Research Article

Importance Degree Evaluation of Spare Parts Based on Clustering Algorithm and Back-Propagation Neural Network

Shoujing Zhang,[1] **Xiaofan Qin,**[1] **Sheng Hu** (iD),[1] **Qing Zhang** (iD),[2] **Bochao Dong,**[1] **and Jiangbin Zhao** (iD)[3]

[1]*Department of Industrial Engineering, Xi'an Key Laboratory of Modern Intelligent Textile Equipment, Xi'an Polytechnic University, Xi'an, Shaanxi 710600, China*
[2]*Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an 710049, China*
[3]*Department of Industrial Engineering, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China*

Correspondence should be addressed to Qing Zhang; zhangq@mail.xjtu.edu.cn

The quantitative evaluation of the importance degree of spare parts is essential as spare parts' maintenance is critical for inventory management. Most of the methods used in previous research are subjective. For this reason, an accurate method for the evaluation of the importance degree combining an improved clustering algorithm with a back-propagation neural network (BPNN) is proposed in the present paper. First, we classified the spare parts by analyzing their historical maintenance and inventory data. Second, we evaluated the effectiveness of classification using the Davies–Bouldin index and the Calinski–Harabasz indicator and verified it using the training data. Finally, we used BPNN to determine the training data necessary for an accurate assessment of the importance degree of spare parts. The previous importance evaluation methods were susceptible to subjective factors during the evaluation process. The model established in this paper used the actual data of the company for machine learning and used the improved clustering algorithm to implement training and classification of spare parts data. The importance value of each spare part was output, which additionally reduced the impact of subjective factors on the importance evaluation. At the same time, the use of less data to evaluate the importance of spare parts was achieved, which improved the evaluation efficiency.

1. Introduction

Spare parts management [1–3] is an important part of inventory management. It has a significant impact on overall business and cost control. Spare parts are necessary materials for manufacturing equipment maintenance. The timeliness of providing spare parts is an important factor to ensure the good condition of the equipment. And with the development of science and technology, components have became more and more complex, the types and quantities of parts are increasing, and the cost of ordering and storing has also increased. Therefore, how to carry out targeted inventory management of spare parts has important research significance. This article proposes a concept to make

corresponding inventory strategies based on the importance degree of spare parts. Therefore, how to achieve accurate, rapid, and objective importance degree assessment is of great significance to the enterprise because it helps to arrange spare parts inventory.

The research on importance is mainly in the field of reliability. Birnbaum [4] first introduced the concept of component importance measurement in 1969. More and more researchers [5, 6] have begun to study how to analyze the entire system based on changes in the reliability importance of components. Si et al. [7] studied the linear continuous k-out-of-n:F and G systems. They verified that the change in the importance index corresponds to the change in the optimal configuration of the system and found

that the importance index does not change monotonously with the change in component reliability. Their research has made a deeper exploration of the importance theory in reliability research. At the same time, following Birnbaum, Natvig [8] first proposed the concept of the life importance measure in 1979 and used it to assess the impact of components on the entire system along with the life cycle. Hong [9] and Borgonovo [10] proposed the concepts of joint importance measure and differential importance measure.

Spare parts have their own attributes. Every spare part has its importance degree, and the evaluation of the importance degree depends on the attributes. It is a process of comprehensive analysis to the spare parts. Therefore, it is necessary to analyze the relevant factors of the importance degree evaluation of spare parts. However, among the aforementioned methods of importance degree evaluation in the field of reliability, the environment of importance degree evaluation is more suitable for a system with components or components connected in series, parallel, or series-parallel. Structural evaluation of components is not applicable for the evaluation of the comprehensive attributes of a spare part; this also means the method cannot be applied to the evaluation of the importance degree of the spare part.

In the research on the importance degree of spare parts, most of the studies have mainly focused on the classification of spare parts and determined the level classification of the importance of spare parts. The most commonly used method is the ABC classification [11, 12], which mainly determines the importance of spare parts based on the price of spare parts. Güven et al. [13] used cost analysis to determine the importance of equipment. Ishizaka et al. [14] proposed a DEA ranking method based on this and divided a large number of items into three categories: A, B, and C. At the same time, some other researchers no longer classify parts based on their price. Chandima Ratnayake and Antosz [15] began to develop some classification standards and used the standards to achieve a more accurate classification process. Roda et al. [16] proposed a multistandard part classification method to deal with the diversity of part manufacturing equipment and provided suggestions for implementing the classification method in actual engineering. Yousefi et al. [17] increased accuracy of the clustering algorithm combined with PCA. Wang et al. [18] proposed an artificial neural network (ANNS) for critical evaluation of power plant spare parts. To a certain extent, the accuracy of classification of spare parts has been greatly improved. Although an accurate classification can achieve a rough grasp of the importance of spare parts, but how to accurately assess the importance degree of each spare part is a problem that current researchers mainly research. For this reason, the evaluation method of the importance degree of spare parts based on the AHP method [19] has been proposed, and it is mainly used in the inventory management of spare parts. Wu and Tsai [20] used both the analytic hierarchy process (AHP) and decision-making trial and evaluation laboratory (DEMATEL) methods to evaluate the criteria in the auto spare parts industry. Huang et al. [21] proposed the BPNN importance degree evaluation model, which established indicators for evaluating the importance degree of spare parts, including functional indicators, wear indicators, and economic indicators. However, the evaluation method mainly uses BP to learn and predict the importance degree of the AHP evaluation value. In essence, the entire evaluation process still mainly depends on the decision of subjective factors. Yang and Du [22] proposed to use FMECA technology to identify faulty spare parts first and used RPN (risk priority number) × MTAT (maintenance turn-around time), GRN (grey relational number), and GRN with MTAT to determine the importance degree of spare parts, respectively. However, in the analysis process, subjective experience needs to be used to score the risk priority number. Therefore, it is inevitable that there will be a large amount of nonobjectivity in the importance degree evaluation process. The comprehensive evaluation and analysis of the importance degree of spare parts not only need to achieve the accuracy of spare parts evaluation but also need to avoid large subjectivity; however, most of the abovementioned research methods have many subjective factors. Although such an evaluation method can realize a comprehensive evaluation process, it is highly subjective.

This article introduces an improved clustering algorithm to effectively and accurately classify the spare parts. By using the real data given by the enterprise, the classified data is filtered into the BP neural network for training so that the data of each type of spare parts can be learned and trained, and through this method, the data of relevant influencing factors is input, and the importance degree value of spare parts can be output by the trained neural network. This method largely overcomes the subjectivity of importance degree in the evaluation process, and this method has great practical value.

An improved clustering algorithm has been introduced to process the data. The Mahalanobis distance clustering algorithm is used to classify repair spare parts; it eliminates data correlation and the influence of the dimension on maintenance spare parts classification, which makes data classification more accurate. The Davies–Bouldin (DB) and Calinski–Harabasz (CH) indices are being used to evaluate the classification effect. Moreover, this method can reduce the influence of subjective human factors and can be used to evaluate the importance degree in a timely, rapid, and accurate manner when new repair parts enter the inventory system. Therefore, the present paper proposes an evaluation method based on the clustering algorithm and BPNN without considering subjective factors to assess the importance degree of maintenance spare parts. The proposed method can be used to quickly and accurately assess the importance degree of spare parts by analyzing historical data once new spare parts are added. The present proposed evaluation method has the following advantages: (1) it quantifies the importance degree of spare parts directly from raw data, without considering expert experience, and (2) when new spare parts enter the inventory system, it evaluates their importance values quickly and efficiently according to the corresponding historical data.

The remaining structure of the present paper is organized as follows. Section 2 outlines the research problems to

be faced and describes the evaluation method in detail. Section 3 describes the verification of an example and an assessment of the accuracy of the evaluation method by analysis of the raw inventory data. Section 4 summarizes the conclusions.

2. Research on the Importance Degree Evaluation Method for Spare Parts

2.1. Problem Description. In the past, assessment of the importance degree of spare parts in inventory management has depended on the experience of inventory managers. Sometimes, the importance degree always evaluated by subjective methods, such as an AHP, and the methods described above are not objective. At the same time, with the development of technology, the technical complexity of spare parts has gradually increased, and the types have shown a trend of diversified development. How to achieve rapid and accurate spare parts importance assessment with less data and information has become an urgent issue. Furthermore, it is difficult to evaluate the importance degree of spare parts once a company adds new spare parts. For this reason, it is necessary to consider how to evaluate the importance degree of spare parts objectively and effectively once new spare parts have been introduced to the inventory system.

2.2. Method for the Evaluation of Spare Parts Importance Degree. For the more accurate evaluation of the importance degree of maintenance spare parts, an evaluation model based on improved cluster analysis and BPNN is proposed in the present paper. The technique overcomes the shortcomings of traditional evaluation methods, which are more reliant on the experience of experts. The proposed method is used to evaluate the spare parts importance degree; the method includes three stages from the perspectives of the main process, as shown in Figure 1.

From the general procedures of the evaluation model, in the first stage, the PCA method is used to deal with data dimension reduction, and then the sample dataset is autonomously divided into different classifications by the improved clustering algorithm, and the rationality of the classifications is evaluated using the DB and CH indices, and the data processed by the improved clustering algorithm is used for the neural network learning and training data. The second stage constructed a BPNN evaluation model; the model was used to predict the importance degree of spare parts by learning the historical data after processing. The third stage verified the effectiveness of the proposed evaluation method by comparing examples with the sample importance degree.

Moreover, there is the advantage of using the data processing method described above in the BPNN. Through the preliminary training of the data, the clustering algorithm performs initial clustering on the sample data, randomly extracts a certain amount of sample data according to the classification ratio to extract the sample data for each cluster, and uses the extracted data for the learning of the neural network. The purpose of the cluster is to make the data sample types at the input end of the neural network hierarchical, and the spare parts type coverage is complete, reducing prediction errors.

2.2.1. Spare Parts Classification Based on an Improved Clustering Algorithm. The present paper proposes a method that classifies repair spare parts and randomly extracts a certain amount of sample data according to the classification ratio to extract the sample data for each cluster; the training samples can cover various sample data. PCA data dimension reduction and k-means cluster analyses are the two essential components of the improved clustering algorithm. The general process of the improved clustering algorithm can be described as follows:

(I) PCA is used to perform dimension reduction on the data

(II) A clustering algorithm based on the Mahalanobis distance is used to classify the processed data

(III) The rationality of the classifications is evaluated using the DB and CH indices to check the number of clusters

Because the traditional k-means algorithm is unsuitable for nonconvex datasets or data with significant differences in class size and there is a specific correlation between influencing factors and the influence of dimension, the Mahalanobis distance was used for classification in the present paper. To provide a better understanding of the improved clustering method, PCA data processing analysis and the improved k-means cluster analysis are described in detail in the following sections.

(1) PCA Data Processing Analysis. Principal component analysis is mainly aimed at the same object when there are multiple influencing factors to affect it; there are some correlations between these influencing factors. In this case, the process became more and more complicated in subsequent research. In order to make the subsequent research more simple and consistent with the purpose of the research, it is hoped that one or several representative indicators can be found and the statistical methods can be independent of each other.

The main steps of principal component analysis are as follows:

(I) Standardize the data.

(II) Find the correlation matrix.

(III) Solve the eigenvalues and eigenvectors.

(IV) Find the principal component: select the number of principal components based on the principle that the factor contribution rate is greater than 85%.

(V) Calculate the factor score: when new factors after dimension reduction are determined, their scores can be calculated by calculating the specific values of these factors. These values can become factor
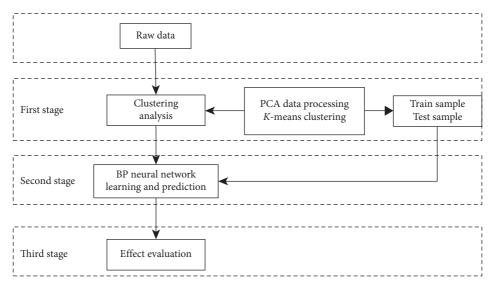
FIGURE 1: The general procedures of the evaluation model.

scores. These new variables are called factor variables. With these new variables, you can use them to replace the original variables, thereby reducing the dimensionality of the data to become more information intensive.

Matrix X processed by PCA dimension reduction is transformed into new comprehensive variables as follows:

$$\begin{cases} F_1 = \beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1p}x_p \\ F_2 = \beta_{21}x_1 + \beta_{22}x_2 + \cdots + \beta_{2p}x_p \\ \qquad \cdots \\ F_i = \beta_{i1}x_1 + \beta_{i2}x_2 + \cdots + \beta_{ip}x_p \end{cases}, \qquad (1)$$

where F_i is principal component 1, principal component 2, ... principal component p. Therefore, F_1 is called the first principal component and F_2 is the second principal component and until the i-th principal component. β_{ij} is the principal component coefficient.

(2) The Improved k-Means Algorithm for Classifying Spare Parts. The k-means algorithm [23] is a typical distance-based clustering algorithm. Distance is used to evaluate similarity. The smaller the distance between two objects, the higher the similarity.

The k-means cluster algorithm is used to classify data with different numbers of clusters so that they can be assigned to the group with the highest similarity. The classification method mainly uses the Mahalanobis distance to divide the clusters of each point. The traditional k-means algorithm often uses the continental distance, which depends on the magnitude scale to classify the sample data. The Mahalanobis distance considers the dimension and excludes the interference of the correlation between the sample data, which is not affected by the linear transformation. Compared to the continental distance, the Mahalanobis distance is used in the clustering algorithm to improve the accuracy of the cluster, and the distance can be calculated using the following equation:

$$D_{\text{mahananobis}}(r, G) = \sqrt{(r - u)^T C^{-1} (r - u)}, \qquad (2)$$

where D represents the Mahalanobis distance and G represents an overall sample space in the m-dimensional sample space; m is the dimension; n is the number of samples; r represents a sample point in the m-dimensional space; u represents the mean vector of the population sample; and C is the covariance of the population sample matrix.

The proper number of clusters can be determined based on the values of the DB and CH indices. The DB index is the average of the distances from all other sample points in the cluster, i.e., the intraclass distance; the smaller the DB index value, the better the clustering effect:

$$\text{DB} = \frac{1}{n}\sum_{i=1}^{n}\max_{j \neq i}\left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)}\right). \qquad (3)$$

The DB index is used to measure the tightness within the cluster by calculating the sum of the squares of the distances between the points in the cluster and the cluster center. σ represents the average of the distance from all other sample points in the cluster to the center of the c_i cluster, and d represents the distance between samples.

The separation of the dataset is measured by calculating the quadratic sum of the distances between the center point of this cluster and the center points of all the clusters. The CH index is the ratio of separation to tightness, as shown in the following equation:

$$\text{CH}(K) = \frac{B(K)(N - K)}{W(K)(K - 1)}. \qquad (4)$$

Here, N represents the number of training set samples and K represents the number of categories, where $W(K)$ refers to the separation degree and $B(K)$ refers to the tightness degree. The larger the CH index, the tighter the cluster. Thus, a lower CH index means that the clusters are more dispersed, and the classifications are better. Therefore, better data classifications should have lower DB and CH indices.

2.2.2. Prediction of Importance Degree Based on a Neural Network. The importance degree evaluation workload is also heavy when there are large volumes of spare parts data. In the current paper, a neural network was used to carry out a quantitative evaluation of the importance degree of spare parts. At the same time, in order to overcome the problem of strong subjectivity in the process of importance degree evaluation, this article used the actual data of the company to train and learn BPNN. By learning the actual importance value of enterprise spare parts, the training rules of neural networks can be more objectively trained.

BPNN training uses existing training samples to train network input and outputs the training results, which are ultimately verified by the test samples. Neural network prediction uses the trained network and the analysis of the prediction results to predict the output. The flowchart of the BPNN algorithm is shown in Figure 2.

First, we constructed a suitable BP neural network structure, mainly including the settings of the input layer, hidden layer, and output layer parameters. Second, we trained the BP neural network and continuously optimized the neural network. Finally, we used the trained BP neural network. We implemented neural network data prediction function on test data.

(1) Single-Layer BPNN Structure. The current paper mainly describes the use of a single-layer BPNN structure to train the importance degree, as shown in Figure 3. The input is new factors of impact importance degree evaluation after dimensionality reduction, and the output is the importance degree.

The operational steps are as follows:

Step 1: initialize the BPNN, determine the number of layer nodes, and set each weight and the initial value of the threshold to a relatively small random number.

Step 2: input the sample and the corresponding output, and learn from each training sample, i.e., perform steps 3 and 4 using the data for each sample.

Step 3: calculate the error of the output layer and the error of the hidden layer according to the expected output value in the sample and the output value obtained by the training.

Step 4: continuously perform the inverse correction to the weight and the threshold such that the output value from the training is infinitely close to the expected output from the sample;

Step 5: the output error value can be used to judge whether the error value converges within a given learning precision; if it is satisfied, learning ends; otherwise, the process returns to Step 2.

The learning rules of the neural network are gradually established by continuously learning the output from the sample data, and the ability to learn through continuous correction improves. PCA analysis can realize the dimensionality reduction of the data, making the data simpler, and the input of the influencing factors can be used to obtain the
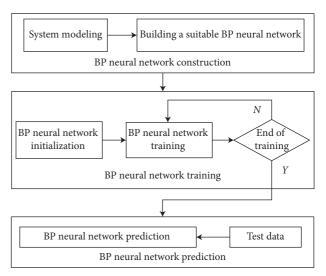

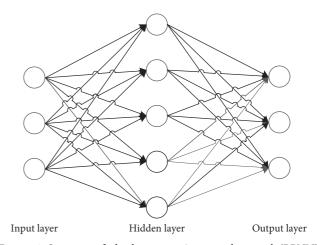FIGURE 2: Flowchart of the back-propagation neural network (BPNN) algorithm.

FIGURE 3: Structure of a back-propagation neural network (BPNN) with a single hidden layer.

importance value of the corresponding spare parts. Therefore, when a company introduces new spare parts by inputting the influencing factors with the established learning rules, it is possible to quickly evaluate the importance degree value of the spare parts. Using this model, it is possible to accurately and efficiently evaluate the importance degree of spare parts.

(2) Parameter Setting of BPNN. The BP network input layer node is h_i, which is the importance degree evaluation index.

(I) Determination of the number of hidden layer nodes [24]:

$$j = \sqrt{h + l} + a. \tag{5}$$

The h represents the number of input nodes, l represents the number of output nodes, and a represents the constant ranging from [1, 10]. The number of nodes in the hidden layer of the neural

network can be determined from the output and input nodes. The analysis can be used to determine the basic construction of the neural network structure.

(II) The transfer function of the hidden layer neurons uses the tan-sigmoid function [25], namely,

$$\tan \text{sig}(n) = \frac{2}{1 + e^{-2n}} - 1. \tag{6}$$

3. Examples and Results

The historical maintenance data and the inventory information for the maintenance of various spare parts are used to evaluate the importance degree of the spare parts, and the information is about heavy truck vehicles maintenance company.

3.1. Data Training

3.1.1. PCA Dimensionality Reduction.
There are many influential factors that affect the evaluation of the importance of spare parts. In order to alleviate the dimensional disaster problem, we performed PCA dimensionality reduction processing on the data of the influencing factors. On the one hand, it reduces the amount of data processing.

In the original data, there are eight factors that affect the importance degree evaluation. The dimensional information included the purchase lead time, the annual failure number, the management fee, the maintenance difficulty, the unit price, the service satisfaction, the working hours, the maintenance price for a unit hour, and the importance degree. Among them, the data on the importance evaluation factors mainly came from the historical maintenance data and inventory data information of the enterprise. The importance degree mainly comes from the actual value of the actual verification of the enterprise (see Appendix 1 in Supplementary Materials for details of the raw data).

After normalizing the original data, using SPSS to perform principal component analysis on these data, we can obtain the total variance explained by the original variables through these indicators, as shown in Table 1.

It can be seen from Table 1 that the extracted feature values are three principal components of 3.929, 2.551, and 1.033, respectively. The variance contribution rate of the first principal component is 49.127%, and the variance contribution rate of the second principal component is 29.391%. The variance contribution rate of the three principal components is 12.908%, and the cumulative percentage of the three factors to the initial eigenvalue reaches 91.426%. The three-component index score coefficient is shown in Table 2.

From this, new factors F_1, F_2, and F_3 are obtained, and we can get the data information of the new factor (for details, see Appendix 2 in Supplementary Materials):

$$F_1 = 0.314x_1 + 0.004x_2 + 0.111x_3 + 0.008x_4 + 0.107x_5 + 0.003x_6 + 0.313x_7 + 0.314x_8,$$

$$F_2 = -0.091x_1 + 0.36x_2 + 0.448x_3 + 0.451x_4 + 0.293x_5 - 0.01x_6 - 0.093x_7 - 0.09x_8, \tag{7}$$

$$F_3 = -0.036x_1 - 0.371x2 + 0.206x_3 + 0.035x4 + 0.085x_5 + 0.881x_6 - 0.035x_7 - 0.036x_8.$$

3.1.2. K-Means Sample Training.
After dimensionality reduction, the three-dimensional data for each sample comprised three factors: F_1, F_2, and F_3. The total number of data was 500. Figure 4 shows the three-dimensional coordinates of the sample data for 500 spare parts. The coordinate information comprises F_1, F_2, and F_3.

First, the samples were classified by using the improved k-means clustering method. Figure 5 shows the classification of 500 spare parts. It is clear from Figure 5 that the 500 samples are divided into three clusters. The first cluster of sample data consists of green cross, the second cluster consists of blue six-pointed stars, and the third cluster consists of red five-pointed stars. The black circles represent the centers of each cluster.

Secondly, cluster data and cluster centers C_1, C_2, and C_3 were obtained. The cluster center coordinates are shown in Table 3.

Finally, using CH and DB index values, the average number of clusters was found by averaging multiple times. The clustering effects of ten simulations when $k = 2$, $k = 3$, $k = 4$, and $k = 5$ are shown in Table 4.

In Table 4, the DB value represents the average of the distances from all other sample points in the cluster. The smaller the DB value, the smaller the intraclass distance; the larger the distance between classes, the better the clustering effect. The larger the CH value, the tighter the class. The average of 10 simulation results can be used to evaluate the optimal number of clusters, as shown in Table 5. The best clustering effect was obtained when $k = 3$. At this time, the points in the cluster were closest together, but the distance between the classes was larger.

3.2. BPNN Training and Evaluation.
By establishing a neural network model, using real data obtained from the enterprise, including historical spare parts maintenance information and inventory information, and the importance value of spare parts, the neural network is trained to achieve its accurate prediction function.

PCA analysis determined that the BPNN input comprised three nodes, which were the three factors that had the greatest impact on the importance degree evaluation,

TABLE 1: Total variance explained.

Ingredient	Initial eigenvalue			Extract load sum of squares		
	Sum	Variance (%)	Sum (%)	Sum	Variance (%)	Sum (%)
1	3.929	49.127	49.127	3.929	49.127	49.127
2	2.351	29.391	78.518	2.351	29.391	78.518
3	1.033	12.908	91.426	1.033	12.908	91.426
4	0.382	4.7755	96.2015			
5	0.209	2.6128	98.8143			
6	0.055	0.6875	99.5018			
7	0.038	0.4751	99.9769			
8	0.002	0.025	100.000			

TABLE 2: Component score coefficient matrix.

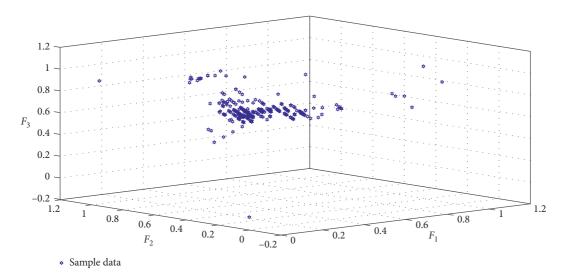| Ingredient	New factors		
	F_1	F_2	F_3
$x1$	0.314	−0.091	−0.036
$x2$	0.004	0.360	−0.371
$x3$	0.111	0.448	0.206
$x4$	0.008	0.451	0.035
$x5$	0.107	0.293	0.085
$x6$	0.003	−0.010	0.881
$x7$	0.313	−0.093	−0.035
$x8$	0.314	−0.090	−0.036

FIGURE 4: Distribution of raw data for the spare parts.

namely, F_1, F_2, and F_3, respectively, set to h_1, h_2, and h_3 (a total of three). Because the output is the importance degree of the spare parts, the l value was 1, and the number of hidden layer nodes was verified using equation (5) and continuous simulation.

Figure 6 is a comparison of the sample output values and the output errors after sample learning at various numbers of hidden nodes. There error is lowest, and the training effect was at its most efficient when the number of hidden layer nodes was 10. The transfer function of the hidden layer neurons uses equation (6). When the BPNN parameters have been determined, the entire BPNN can be constructed.

After K-means initially classifies the original data, the number of data in the first data cluster is 20, the number of data in the second data cluster is 80, and the number of data in the third data cluster is 400. We can find that the data ratio of each type of data cluster is $1:4:20$, and this ratio is also used in MATLAB to randomly extract data for training. In this paper, 250 data are randomly selected from the various types of spare parts classified by k-means according to the ratio of $1:4:20$ for neural network learning and training, and the rest of the data is test data; we used the training data to train the BPNN, and continuously updated the weight values to reduce the output
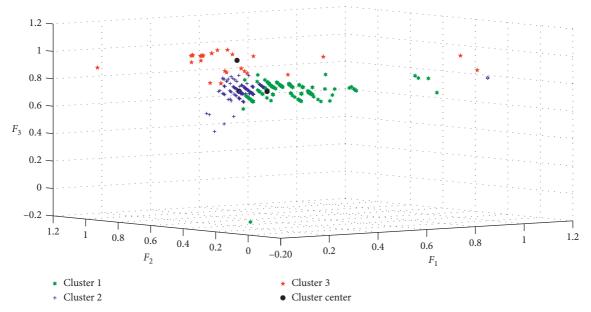
FIGURE 5: Distribution of the three clusters.

TABLE 3: Position information of the three cluster centers.

Factor	C_1	C_2	C_3
F_1	0.1223	0.0654	0.8247
F_2	0.0240	0.0894	0.8337
F_3	0.2834	0.4919	0.9852

TABLE 4: The Davies–Bouldin (DB) and Calinski–Harabasz (CH) indices of 10 simulations with different k values.

Serial number	$k = 2$		$k = 3$		$k = 4$		$k = 5$	
	DB	CH	DB	CH	DB	CH	DB	CH
1	1.6347	71.6465	0.6859	226.6673	1.1873	122.1868	1.2345	77.5856
2	1.3394	179.4807	0.8963	280.6704	1.1071	165.3401	1.1626	100.0095
3	1.2098	186.3162	0.8821	274.6304	1.2208	112.8648	1.3117	83.3012
4	1.3445	161.4082	1.0001	232.5309	1.2754	132.9948	1.1375	183.1147
5	1.1514	215.2821	0.8887	262.2657	1.2790	154.3030	1.2678	89.4779
6	1.3001	198.2829	0.7225	298.6832	1.1551	174.1512	1.2404	154.8071
7	1.2883	190.4038	0.9551	255.7667	1.1216	126.2768	1.0492	151.5377
8	1.2983	196.1386	1.0411	242.0565	1.1299	140.6781	1.2082	93.3379
9	1.3271	173.8368	0.7084	297.5069	1.0789	214.7179	1.1228	134.8111
10	1.4925	118.0285	0.7278	293.1378	1.1847	133.5798	1.0530	216.0843

TABLE 5: Average Davies–Bouldin (DB) and Calinski–Harabasz (CH) indices with various k values for 10 simulations.

k	DB	CH
2	1.33861	169.08243
3	0.8508	266.39158
4	1.17398	147.70933
5	1.17877	128.4067

error of the training data. In order to further verify that BPNN can achieve accurate predictions, we used the test data to verify the BPNN to prove the accuracy of the model.

In Figure 7, the abscissa indicates the number of spare parts, and the ordinate indicates the corresponding absolute error. The figure reveals that the data prediction error of the training sample and test sample was very close to 0.

The importance degree values predicted by factor screening and clustering are shown in Figure 8. The expected importance degree values of the samples are represented in blue, and the output importance degree values after learning through the neural network are represented in red. By comparing the fit between the output values and the expected values, it is possible to determine whether learning produces accurate results. The figure demonstrates that the fit between the values is quite satisfactory, and therefore the results of training and testing are accurate.
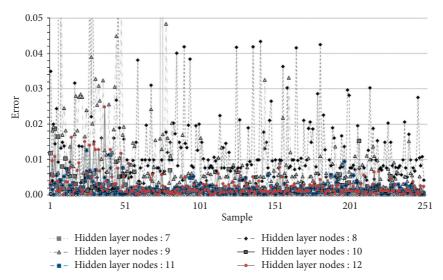
Figure 6: Prediction errors in situations with various hidden layer nodes.
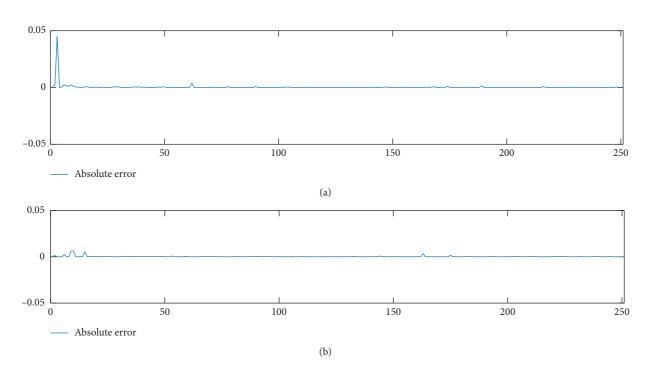

(a)

(b)

Figure 7: Absolute error of prediction based on the (a) training and (b) test samples.

To further determine the validity of the model, it was necessary to examine the unprocessed data in the neural network. The learning effect is shown in Figure 9, which demonstrates that the accuracy was poor, with regard to both the training sample and the test sample.

Figures 8 and 9 reveal an obvious difference between the predicted results obtained with or without model processing. The learning effect was more apparent in the predicted results following model processing than in the unprocessed results. The learning effect was greater and the results were more accurate, which verifies the validity of the model.

Figure 10 also reveals that during learning simulation process, the error value decreased continuously. When the training reached epoch 52, the error stabilized and there were no subsequent large fluctuations. Therefore, we could get 52 as the optimal number of iterations.

In summary, Figures 7–10 prove that the errors between the real values and the expected values arising from the proposed evaluation method were small and the prediction accuracy was high.

To verify the validity of the established importance degree evaluation model, the relative errors between the real importance degree values and the evaluation importance degree values are presented in Table 6 (see Appendix 3 in Supplementary Materials for all data). The table reveals that the importance degree values obtained by the two methods
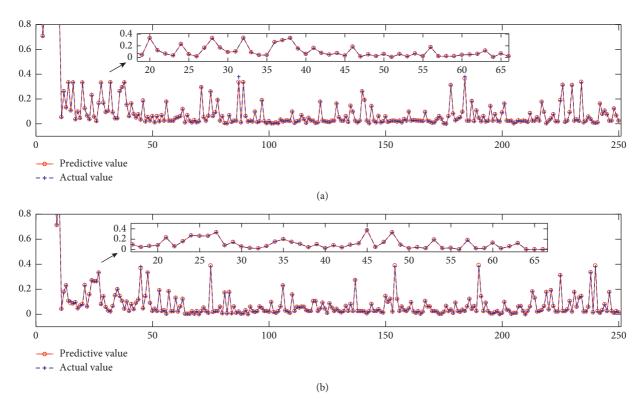
(a)

(b)

FIGURE 8: Results of a comparison between the actual values and the predicted values following data screening. (a) Training sample prediction. (b) Test sample prediction.
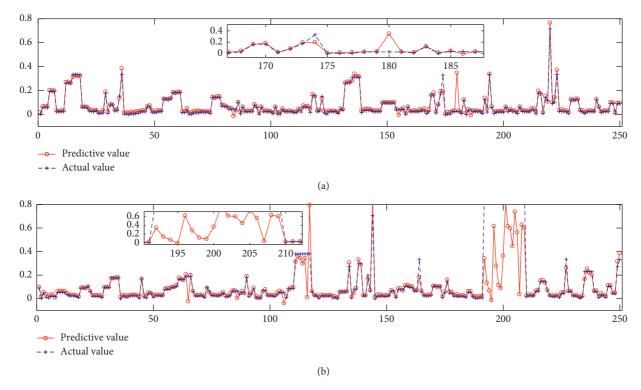

(a)

(b)

FIGURE 9: Results of a comparison between the actual values and the predictive values without data screening. (a) Training sample prediction. (b) Test sample prediction.
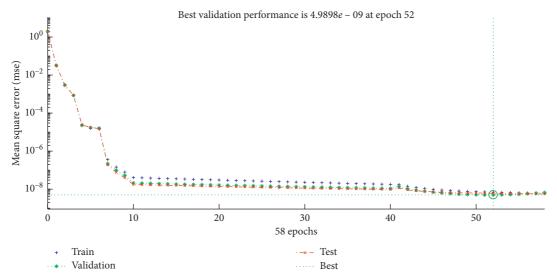
FIGURE 10: Comparison of mean square errors of the results of training, validation, and testing.

TABLE 6: Comparison of importance degree values obtained from the samples and from the proposed method.

Part number	Part number	Samples-based importance degree	The proposed method-based importance degree	Relative error (%)
1	A 881 316 05 52	0.05708794	0.05708713	0.0000809
2	A 881 350 00 06	0.056734144	0.056682938	0.0051206
3	A 881 316 04 52	0.05708794	0.057078051	0.0009889
4	A 520 541 05 01	0.192090071	0.192017238	0.0072833
5	A 881 500 03 55	0.191947037	0.191609507	0.0337530
6	A 518 820 25 75	0.191947037	0.191933056	0.0013981
7	A 881 331 02 28	0.261637417	0.261629909	0.0007507
8	A 881 331 01 28	0.261668054	0.261572242	0.0095811
9	A 520 541 05 02	0.261305535	0.261270967	0.0034569
10	A 518 750 00 18	0.333885659	0.333885496	0.0000164
11	A 518 671 02 10	0.334024112	0.333950642	0.0073469
12	A 881 820 04 61	0.333951472	0.333906344	0.0045128
13	A 881 820 03 61	0.333951472	0.333903023	0.0048449
14	A 881 698 04 75	0.060942867	0.060938979	0.0003888
15	A 881 820 02 56	0.060905256	0.06084617	0.0059087
16	A 881 820 01 56	0.060905256	0.060873403	0.0031853
17	A 518 891 02 17	0.035284462	0.035013288	0.0271174
18	A 520 541 05 03	0.17749964	0.177416624	0.0083016
19	A 881 331 02 29	0.080136	0.080128215	0.0007785
20	A 881 331 01 29	0.080145812	0.080124858	0.0020954
21	A 346 353 07 35	0.154651334	0.154597296	0.0054038
22	A 881 500 00 99	0.333385668	0.33337566	0.0010008
23	A 070 490 06 61	0.062971631	0.062939744	0.0031887
24	A 881 490 04 40	0.062854104	0.062802206	0.0051898
25	A 518 544 15 03	0.130196477	0.130188692	0.0007785
26	A 818 522 00 03	0.13028396	0.130272664	0.0011296
27	A 881 320 00 02-03	0.130572149	0.130458202	0.0113948
28	A 818 520 00 19	0.130225144	0.13015753	0.0067615
29	A 881 890 00 05	0.181999949	0.181893667	0.0106282
30	A 881 323 01 00	0.181868101	0.181867296	0.0000805
. . .	. . .	. . .	. . .	. . .

were very close, and the order of the importance degree of the spare parts was the same. Therefore, the established importance degree evaluation model can be used to evaluate the importance of spare parts and produces accurate results.

4. Conclusions

The present paper proposes an accurate importance degree evaluation method that combines an improved clustering

algorithm with BPNN. First, we used PCA for data dimensionality reduction, and the training sample set is then classified into different clusters by the improved k-means clustering algorithm. Finally, a single-layer BPNN with multiple inputs and a single output was established to evaluate the importance degree of the spare parts. After training the extract data, the BPNN automatically generated an evaluation of the importance degree of the spare parts from inputs to output. Compared to the importance degree results obtained from the sample maintenance spare parts, the evaluation results obtained from the proposed method are accurate, and to a certain extent, it overcomes the interference of subjective factors in the evaluation of importance. At the same time, it has the ability to evaluate the importance degree of spare parts for processing large amounts of data.

Notations

X_i:	Importance degree factors
x_i:	Normalized importance degree factors
β_{ij}:	Principal component coefficient
F_i:	Principal components
y:	Importance degree
$D_{\mathrm{mahalanobis}}$:	The mahalanobis distance
G:	An overall sample space in the m-dimensional sample space
m:	The dimension of a sample
n:	The number of samples
r:	A sample point in the m-dimensional space
u:	The mean vector of the population sample
C:	The covariance of the population sample matrix
σ_i:	The average of the distance from all other sample points in the cluster to the center of the c_i cluster
$d(c_i, c_j)$:	The distance between samples
N:	The number of training set samples
K:	The number of categories
$B(K)$:	The trace of the interclass dispersion matrix
$W(K)$:	The trace of the intraclass dispersion matrix
h:	The number of input nodes
l:	The number of output nodes
a:	The constant ranging from [1, 10].
j:	The number of hidden nodes.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Supplementary Materials

Appendix 1: the original data of the enterprise's spare parts information, where the data information mainly includes the factors that affect the importance assessment and the importance value of the spare parts. F_1, maintenance difficulty; F_2, annual failure number (pieces); F_3, unit price (yuan); F_4, purchase lead time (months); F_5, management fee (yuan); F_6, service satisfaction; F_7, working time (hours); F_8, price per hour (yuan/hour); y, importance degree. Appendix 2: the data after dimensionality reduction. The dimensionality reduction process uses the PCA method. The 8 influencing factors contained in the original data become 3 types of influencing factors through dimensionality reduction. Appendix 3: the specific data information of Table 6, the comparison of importance degree values obtained from the samples and proposed method. (*Supplementary Materials*)

References

[1] A. A. Aderoba, B. Kareem, and T. I. Ogedengbe, "A model for progressive inventory management for job–shop," *Nigerian Journal of Engineering Management*, vol. 4, no. 1, pp. 28–34, 2003.

[2] B. Kareem, T. I. Ogedengbe, and A. O. Jewo, "A condition based model for maintenance planning on critical equipment in the petrochemical industry," in *Proceedings of the 2010 International Conference on Test and Measurement (ICTM 2010)*, Phuket, Thailand, 2010.

[3] C. Teixeira, I. Lopes, and M. Figueiredo, "Multi-criteria classification for spare parts management: a case study," *Procedia Manufacturing*, vol. 11, pp. 1560–1567, 2017.

[4] Z. W. Birnbaum, *On the Importance of Different Components in a Multi-Component System*, Academic Press, New York, NY, USA, 1969.

[5] M. Catelani, L. Ciani, and M. Venzi, "Component reliability importance assessment on complex systems using credible improvement potential," *Microelectronics Reliability*, vol. 64, pp. 113–119, 2016.

[6] H. Dui, S. Si, and R. C. M. Yam, "Importance measures for optimal structure in linear consecutive-k-out-of-n systems," *Reliability Engineering & System Safety*, vol. 167, pp. 1–9, 2017.

[7] S. Si, G. Levitin, H. Dui, and S. Sun, "Importance analysis for reconfigurable systems," *Reliability Engineering & System Safety*, vol. 126, pp. 72–80, 2014.

[8] B. Natvig, "A suggestion of a new measure of importance of system components," *Stochastic Processes and their Applications*, vol. 9, no. 3, pp. 319–330, 1979.

[9] J. S. Hong and C. H. Lie, "Joint reliability-importance of two edges in an undirected network," *IEEE Transactions on Reliability*, vol. 42, no. 1, pp. 17–23, 1993.

[10] E. Borgonovo and G. E. Apostolakis, "A new importance measure for risk-informed decision making," *Reliability Engineering & System Safety*, vol. 72, no. 2, pp. 193–212, 2001.

[11] F. Y. Partovi, J. Burton, and J. Burton, "Using the analytic hierarchy process for ABC analysis," *International Journal of*

Operations & Production Management, vol. 13, no. 9, pp. 29–44, 1993.

[12] B. Kareem and A. S. Lawal, "Spare parts failure prediction of an automobile under criticality condition," *Engineering Failure Analysis*, vol. 56, pp. 69–79, 2015.

[13] G. Bektemur, N. Muzoğlu, M. A. Arıcı, and M. K. Karaaslan, "Cost analysis of medical device spare parts," *Pakistan Journal of Medical Sciences*, vol. 34, no. 2, pp. 472–477, 2018.

[14] A. Ishizaka, F. Lolli, E. Balugani, R. Cavallieri, and R. Gamberini, "DEASort: assigning items with data envelopment analysis in ABC classes," *International Journal of Production Economics*, vol. 199, pp. 7–15, 2018.

[15] R. M. Chandima Ratnayake and K. Antosz, "Risk-based maintenance analysis with fuzzy logic," *Procedia Engineering*, vol. 182, pp. 602–610, 2017.

[16] I. Roda, M. Macchi, L. Fumagalli, and P. Viveros, "On the classification of spare parts with a multi-criteria perspective," *IFAC Proceedings Volumes*, vol. 45, no. 31, pp. 19–24, 2012.

[17] B. Yousefi, S. Sfarra, C. Ibarra-Castanedo, N. P. Avdelidis, and X. P. V. Maldague, "Thermography data fusion and non-negative matrix factorization for the evaluation of cultural heritage objects and buildings," *Journal of Thermal Analysis and Calorimetry*, vol. 136, pp. 943–955, 2019.

[18] L. Wang, Y. Zeng, J. Zhang, W. Huang, and Y. Bao, "The criticality of spare parts evaluating model using artificial neural network approach," in *Proceedings of the 2006 International Conference on Computational Science*, Glasgow, UK, May 2006.

[19] Z. Cai, J. Jin, and Y. Chen, "Optimal inventory modeling of spare parts under the criticality," *Systems Engineering and Electronics*, vol. 39, no. 8, pp. 1765–1773, 2017, in Chinese.

[20] H.-H. Wu and Y.-N. Tsai, "An integrated approach of AHP and DEMATEL methods in evaluating the criteria of auto spare parts industry," *International Journal of Systems Science*, vol. 43, no. 11, pp. 2114–2124, 2012.

[21] Y. Huang, X. Guo, and D. Sun, "Criticality evaluation of materiel spare parts based on BP Neutral Network," *Aircraft Design*, vol. 31, no. 2, pp. 64–66, 2011, in Chinese.

[22] S. C. Yang and Z. W. Du, "Criticality evaluation for spare parts initial provisioning," in *Proceedings of the Annual Symposium Reliability and Maintainability, 2004-RAMS*, IEEE Xplore, Los Angeles, CA, USA, January 2004.

[23] J. Macqueen, "Some methods for classification and analysis of multi variate observations," in *Proceedings of the Berkeley Symposium on Mathematical Statistics & Probability*, Dayton, OH, USA, 1965.

[24] T. Lei, X.-H. Sun, X.-H. Guo, and J.-J. Ma, "Quantifying the relative importance of soil moisture, nitrogen, and temperature on the urea hydrolysis rate," *Soil Science and Plant Nutrition*, vol. 63, pp. 225–232, 2017.

[25] M. Chen, *Principles and Examples of MATLAB Neural Network*, Tsinghua University Press, Beijing, China, 2013, In Chinese.