Journal of Network Intelligence ©2023 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **8**, Number **2**, May **2023**

Improved Elman Neural Network Mechanical System Failure Prediction for Data Integration

Yang Shen*

College of Information Engineering
Guangzhou Institute of Technology
Guangzhou 510075, China
sy9301@163.com

Xiefei He

College of Information Engineering
Guangzhou Institute of Technology
Guangzhou 510075, China
hexiefei_2017@163.com

Lizhu Ye

Graduate School
Management and Science University
Shah Anam 40100, Malaysia
20180160@ gcc.edu.cn

*Corresponding author: Yang Shen
Received October 15, 2022, revised December 14, 2022, accepted February 19, 2023.

Abstract. *With the increasing performance and complexity of modern large mechanical systems, the traditional fault diagnosis techniques can no longer meet the actual operational needs of mechanical systems. Most of the data sets selected by traditional fault prediction methods come from a single information source, which can only reflect partial fault information and lead to large errors in prediction results. Therefore, this paper proposes a data integration-oriented fault prediction method for mechanical systems. Firstly, the full vector spectrum technology is used to effectively fuse multi-channel signals, so as to reflect the vibration characteristics of mechanical equipment more comprehensively and accurately. Secondly, the Elman neural network-based fault prediction model is built to address the characteristics of mechanical equipment systems, such as non-linearity, strong temporality and small amount of data, and the weight parameters of the Elman are optimized by using genetic algorithm (GA). Finally, the rolling bearing data set from the Bearing Data Center of CWRU was used as the experimental data, and the proposed fault prediction method was simulated using MATLAB software. The results of the experiment results show that the GA-Elman has a better prediction effect compared with the prediction effect of the Elman model. The prediction method using data integration technology is more effective than the prediction method using a single information source, and has a greater practical value in the field of mechanical system fault diagnosis.*
Keywords: Elman neural network, Fault prediction, Full vector spectrum technique, Data integration, Genetic algorithm

---

1. **Introduction.** With the increasing performance and complexity of modern large mechanical systems, the traditional fault diagnosis techniques have failed to meet the actual

operational needs of mechanical systems, resulting in serious accidents caused by mechanical equipment operational failures still occurring frequently in recent years. Most of the data sets selected by traditional fault prediction methods come from a single information source, which can only reflect part of the fault information, resulting in large errors in the prediction results [1,2,3,4]. How to predict the probability of failure of mechanical systems to achieve "predictive" maintenance of equipment has become the focus of research in the field of fault diagnosis. The study of failure prediction technology for mechanical equipment systems is of great value both in reducing economic losses and in extending the service life of equipment.

For critical equipment in heavy industry, the structure of most mechanical systems is very complex. In addition, the harsh operating environment of large mechanical systems leads to frequent failures, which can lead to downtime in minor cases or major safety accidents. Accurate prediction of the probability of mechanical equipment failure is fundamental to ensure the lasting operation of critical equipment and to improve equipment productivity [5,6,7,8]. Mechanical failure prediction technology is a key technology to achieve "predictive" maintenance of equipment and ensure the long-term safe operation of machinery and equipment, and is an emerging interdisciplinary discipline involving machinery, automation, computers, communications, and control. The core of mechanical fault prediction technology is to establish a fault prediction model that is consistent with the actual operating state and development trend of mechanical equipment. The failure prediction model can predict the future operation status of the system, thus providing a strong basis for equipment managers to correctly develop maintenance plans [9,10]. At present, the mainstream mechanical failure prediction models mainly include: time series prediction model, gray prediction model, machine learning, etc. However, for most large machinery and equipment systems usually have the characteristics of nonlinearity, strong time-series and small data volume [11], how to establish a fault prediction model with high prediction accuracy and real-time online monitoring will be the key problem to be solved in the study of fault prediction technology.

The aim of this study is to implement mechanical system fault prediction using a full vector spectrum technique and an improved Elman neural network to fuse the results of nonlinear, small-sample multi-source data in order to effectively improve the accuracy, reliability, and generalization of equipment fault prediction. Therefore, this paper proposes a data integration-oriented fault prediction method for mechanical systems. Meanwhile, the full vector spectrum technique and the improved Elman neural network model are used to realize the whole process of data integration. Compared with general prediction methods, the proposed method overcomes the limitation that a single information source can only obtain partial information of the object under test. By reasonably using sensor information, the utilization of information resources is maximized, thus improving the accuracy of machinery fault prediction and the reliability of prediction results. The prediction method using data integration technology is more effective than the prediction method using a single information source. In addition, the proposed method is more generalizable and can effectively predict the operating trends under different operating conditions. In other words, only the training data of the model need to be changed for different study subjects. The bearing fault prediction results show that the proposed method has good generalization and effectiveness. Therefore, the proposed method has great theoretical significance for the field of equipment failure prediction. At the same time, the proposed method can provide a strong basis for scientific management of mechanical equipment and reduce equipment management costs in practical engineering, which has certain engineering application value.

1.1. **Related Work.** In the late 1970s, with the rapid rise of nonlinear scientific theories such as artificial intelligence, fuzzy systems and gray systems, the study of fault prediction techniques gradually became one of the hot issues of interest to mechanical system experts and scholars. It was found that data-based fault prediction methods can be used in almost all situations with high generalization. At the same time, compared with knowledge- and model-based fault prediction methods, the accuracy of data-based fault prediction methods has certain absolute advantages [12,13]. By analyzing the historical operation data of the equipment and the change process of the current operation status, the data-based failure prediction method uses statistical and mathematical methods to predict the future status of the equipment. Equipment managers can refer to the prediction results to develop the most reasonable maintenance strategy.

As a typical representative of data-based fault prediction methods, artificial neural networks can accurately reflect the nonlinear correspondence between two or more things without the need to build a mathematical model. Therefore, artificial neural network models have been widely used in equipment fault prediction. Artificial neural networks have excellent ability to handle random and nonlinear relationships, and are very suitable for building data-based mechanical fault prediction models. According to the limitations of the Long Short-Term Memory (LSTM) neural network model and the data characteristics of aero engines, Li et al [14] proposed a fault prediction model based on parameter optimized LSTM neural network, which effectively improved the prediction accuracy. To address the limitations of a single wavelet threshold processing function, Huang et al [15] used an adaptive threshold function to reconstruct the signal and used a BP neural network for fault prediction. Wang et al [16] proposed a fault prediction method based on Principal Component Analysis (PCA) and BP neural network to reduce the influence of input sample correlation on BP neural network. Chen et al [17] established a fault prediction model combining Convolutional Neural Network (CNN) and LSTM, which effectively improved the accuracy of fault diagnosis and prediction of equipment. Xu et al [18] combined gray theory and artificial neural network, and proposed a fault prediction model based on gray neural network, which effectively solved the problem of small sample problem of low accuracy rate.

1.2. **Motivation and contribution.** However, most of these fault prediction methods only analyze the information from a single data source. Since faults usually need to be described by multiple data sources together, the above fault prediction methods cannot obtain complete fault information, which leads to easy misjudgment and omission in the process of fault prediction. In order to solve this problem, this paper uses data integration technology to fuse data from multiple sources, which can effectively ensure the integrity of information. The use of data integration technology in the field of fault prediction can not only improve the prediction accuracy to a certain extent, but also increase the confidence level of the prediction results. Since the hidden layer of Elman neural network is equipped with a takeover layer with a delay effect (which acquires the function of memory), it makes the whole network have the ability to adapt to time change. As a kind of neural network, Elman neural network has a better analysis ability for nonlinear and time-sensitive small sample data [19,20,21,22]. Therefore, a fault prediction model based on Elman neural network is built in this paper. To address the problem of low output stability due to randomly generated initial weights of Elman neural network, this paper uses genetic algorithm to optimize the initial weights of Elman neural network, so as to further improve the accuracy of fault prediction. Finally, the equipment failure prediction method based on data integration technology is applied to a specific research object of rolling bearings. The prediction results of different working conditions are analyzed to verify the effectiveness and generalization of the proposed method.

The main innovations and contributions of this study are shown below.

(1) In order to fuse multi-source data using data integration techniques, a feature extraction method based on the full vector spectrum technique is proposed, thus effectively ensuring the integrity of the information.

(2) To address the problem that the Elman neural network has low output stability due to randomly generated initial weights, a genetic algorithm is used to optimize the above parameters of the Elman neural network, so as to further improve the accuracy of fault prediction.

2. **Fault feature extraction based on full vector spectrum technique.** The operating environment of mechanical equipment is mostly complex, which makes the vibration signals collected by sensors very messy. It is difficult to get the characteristic information that can accurately describe the operation status of the equipment from the messy signals. Therefore, in equipment fault prediction, it is crucial to extract accurate and reliable fault characteristic information from the vibration signals collected by sensors.

Conventional signal processing methods make judgments by virtue of individual channel signals only, so the prediction results are unreliable. For condition monitoring of large machinery and equipment, two identical sets of sensors, arranged vertically in the same plane, are generally used to obtain the same vibration signal. However, there are some differences between the characteristic information collected by the sensors at two different locations. The full vector spectrum technique [23,24] is able to integrate the homogenous signals acquired by two sensors vertically arranged in the same plane, which in turn enables more complete information. The complete information helps to accurately describe the equipment operating conditions, thus gaining a greater advantage in fault prediction. In this study, we use full vector spectrum technique to extract features. The accuracy and effectiveness of the proposed method is verified using multiple sources of fault data collected under multiple operating conditions using rolling bearing datasets.

The path of motion of a rotating mechanism in the same cross section during the operation of a machine is considered as an ellipse (same harmonic frequency). However, the path of motion is different for different harmonic frequencies. The vibration signals measured in multiple directions of the same cross-section are necessarily different and also inseparably linked. Since the motion paths in different directions are not the same, if multiple sensors of the same type are used and arranged in different directions of the same cross-section for signal acquisition, the structure of the obtained spectrum is bound to differ as well. If only one sensor is used in the process of equipment fault prediction, it will inevitably lead to misjudgment and omission.

In order to identify mechanical equipment faults more precisely, two sensor information arranged vertically in the same cross section must be considered simultaneously. The acquisition of new information is a comprehensive process of extracting the original information, not a simple superimposed combination of all original information. The rotor system of mechanical equipment generates eddy currents at different harmonic frequencies, and the full vector spectrum technique is used to discern faults by the intensity of eddy currents at different harmonic frequencies [25].

In the full vector spectrum technique, the path of a rotating mechanism operating at a single harmonic frequency is represented as an ellipse. First, we need to define the relevant geometric parameters involved in the elliptical path. R_L denotes the primary vibration variable in the harmonic, which is geometrically represented as the long half-axis of the path. R_S denotes the secondary vibration variable in the harmonic, which is geometrically represented as the short half-axis of the path. α denotes the angle between the long half-axis and the x-axis.

Suppose $\{x_n\}$ and $\{y_n\}(n = 0, 1, 2, ..., N-1)$ are discrete sequences in the x and y directions, respectively. To simplify the computational steps, the set of these two discrete sequences is made into a plural sequence $\{z_n\} = \{x_n\} + \{y_n\}, (n = 0, 1, 2, \ldots, N-1)$. The Fourier transform of the complex sequence $\{z_n\}$ yields $\{Z_k\} = \{Z_{Rk}\} + \{Z_{Ik}\}(k = 0, 1, 2, ..., N-1)$. $\{Z_{Rk}\}$and $\{Z_{Ik}\}$ are the real and imaginary part sequences of $\{Z_k\}$, respectively. Therefore, we simply compute the other geometric parameters

$$\begin{cases} R_{Lk} = X_{pk} + X_{rk} = \frac{1}{2N}\left[|Z_k| + |Z_{N-k}|\right] \\ R_{Sk} = X_{pk} - X_{rk} = \frac{1}{2N}\left[|Z_k| + |Z_{N-k}|\right] \\ \tan\varphi_{pk} = \frac{Z_{Ik}}{Z_{Rk}} = \tan\varphi_{\alpha k} \\ \tan\varphi_{rk} = -\frac{Z_{I(N-k)}}{Z_{R(N-k)}} \\ \alpha_k = \frac{\varphi_{pk}+\varphi_{rk}}{2} \end{cases}, k = 0, 1, 2, \ldots, \frac{N}{2} - 1 \tag{1}$$

Theoretically, at least two sensors need to be used for signal acquisition in order to ensure comprehensive information. The full loss-of-spectrum technique is used to fuse the two-channel signals to obtain a new basis for fault identification.

3. Fault prediction modeling based on GA-Elman.

Among the fault diagnosis techniques, the fault prediction method is an advanced equipment maintenance tool. Fault prediction can predict the future operation of the equipment, thus providing a scientific basis for future maintenance. Under the premise of ensuring lasting operation of equipment, fault prediction can minimize the number of equipment maintenance, reduce equipment maintenance costs, and reduce downtime.

At present, many effective fault prediction methods have been proposed by numerous experts and scholars. Artificial neural networks can accurately reflect the nonlinear correspondence between two or more things without the need to build a mathematical model. Therefore, the artificial neural network can better reflect the dynamic trend of system operation. Color. Since the hidden layer of Elman neural network is set up with a takeover layer with delay effect (the function of memory is acquired), which makes the whole network have the ability to adapt to time change. Therefore, the Elman neural network is used as the main body of the prediction model in this paper. For the problems of Elman neural network such as low fitting degree and easy to fall into local optimum, we use genetic algorithm to improve it.

3.1. Elman Neural Network.

Since the 21st century, artificial neural networks have had more successful applications in big data, image processing and recognition, and prediction and classification in related fields. The processing of an artificial neural network (ANN) is a nonlinear system fitting, and the ANN has a structure and processing order similar to that of the brain. The model of an ANN is shown in Figure 1.

$$X = \sum_{i=1}^{n} w_i x_i \tag{2}$$

Where X is the input to the ANN, x is a one-dimensional column vector, and w is the strength (weight) of the input's action on the neuron

$$O_j = f(X - \theta) \tag{3}$$

Where O is the ANN output, θ denotes the threshold between neurons, and f denotes the activation function. The range of ANN output is limited by the type of activation

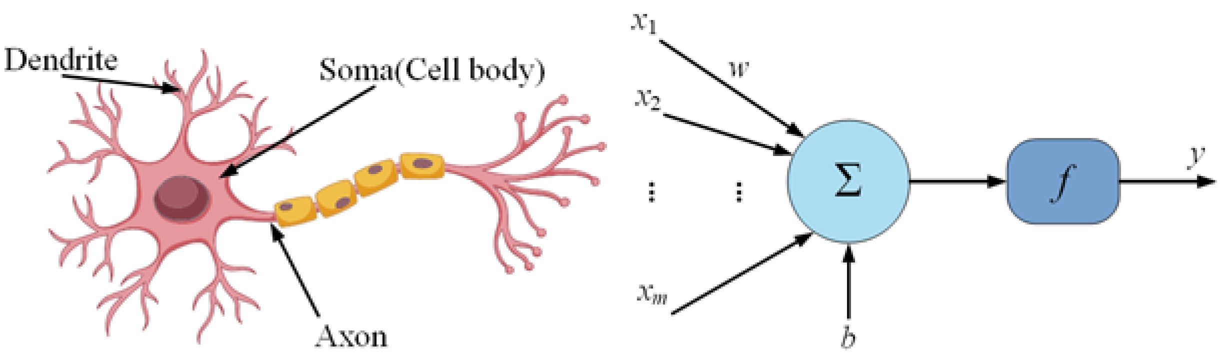


Figure 1. Model of ANN.

function(Sigmoid). The principle of the Sigmoid function is shown in Figure 2.

$$f(z) = \frac{1}{1+e^{-t}} \tag{4}$$

However, the original ANN model lacks the ability to handle dynamic information and

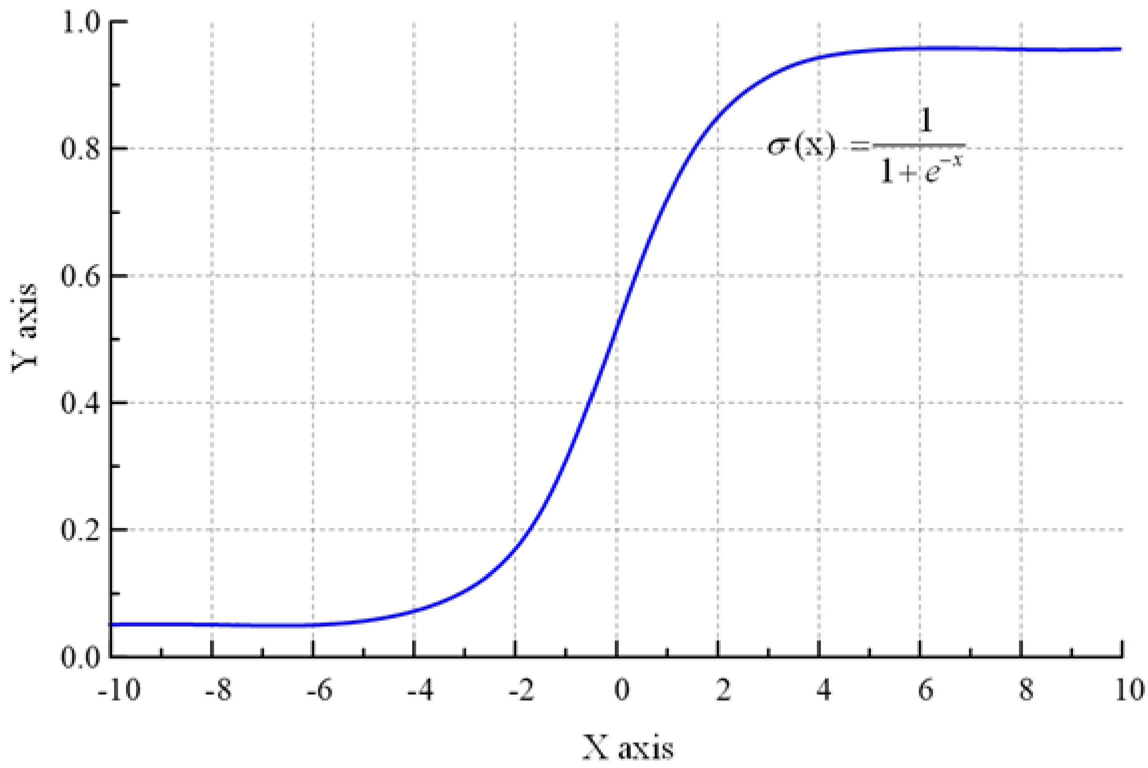


Figure 2. Principle of Sigmoid function.

is therefore weak for nonlinear, time-series stronger signals. Elman is a typical dynamic recurrent neural network that propagates in the same way as many feed-forward neural networks [26]. Compared with BP neural network, Elman has set a takeover layer with delay effect on the hidden layer. Due to its memory function, the whole Elman network has acquired the ability to adapt to time variation, so it has a strong ability to deal with timing information.

Elman neural network is a feedback neural network model divided into four layers: input layer, hidden layer, take-up layer and output layer. After training, the output values of the hidden layer are partially fed back to the units in the upper and lower layers and kept until the next training moment. This approach makes the Elman neural network sensitive to data from historical states.

The structure of the Elman neural network is shown in Figure 3. The input vector is the r-dimensional vector $x, x = [x_1, x_2, ..., x_r]$. The output vector of the implicit layer is the n-dimensional vector u, $u = [u_1, u_2, ..., u_n]$. The output vector is the m-dimensional vector $y, y = [y_1, y_2, ..., y_m]$. The output vector of the undertaking layer is an n-dimensional $x_c, x_c = [x_{c1}, x_{c2}, ..., x_{cn}]$. $w(i,k), w(k,j)$ and $w(k,j)$ are all weight matrices between different layers.$f()$ and $g()$ are the activation functions of the hidden layer and the output layer, respectively, and $h()$ is the takeover layer activation function.

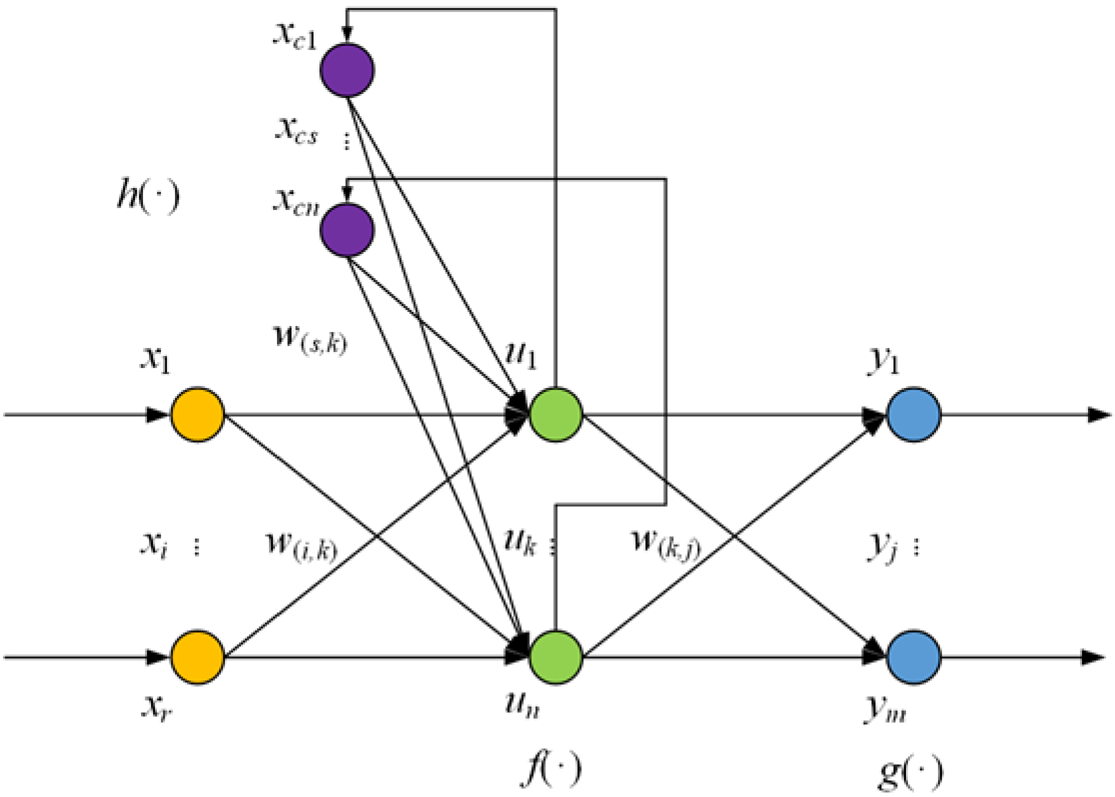


Figure 3. Structure of Elman neural network.

$$y(t) = g\left(f(t)w_{(k,j)}\right) \tag{5}$$

$$u(t) = f\left(x(t)w_{(i,k)} + x_c(t-1)w_{(s,k)}\right) \tag{6}$$

$$x_c = h(u(t-1)) \tag{7}$$

First, the weights of the nodes in each layer are initialized. Then, the training data is input and the input and output values of each layer are calculated. In the learning process of Elman neural network model, the output of the previous round of the implicit layer needs to be fed back to the takeover layer. The data processed by the takeover layer is input to the implicit layer together with the input layer data. Finally, the error is calculated based on the results of the output layer and the error function. If the size of the error meets the requirements or the training times reach the maximum, the training is stopped, otherwise the weights are updated and the next round of training is entered. The error function is E.

$$E(t) = \frac{1}{2}(y(t) - y_a(t))^{\mathrm{T}}\left(y(t) - y_a(t)\right) \tag{8}$$

Where $y_a(t)$ is the standard actual output data and $y(t)$ is the model output data. The correction parameter is used to perform the calculation of the weights according to the error back propagation algorithm. In general, the optimal solution of the weights w is found by the learning algorithm of gradient descent.

$$\Delta w_{(k,j)} = \eta_3 \delta_j u_k(t) \tag{9}$$

$$\Delta w_{(i,k)} = \eta_2 \delta_k x_{cs}(t) \tag{10}$$

$$\Delta w_{(s,k)} = \eta_1 \sum_{k=1}^{n} \left(\delta_j w_{(k,j)}\right) \frac{\partial x_{cs}(t)}{\partial w_{(s,k)}} \tag{11}$$

$$\delta_j = \left(y_j(t) - y_{aj}(t)\right) g'_j\left(\cdot\right) \tag{12}$$

$$\delta_k = \sum_{j=1}^{m} \left(\delta_j w_{(k,j)}\right) f'_k\left(\cdot\right) \tag{13}$$

$$\frac{\partial x_{cs}(t)}{\partial w_{(s,k)}} = f'_k(.) x_{cs}(t-1) + \alpha \frac{\partial x_{cs}(t-1)}{\partial w_{(s,k)}} \tag{14}$$

Where, η_1 ,η_2 , and η_3 are the learning rates of $w(i,k), w(k,j)$, and $w(k,j)$, respectively. δ_j is the gradient term of the neurons in the output layer, δ_k is the gradient term of the neurons in the hidden layer. $g'_j\left(\cdot\right)$ is the derivative of the output layer, and $f'_k(.)$ is the derivative of the hidden layer.

In the learning process of Elman network, the features of the data are extracted through a series of nonlinear mappings. Using the acquired features, the weights of the neurons are updated uninterruptedly after initializing the parameters in order to get the output value as close to the real value as possible. When the error between the output value and the actual value is greater than the set error, the Elman network uses back propagation to provide feedback from the output layer in order to update the weights among the neurons in each layer.

3.2. **Genetic Algorithm.** Genetic algorithm (GA) is an optimization algorithm based on genetic and mutation mechanisms [27], which is essentially an adaptive probabilistic optimization algorithm. The process of GA is similar to the process of artificial population evolution.GA generates new subpopulations through operations such as selection, crossover, and mutation to select well-adapted individuals. The basic process of GA is shown in Figure 4.

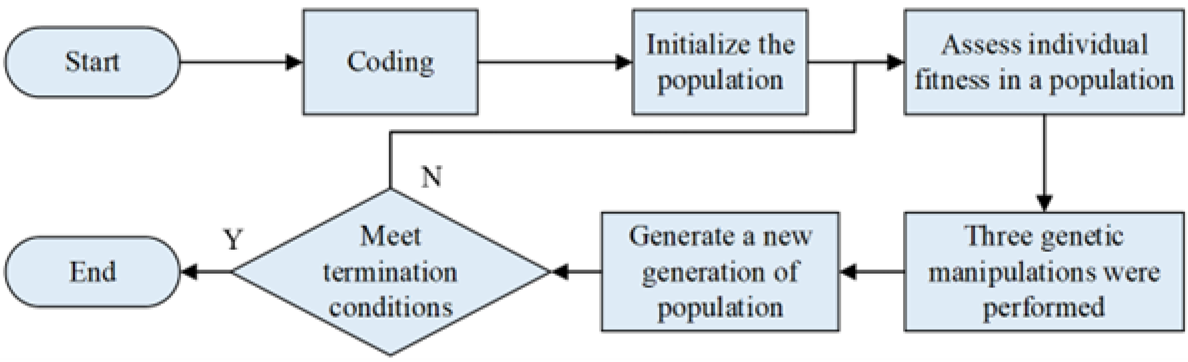


Figure 4. Basic flow of genetic algorithm.

3.3. **Genetic algorithm to optimize the parameters of Elman.** Unlike gradient descent algorithms, GA is good at handling discontinuous objective value optimization [28].GA uses fitness functions to judge the degree of individual merit. Choosing the right function can improve the speed of the merit search. For Elman neural network, the GA needs to find the merit is the initial weights and thresholds. Therefore, the designed fitness function F is shown as follows.

$$F = k\left(\sum_{i=1}^{L} abs\left(y_i - o_i\right)\right) \tag{15}$$

Where L is the number of individual samples, y_i is the expected output of individual i in the Elman neural network. Here a lower fitness means a better individual. When the optimal individual is used as the parameter of the initial model, the accuracy of the model will be improved. O_i is the predicted output of individual i.

There are many methods of selection operations in genetic algorithms, and the commonly used one is the roulette method. According to the proportional selection strategy of fitness, we need to calculate selection probability P_i of individual i.

$$f_i = k/F_1 \tag{16}$$

$$p_i = \frac{f_i}{\sum\limits_{j=1}^{N} f_i} \tag{17}$$

Where F_i is the fitness value. For network training, Gray's code is used for each individual. The crossover operation is performed at position k. N is the number of populations. The coded values corresponding to two consecutive numbers differ by only one code bit.

$$f_i = k/F_1 \tag{18}$$

$$p_i = \frac{f_i}{\sum\limits_{j=1}^{N} f_i} \tag{19}$$

In this paper, Gray code is used for encoding, which can effectively improve the local search ability of GA and avoid too much change after the occurrence of variation, leading to the situation of far from the optimal solution. Finally, the j-th gene a_{ij} of individual i was selected for the mutation operation. The predictive ability of the Elman neural network model is mainly influenced by the weights from the input layer to the hidden layer. Therefore, in this paper, GA is used to select the optimal parameters of the Elman neural network model, and the specific process is shown in Figure 5.

3.4. **GA-Elman based fault prediction process.** The collected data are divided into a training set and a test set. The parameters of the Elman model are optimized using the training set. The optimized Elmam model is used as the final prediction model. The feasibility of the model is verified using the test set, and the process is as follows.

Step 1: The original signal will be pre-processed, such as normalization, so as to ensure the validity of the prediction results.

Step 2: Feature extraction of the raw data using full vector spectrum technique. The feature data are arranged in the order of frequency bands to obtain a set of feature vectors for describing the device state. The feature vector data are randomly sorted. The top 90% of the data are selected in order to form the training set, and the remaining data are formed into the test set.

Step 3: Input training data, so that the Elman model can fully learn each fault feature, and perform parameter optimization of the Elman model by GA.

Step 4: judge whether the target accuracy is reached, if it is, then save the GA-Elman neural network that has finished learning; if it does not reach the target accuracy, then return to Step 3, reset the parameters and make the GA-Elman neural network learn again until it reaches the target accuracy.
Step 5: Build the final GA-Elman prediction model with optimal parameters.
Step 6: Use the test set data as input data, apply the final GA-Elman neural network model to make predictions, and save the prediction results.

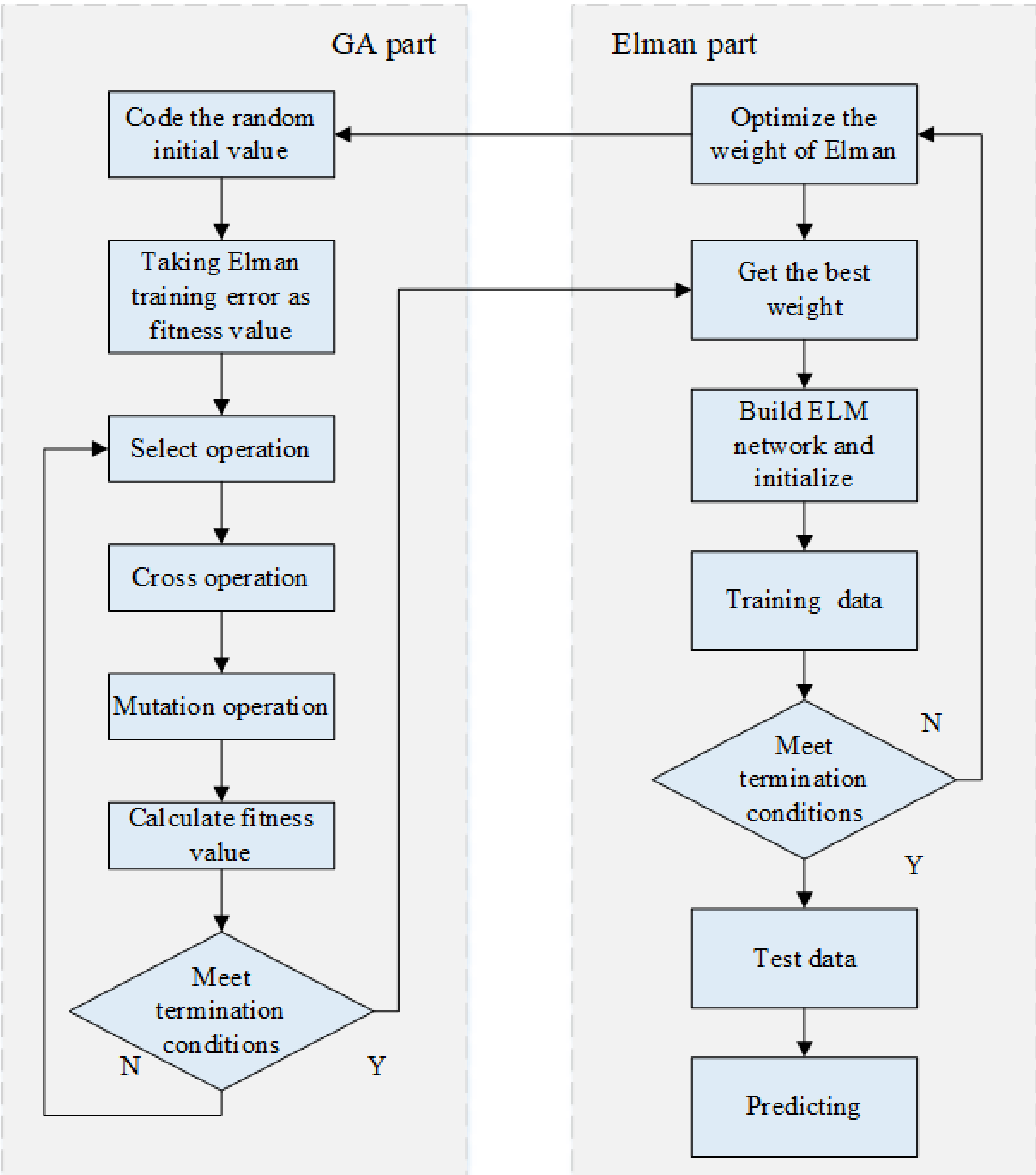


Figure 5. Flow of GA-Elman neural network.

4. Experimental results and analysis.

4.1. **Experimental data set.** The test bench [29] used by CWRU is shown in Figure 6. The bearing type used for the test is 6205-2RS JEM SKF. The data set consisted of normal and fault data. A total of nearly 60 million data were selected for this study. All the selected data were normalized and pre-processed before feature extraction. In the

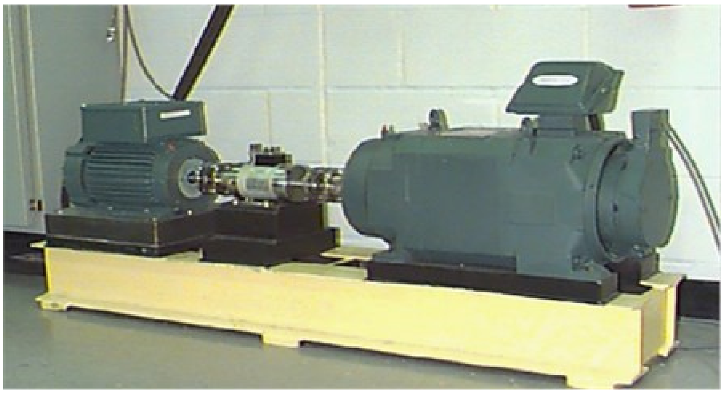


Figure 6. Rolling bearing test bench.

MATLAB programming environment, the Elman neural network was trained using the training set, and the parameters of the Elman neural network were optimized by GA. Mean Absolute Percentage Error (MAPE) is used as the fitness function. The number of hidden layer neurons is 160. The relevant parameters of GA were set as shown in Table 1. Figure 7 gives the relationship curve between the number of genetic generations M and the

Table 1. Relevant parameters of GA.

Parameters	Numerical value
Number of Individuals N	20
Maximum number of genetic generations M	30
Generation G	0.95
The crossover probability P1	0.7
The probability of variation P2	0.01

fitness function in the GA search process. As the number of genetic generations' increases, the fitness gradually decreases, that is, the error between the predicted and actual values gradually decreases, which indicates that the selection of parameters becomes more and more superior. The fitness reaches stability when M is approximately equal to 21.

4.2. **Fault feature extraction results.** The vibration signals of the driving end of the rolling bearing were acquired at a sampling frequency of 12 kHz, a fault diameter of 0.1778 mm, and a motor load of 1 HP. For feature extraction using the full vector spectrum technique, for the two-channel signal, the X-channel signal X(t) was acquired in the 3 o'clock direction, while the Y-channel signal Y(t) was acquired in the 6 o'clock direction. The histograms of the energy occupancy ratios for normal and three fault states are plotted as shown in Figure 8. It can be seen that the full vector spectrum technique can effectively extract the feature vectors under different operating conditions. The energy distribution of each frequency band can then be used to effectively identify the rolling bearing fault type.

4.3. **Analysis of prediction results for different working conditions.** Using the proposed method, the fault data collected under different working conditions are predicted separately and the prediction accuracy of the three fault states (inner ring, rolling body and outer ring fault) is tallied. The parameters of various working conditions are shown in

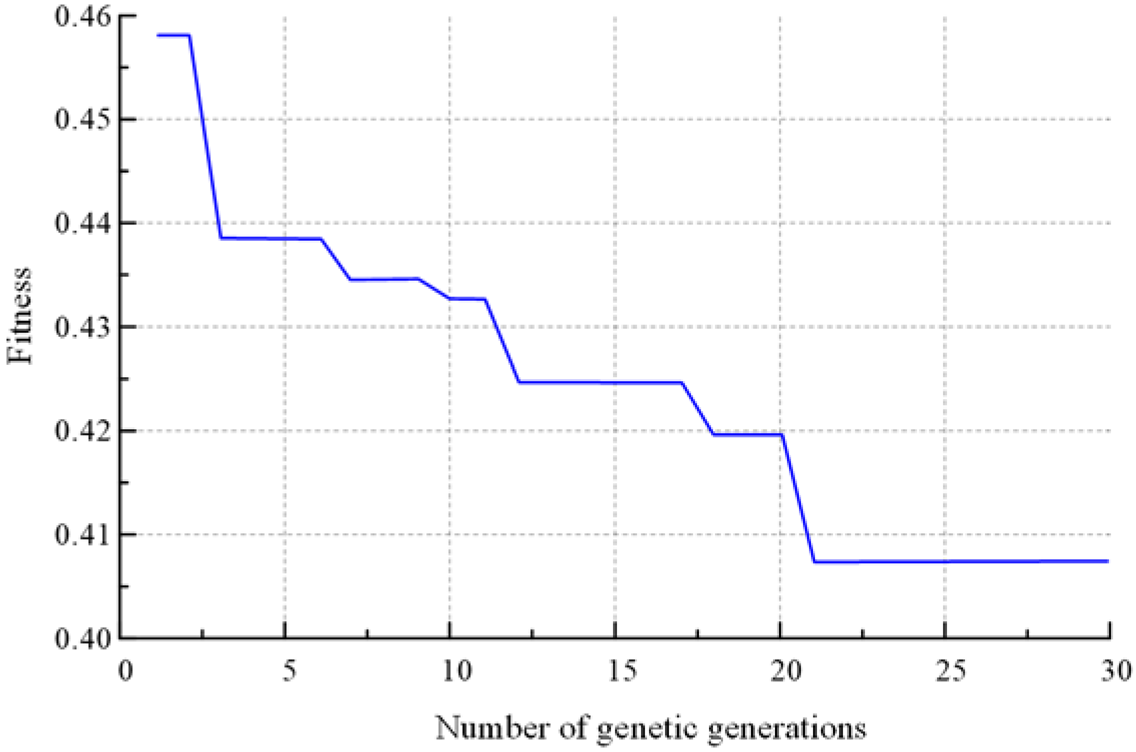
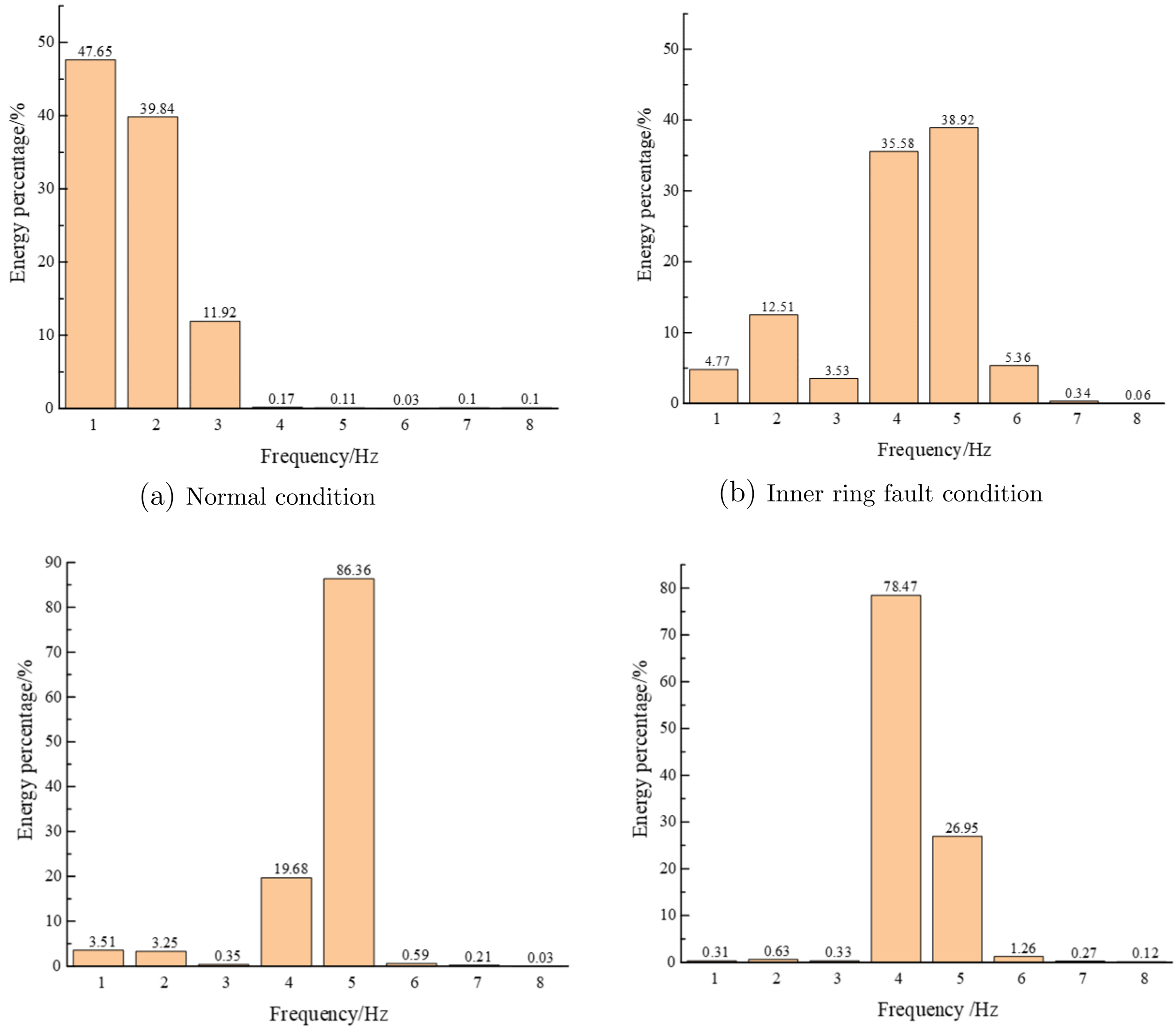


Figure 7. Optimization search process of genetic algorithm

Table 2. Various working conditions parameters.

Type of working condition	Sampling frequency/kHz	Fault diameter/mm	Motor load/HP	Motor approximate speed/r·min-1
1	12	0.1778	0	1797
2	12	0.1778	1	1772
3	12	0.1778	2	1750
4	12	0.1778	3	1730
3	12	0.3556	0	1797
6	12	0.3556	1	1772
7	12	0.3556	2	1750
8	12	0.3556	3	1730
9	12	0.5334	0	1797
10	12	0.5334	1	1772
11	12	0.5334	2	1750
12	12	0.5334	3	1730
13	48	0.1778	0	1797
14	48	0.1778	1	1772
15	48	0.1778	2	1750
16	48	0.1778	3	1730
17	48	0.3556	0	1797
18	48	0.3556	1	1772
19	48	0.3556	2	1750
20	48	0.3556	3	1730

Table 2. The failures of inner ring, rolling body and outer ring under different operating conditions were tested using the proposed fault prediction method. The overall average prediction accuracies of 96.45%, 95.67% and 96.63% were obtained for inner ring, rolling body and outer ring failures, respectively. The average prediction accuracy for all fault types under all operating conditions was as high as 96.25%. It can be seen that both the single fault and the overall average prediction accuracy are relatively satisfactory. If



(a) Normal condition

(b) Inner ring fault condition

(c) Rolling body fault condition

(d) Outer ring fault condition

Figure 8. Histogram of the energy share of rolling bearings for different failure states

the data integration based on the full vector spectrum technique is not used, but the prediction method using a single information source, the average prediction accuracy of all fault types under all working conditions is only 89.37%. Therefore, the prediction method using the data integration technique is more effective than the prediction method using a single information source.

Therefore, the proposed fault prediction method can effectively predict faults under different operating conditions with strong generalization. The main reason is that the feature extraction part uses full vector spectrum technology to fuse the data collected by sensors, and the fault prediction part uses GA-Elman neural network to make integrated judgment of state features.

4.4. **Experimental comparison analysis.** In addition, fault prediction models such as GA-BPNN[30], Elman,GA-Elman,GA-SVM[31], and PSO-SVM[32] were compared on the same dataset. The performance comparison results of five mechanical failure prediction models are shown in Table 3. It can be seen that the GA-Elman model has better

Table 3. Various working conditions parameters.

Predictive Models	RMSE	Accuracy	MAPE	Time/sec
GA-BPNN	5.4415	0.93847	0.10891	66
Elman	6.1912	0.92035	0.1077	30
GA-Elman	4.0312	0.96253	0.074551	38
GA-SVM	5.1008	0.94593	0.087885	90
PSO-SVM	5.0977	0.946	0.087871	103

prediction results compared to the prediction results of the Elman model. In the ANN-based prediction model, the RMSE of GA-Elman is significantly lower compared to GA-BPNN and Elman, which is only 4.0312. The MAPE of GA-Elman is also significantly lower compared to the improved SVM-based prediction model, which is only 0.074551. Compared to the Elman model, the running time of GA-Elman model increased, but still significantly smaller than other prediction models. Therefore, on the whole, the GA-Elman-based fault prediction model can achieve high prediction accuracy and running efficiency.

5. **Conclusion.** In this paper, we use data integration technology to fuse data from multiple sources, which can effectively ensure the integrity of information. Using data integration techniques in the field of fault prediction can not only improve the prediction accuracy to a certain extent, but also increase the confidence level of the prediction results. In this paper, a fault prediction model based on Elman neural network is built. To address the problem of low output stability due to the randomly generated initial weights of Elman neural network, this paper uses genetic algorithm to optimize the initial weights of Elman neural network, so as to further improve the accuracy of fault prediction. Finally, the equipment fault prediction method based on data integration technology is applied to the rolling bearing as a specific research object. The prediction results of different working conditions are analyzed to verify the effectiveness and generalization of the proposed method. The average prediction accuracy of all fault types under all working conditions is up to 96.25%. The GA-Elman-based fault prediction model can achieve high prediction accuracy and operational efficiency. The proposed mechanical system fault prediction method is of great value both in reducing economic losses and in extending equipment life.

Acknowledgements. The research is supported by: Tertiary Education Scientific research project of Guangzhou Municipal Education Bureau (202235310); Characteristic innovation projects of colleges and universities in Guangdong Province in 2021 (2021KTSCX268); Characteristic innovation projects of colleges and universities in Guangdong Province in 2021 (2021KTSCX265);

Author introduction. Shen Yang, Born in 1978, PhD. His research interests focus on big data analysis and software engineering; He Xiefei, Born in 1981, M.S., Her main research interests are computer software technology and big data

REFERENCES

[1] R.-N. Liu, B. Yang, E. Zio, and X.-F. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mechanical Systems and Signal Processing*, vol. 108, no. 2, pp. 33-47, Aug. 2018.

[2] S.-M. Zhang, X. Su, X.-H. Jiang, M.-L. Chen, and T.-Y. Wu, "A traffic prediction method of bicycle-sharing based on long and short term memory network," *Journal of Network Intelligence*, vol. 4, no. 2, pp. 17-29, 2019.

[3] H. Chen, D. L. Fan, L. Fang, and W. Huang, "Particle Swarm Optimization Algorithm with Mutation Operator for Particle Filter Noise Reduction in Mechanical Fault Diagnosis," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 10, p. 2058012, Feb. 2020.

[4] Y.-G. Ma, Y.-J. Peng, T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121-2133, 2022.

[5] J. Huo, D. Lin, and W. Qi, "Intelligent fault diagnosis method of mechanical equipment based on fuzzy pattern recognition," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 3657–3664, Apr. 2020.

[6] L.-L. Kang, R.-S. Chen, Y. –C. Chen, C.-C. Wang, X.-G. Li, and T.-Y. Wu, "Using Cache Optimization Method to Reduce Network Traffic in Communication Systems Based on Cloud Computing," *IEEE Access*, vol. 7, pp. 124397-124409, 2019.

[7] Z. Chen, K. Gryllias, and W. Li, "Mechanical fault diagnosis using Convolutional Neural Networks and Extreme Learning Machine," *Mechanical Systems and Signal Processing*, vol. 133, no. 6, p. 106272, Nov. 2019.

[8] H. Cao, X. Wang, D. He, and X. Chen, "An improvement of time-reassigned synchrosqueezing transform algorithm and its application in mechanical fault diagnosis," *Measurement*, vol. 155, no. 10, p. 107538, Apr. 2020.

[9] Z.-T. Zhang, H. Chen, S.-H. Li, and Z. An, "Sparse filtering based domain adaptation for mechanical fault diagnosis," *Neurocomputing*, vol. 393, no. 8, pp. 101–111, Jun. 2020.

[10] S. Kumar, A. Damaraju, A. Kumar, S. Kumari, and C.-M. Chen, "LSTM Network for Transportation Mode Detection," *Journal of Internet Technology*, vol. 22, no. 4, pp. 891-902, 2021.

[11] A. Daniar, Z. Nasiri-Gheidari, and F. Tootoonchian, "Position error calculation of linear resolver under mechanical fault conditions," *IET Science, Measurement & Technology*, vol. 11, no. 7, pp. 948–954, Oct. 2017.

[12] R. K. Patel and V. K. Giri, "Feature selection and classification of mechanical fault of an induction motor using random forest classifier," *Perspectives in Science*, vol. 8, no. 12, pp. 334–337, Sep. 2016.

[13] M. L. Masmoudi, E. Etien, S. Moreau, and A. Sakout, "Amplification of Single Mechanical Fault Signatures Using Full Adaptive PMSM Observer," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 1, pp. 615–623, Jan. 2017.

[14] Z. Li, J. Li, Y. Wang, and K. Wang, "A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment," *The International Journal of Advanced Manufacturing Technology*, vol. 103, no. 1–4, pp. 499–510, Mar. 2019.

[15] M. Huang, Z. Liu, and Y. Tao, "Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion," *Simulation Modelling Practice and Theory*, vol. 102, no. 14, p. 101981, Jul. 2020.

[16] J. Wang, J. Miao, J. Wang, F. Yang, K.-L. Tsui, and Q. Miao, "Fault diagnosis of electrohydraulic actuator based on multiple source signals: An experimental investigation," *Neurocomputing*, vol. 417, no. 15, pp. 224–238, Dec. 2020.

[17] X. Chen, B. Zhang, and D. Gao, "Bearing fault diagnosis base on multi-scale CNN and LSTM model," *Journal of Intelligent Manufacturing*, vol. 32, no. 4, pp. 971–987, Jun. 2020.

[18] J. Xu, X. Zhao, Y. Yu, T. Xie, G. Yang, and J. Xue, "Parametric sensitivity analysis and modelling of mechanical properties of normal- and high-strength recycled aggregate concrete using grey theory, multiple nonlinear regression and artificial neural networks," *Construction and Building Materials*, vol. 211, no. 14, pp. 479–491, Jun. 2019.

[19] L. G. B. Ruiz, R. Rueda, M. P. Cuéllar, and M. C. Pegalajar, "Energy consumption forecasting based on Elman neural networks with evolutive optimization," *Expert Systems with Applications*, vol. 92, no. 6, pp. 380–389, Feb. 2018.

[20] C. Guo, J. Lu, Z. Tian, W. Guo, and A. Darvishan, "Optimization of critical parameters of PEM fuel cell using TLBO-DE based on Elman neural network," *Energy Conversion and Management*, vol. 183, no. 5, pp. 149–158, Mar. 2019.

[21] K. Kolanowski, A. Świetlicka, R. Kapela, J. Pochmara, and A. Rybarczyk, "Multisensor data fusion using Elman neural networks," *Applied Mathematics and Computation*, vol. 4, no. 21, pp. 236–244, Feb. 2018.

[22] S. Kumar Chandar, "Grey Wolf optimization-Elman neural network model for stock price prediction," *Soft Computing*, vol. 25, no. 1, Jul. 2020.

[23] K. V. Starostin, "Creation of an Online Platform for Identification of Microorganisms: Peak Picking or Full-Spectrum Analysis," *Frontiers in Microbiology*, vol. 11, no. 11, Dec. 2020.

[24] H. Yu, H. Li, Y. Li, and Y. Li, "A novel improved full vector spectrum algorithm and its application in multi-sensor data fusion for hydraulic pumps," *Measurement*, vol. 133, no. 3, pp. 145–161, Feb. 2019.

[25] K. Liu, S. Li, L. Wang, Y. Ye, and H. Tang, "Full-Spectrum Prediction of Peptides Tandem Mass Spectra using Deep Neural Network," *Analytical Chemistry*, vol. 92, no. 6, pp. 4275–4283, Feb. 2020.

[26] R. Xu and M. Zhou, "Elman Neural Network-Based Identification of Krasnosel'skii–Pokrovskii Model for Magnetic Shape Memory Alloys Actuator," *IEEE Transactions on Magnetics*, vol. 53, no. 11, pp. 1–4, Nov. 2017.

[27] H. Raeisi-Vanani, M. Shayannejad, and A. Soltani-Toudeshki, "A Simple Method for Land Grading Computations and its Comparison with Genetic Algorithm (GA) Method," *International Journal of Research Studies in Agricultural Sciences*, vol. 3, no. 8, 2017.

[28] L. Zeinalkhani, A. Ali Jamaat, and K. Rostami, "Diagnosis of Brain Tumor Using Combination of K-Means Clustering and Genetic Algorithm," *Iranian Journal of Medical Informatics*, vol. 7, no. 12, p. 6, Nov. 2018.

[29] X. Zhang, B. Zhao, and Y. Lin, Machine Learning Based Bearing Fault Diagnosis Using the Case Western Reserve University Data: A Review," *IEEE Access*, vol. 8, no. 9, pp. 93155-93178, 2021.

[30] X. Chen, Z. Chen, and Y. Zhao, Numerical research on virtual reality of vibration characteristics of the motor based on GA-BPNN model," *Neural Computing and Applications*, vol. 29, no. 5, pp. 1343-1355, Mar. 2017.

[31] C. Sukawattanavijit, J. Chen, and H. Zhang, GA-SVM Algorithm for Improving Land-Cover Classification Using SAR and Optical Remote Sensing Data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 3, pp. 284-288, Mar. 2017.

[32] A. T. Eseye, J. Zhang, and D. Zheng, Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information," *Renewable Energy*, vol. 118, no. 5, pp. 357-367, Apr. 2018.