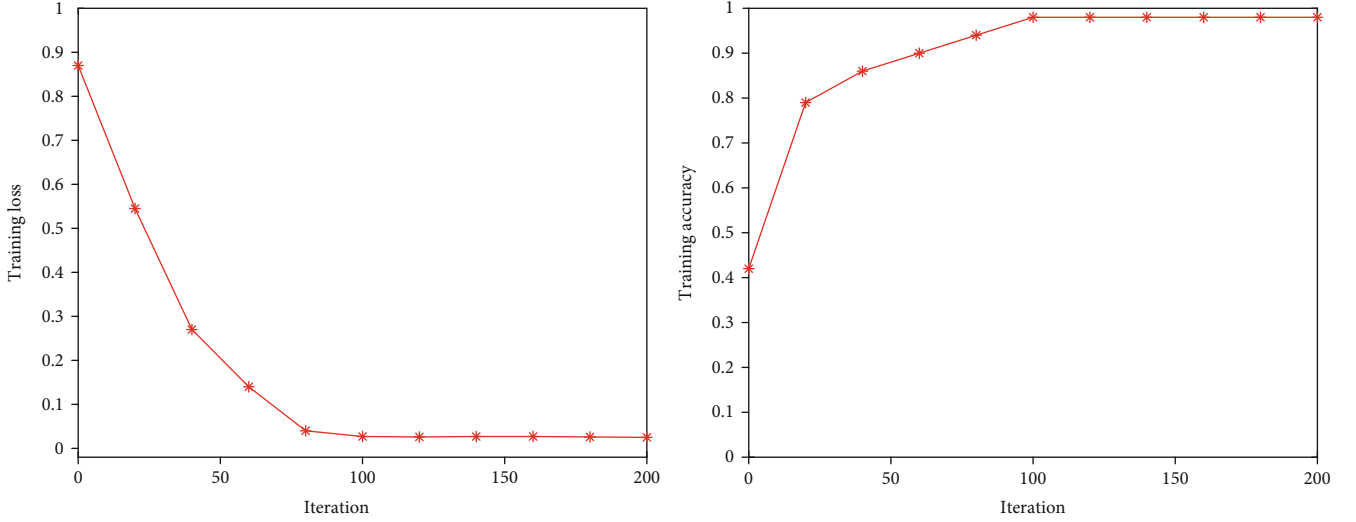


FIGURE 2: Training loss and training accuracy.

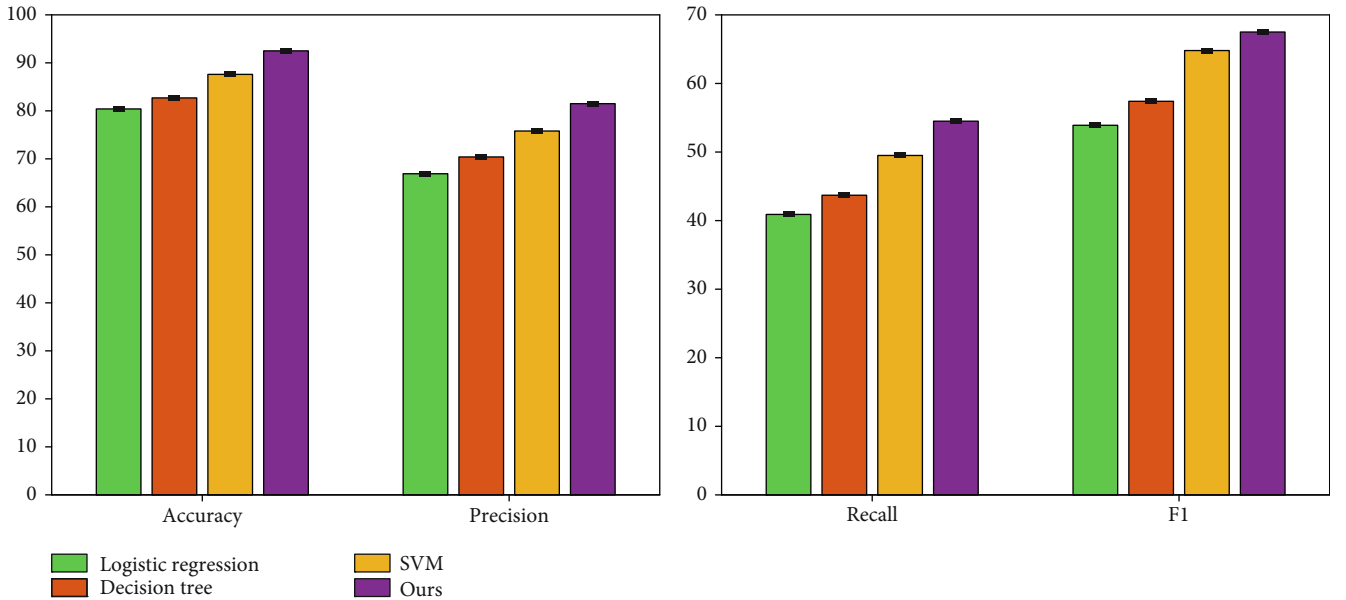


FIGURE 3: Comparison with other method.

condition. When precocity occurs, add interference conditions to the velocity of the particles:

$$v(i) = v(i) + \text{rand}(), \quad (11)$$

$$x(i) = x(i) - 2v(i). \quad (12)$$

When precocious puberty occurs, formula (11) adds a random number to the speed, thereby changing the velocity of the particle, and formula (12) moves the particle to a new position and changes the search direction. Equation (11) is suitable for single-mode functions, and formula (12) is suitable for multimode functions. The interference described in Equations (11) and (12) causes the particles to start searching at a new location at a new speed, jumping out of the orig-

inal premature region. Avoiding the local optimal solution can better find the global minimum.

This PSO algorithm introduces premature judgment and adjustment of particle velocity and position under premature conditions. The global optimization capability has been strengthened, which is called improved PSO (IPSO) here.

The training of the BP algorithm is a supervised method, which uses the method of error back propagation to make the weight of the network drop along the gradient randomly. Each weight will affect the error generated by the output layer; so, in order to reduce the error, the weight matrix should be adjusted according to the error of the output layer. It forms a process in which the error shown by the output terminal is gradually transferred to the input terminal along the direction opposite to the input signal.

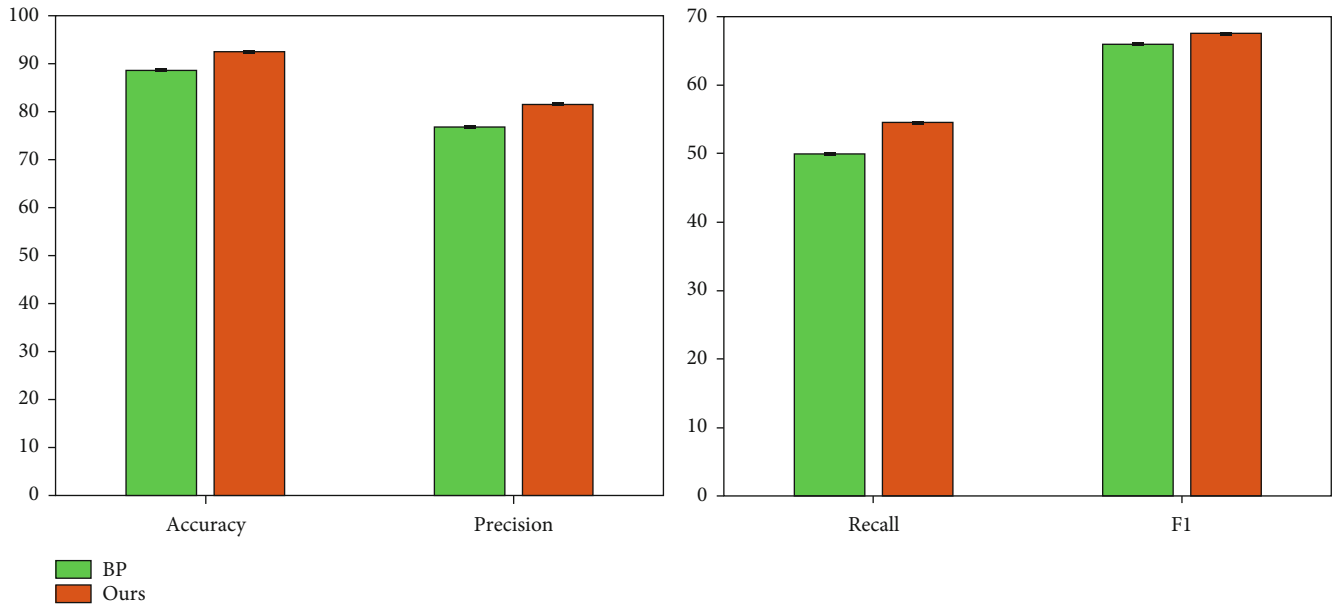


FIGURE 4: Comparison with BP.

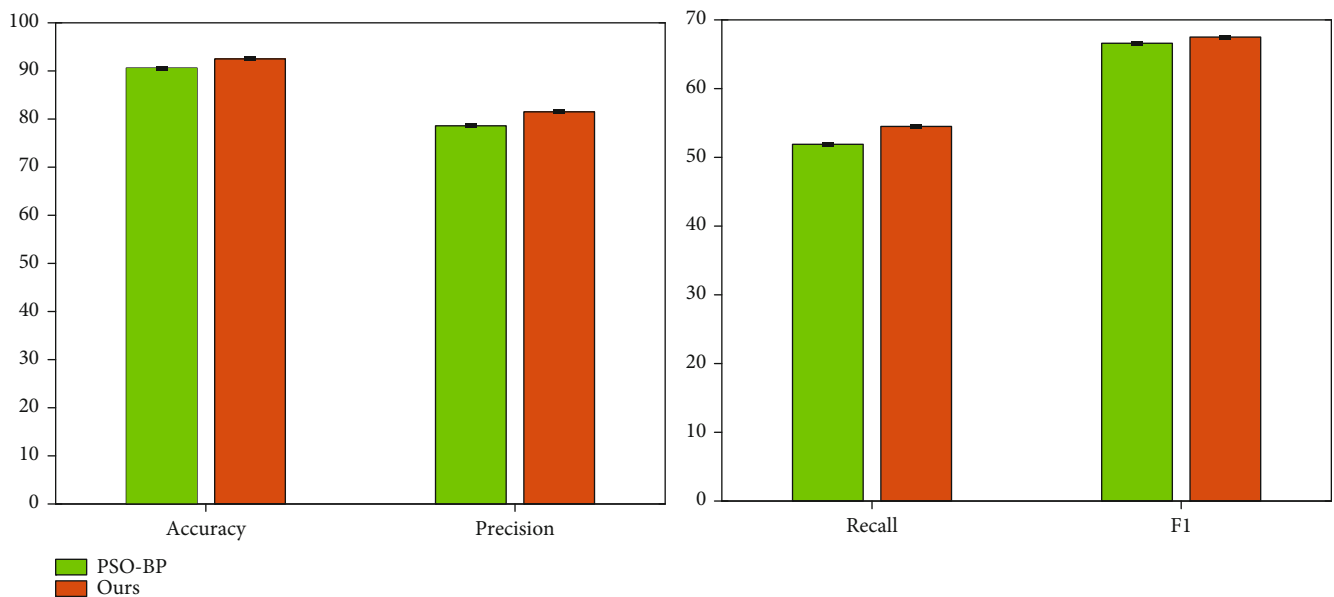


FIGURE 5: Evaluation on IPSO.

IPSO is used to train the BP network in the following way: the position of each particle in the particle swarm represents the set of weights in the current iteration of the BP network, and the dimension of each particle represents the number of weights that are involved in the network. To solve the neural network training problem, the fitness function is applied to the neural network output error from a given training sample set. The neural network’s error threshold is represented by the fitness value. The better the particle performs in the search, the smaller the error must be. A network output layer’s error is minimized by the particles moving and searching in the weight space. In order to reduce the

mean square error, changing the particle speed means updating the network weight. IPSO does this to reduce the MSE by searching and training the neural network’s weights. A new rate is calculated for each iteration, and particles move in a new direction according to their positions. New position is a new weight set, and a new MSE is obtained according to the set of weights. If a particle moves in a new direction such that the new MSE cannot be reduced, the particle does not move in the new direction; that is, the new weight matrix is discarded. The particle with the smallest MSE is the current optimal particle. The training process is repeated on the basis of avoiding premature

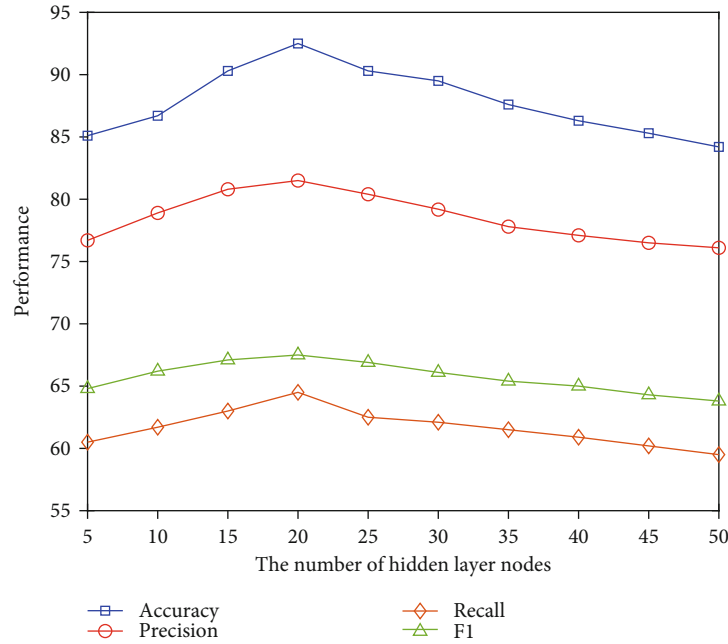


FIGURE 6: Evaluation on the number of hidden layer nodes.

maturity until an error that meets the requirements is generated or the number of calculations set in advance is exceeded. The weights obtained at the end of the algorithm is the final result.

From the above description, it can be seen that the IPSO algorithm is used to train the BP network, which avoids the backward propagation process in the original BP algorithm, accelerates the search speed, and prevents the premature convergence. Aiming at the phenomenon of premature convergence of particle swarms, the BP network trained by the improved particle swarm method with interference conditions has very large randomness in generating the next generation solution set; so, it is difficult to fall a local solution. Moreover, due to the sharing of all solution information in each generation and the improvement of the quality of each solution, the solutions in each generation population have the dual advantages of improving “self”-learning and learning from others. Therefore, it obsesses a faster convergence. In the later stage of the algorithm, new individuals are also trained and optimized, which further improves the convergence speed. Compared with the BP network trained by ordinary PSO, it also has better optimization ability and better generalization performance.

## 4. Experiment and Discussion

**4.1. Dataset.** To evaluate the effectiveness of our method, different educational management theoretical data are collected from universities. These theories include the power of educational teachers, educational software and hardware platforms, the scientific nature of the theory, the validity of the theory, the acceptance of the theory, and the educational results brought by the theory. The data set we collected con-

tains a total of 314 pieces of data, of which 157 pieces of data are training set, 37 pieces of data are validation set, and the remaining 120 pieces of data are the testing set, as shown in Table 1. Then, carry on the network training and test under the hardware and software conditions, as shown in Table 2. And the evaluation of educational management theory adopts a 100-point system, which is divided into five grades, as shown in Table 3.

**4.2. Evaluation on Training Convergence.** In a neural network, the network converges and the convergence effect are important metrics to evaluate the performance. To evaluate this parameter, this paper studies the training loss and training accuracy. The experimental results are shown in Figure 2.

As the network training progresses, the training loss gradually decreases, and the training accuracy rate gradually rises. When the number of iterations reaches 100, the network gradually converges. The algorithm proposed in our paper can achieve the convergence of network training, which is the basis for the reliability and robustness of the network.

**4.3. Comparison with Other Method.** To verify the effectiveness of our method, we compare our method with other methods, including logistic regression, decision tree, and SVM. Comparison indicators include accuracy, precision, recall, and F1 score. The experimental results are shown in Figure 3.

Compared with other methods, the method proposed is significantly better than other artificial neural network methods in accuracy, precision, recall, or F1 score. This proves the effectiveness of this method.



**4.4. Evaluation on IPSO.** As mentioned in the previous part of the article, this article improves the PSO algorithm and combines it with the BP algorithm. To verify the superiority of this combination mechanism, our algorithm is compared with BP algorithm. Experimental results are shown in Figure 4.

Compared with the BP algorithm, after embedding the improved particle swarm algorithm, the evaluation performance of the network can be effectively improved.

In addition, this article improves the traditional PSO algorithm. To verify the effectiveness of our improved method, this paper conducts a corresponding comparative experiment. The unimproved PSO is embedded with the BP and compared with our method. The experimental results are shown in Figure 5.

Compared with the unimproved PSO algorithm, the improved measures in this paper can effectively improve the network performance, which further proves the correctness of the improved measures in this paper.

**4.5. Evaluation on the Number of Hidden Layer Nodes.** In BP network, the number of nodes in the hidden layer is variable. To verify the impact of different node numbers on network, find the best number of nodes. This paper conducts experiments on different numbers of nodes, and the results are shown in Figure 6.

As the number of hidden layer nodes increases, network performance gradually improves. When the number of hidden layers is 20, the performance is optimal. But after reaching the peak, the performance gradually began to decline.

## 5. Conclusion

Education is the cornerstone of a country's national development, designing a comprehensive and accurate educational management theory evaluation model has important theoretical value and practical significance. Artificial neural network research methods have the ability to perform large-scale parallel processing of information. By optimizing the artificial neural network, data mining of characteristic information data can be realized. Therefore, this article uses neural network to conduct data mining on education management theory and comprehensively and systematically evaluate education management theory. Aiming at the two shortcomings of BP neural network, this work introduces the improved PSO algorithm into BP. With the introduction of PSO algorithm, continuous and discontinuous objective functions are globally searched. It effectively solves the shortcomings of long neural network training time and slow convergence due to the large number of gradient calculations in the BP algorithm. Combine the improved PSO algorithm with the BP algorithm and apply it to the evaluation of educational management theory. Through comparative experimental analysis, the algorithm in this paper can make a high-quality evaluation of educational management theory.

## Data Availability

The datasets used are available from the corresponding author on reasonable request.

## Conflicts of Interest

The author declares that he/she has no conflict of interest.

## References

- [1] V. J. Hallenbeck, "Use of high simulation for staff education/development," *Journal for Nurses in Staff Development*, vol. 28, no. 6, pp. 260–269, 2012.
- [2] C. Kai-Ming, "The changing legitimacy in a decentralizing system: the state and education development in China," *International Journal of Educational Development*, vol. 14, no. 3, pp. 265–269, 1994.
- [3] K. Fraser, *Education Development and Leadership in Higher Education*, international encyclopedia of education, 2004.
- [4] F. Khorasgani and Mahdi, "Higher education development and economic growth in Iran," *Education Business & Society Contemporary Middle Eastern Issues*, vol. 1, no. 3, pp. 162–174, 2008.
- [5] I. Kalnins, V. H. Barkauskas, and A. Šešškevičius, "Baccalaureate nursing education development in 2 Baltic countries: Outcomes 10 years after initiation," *Nursing Outlook*, vol. 49, no. 3, pp. 142–147, 2001.
- [6] S. L. Owen, M. A. Kirkpatrick, S. W. Lavery et al., "Selecting and recruiting health programs for the school health education evaluation," *Journal of School Health*, vol. 55, no. 8, pp. 305–308, 2010.
- [7] J. E. Heimlich, "Environmental education evaluation: reinterpreting education as a strategy for meeting mission," *Evaluation & Program Planning*, vol. 33, no. 2, pp. 180–185, 2010.
- [8] J. A. Supovitz and B. S. Taylor, "Systemic education evaluation," *American Journal of Evaluation*, vol. 26, no. 2, pp. 204–230, 2005.
- [9] C. Zhao, Y. Zhao, L. Luo, and Y. Li, "Fractional modeling method research on education evaluation," *Journal of Software*, vol. 6, no. 5, pp. 901–907, 2011.
- [10] J. Yang, Y. Zhang, X. Ye et al., "Dental education evaluation in China: a systematic review," *BMC Medical Education*, vol. 14, no. 1, pp. 1–6, 2014.
- [11] L. K. Olsen, R. Hambleton, R. Simon, D. B. Connell, R. R. Turner, and D. Orenstein, "Development and application of the student test used in the school health education evaluation," *Journal of School Health*, vol. 55, no. 8, pp. 309–315, 2010.
- [12] S. W. Fors and M. E. Doster, "Implication of results: factors for success," *Journal of School Health*, vol. 55, no. 8, pp. 332–334, 1985.
- [13] L. W. Green, T. Cook, M. E. Doster et al., "Thoughts from the school health education evaluation advisory panel," *Journal of School Health*, vol. 55, no. 8, pp. 300–300, 2010.
- [14] L. D. Avery and J. VanTassel-Baska, "Investigating the impact of gifted education evaluation at state and local levels: problems with traction," *Journal for the Education of the Gifted*, vol. 25, no. 2, pp. 153–176, 2001.

- [15] G. G. Wojtowicz, "A secondary analysis of the school health education evaluation data base," *Journal of School Health*, vol. 60, no. 2, pp. 56–59, 2010.
- [16] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958.
- [17] J. Mao, H. D. Zhao, and J. J. Yao, "Application and prospect of artificial neural network," *Electronic Design Engineering*, vol. 19, no. 24, pp. 62–65, 2011.
- [18] D. E. Rumelhart, "Learning internal representation by error propagation," *Parallel Distributed Processing*, vol. 1986, no. 1, p. 1, 1986.
- [19] K. J. Cios and N. Liu, "A machine learning method for generation of a neural network architecture: a continuous ID3 algorithm," *IEEE Transactions on Neural Networks*, vol. 3, no. 2, pp. 280–291, 1992.
- [20] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," *IEEE International Conference on Neural Networks*, vol. 1, pp. 586–591, 1993.
- [21] P. Hao, "Probing modification of forward multilevel neural network," *Computer Engineering and Applications, Computer Engineering and Applications*, vol. 40, no. 7, pp. 17–18, 2004.
- [22] X. Zhang, F. Liu, and W. Mai, "Prediction of loom efficiency based on BP neural network and its improved algorithm," *Journal of Textile Research*, vol. 41, no. 8, pp. 121–127, 2020.
- [23] Y. H. Zweiri, J. F. Whidborne, and L. D. Seneviratne, "A three-term backpropagation algorithm," *Neurocomputing*, vol. 50, no. 1, pp. 305–318, 2003.
- [24] Y. Zhu, "Prediction model of improved BP neural network its application," *Journal of Wuhan University of Technology*, vol. 36, no. 6, pp. 1252–1255, 2012.
- [25] H. Li, H. Lv, and G. Li, "An improving method of BP neural network and its application," *Engineering Science*, vol. 7, no. 5, pp. 63–65, 2005.
- [26] L. Feng and B. P. Optimized, "Neural networks algorithm and its application," *Computer Simulation*, vol. 27, no. 12, pp. 172–175, 2010.
- [27] H. Zhu, J. Wang, and F. Wang, "Optimization of process parameters in the inertial separation chamber of combine harvester-based on improved genetic neural network," *Journal of Agricultural Mechanization Research*, vol. 42, no. 1, pp. 161–166, 2020.
- [28] L. Zhang and X. Lei, "Optimized approach on stomach nourishing decision based on PSO-BP neural network," *Computer Science*, vol. 43, no. 2, pp. 63–66, 2016.
- [29] J. Leonard and M. A. Kramer, "Improvement of the backpropagation algorithm for training neural networks," *Computers & Chemical Engineering*, vol. 14, no. 3, pp. 337–341, 1990.
- [30] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [31] J. T. Tsai, J. H. Chou, and T. K. Liu, "Tuning the structure and parameters of a neural network by using hybrid Taguchi-genetic algorithm," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 69–80, 2006.