Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Fault Diagnosis and Prediction Techniques for Advanced Measurement Systems in Flexible Production of Metrology and Inspection

Yan Zhao[1,✉], Meng Yu[1], Zhenyu Zhao[1], Xuewei Guo[1], Jun Wang[1]

[1] *State Grid Jiangxi Power Supply Service Management Center, Nanchang, Jiangxi, 330000, China*

ABSTRACT

Measurement and verification play a crucial role in flexible production, and with the development of technology, advanced measurement systems in flexible production systems gradually integrate fault diagnosis and prediction techniques to improve production efficiency. In this paper, a deep confidence neural network model, combined with the ISSA-VMD feature fusion model, is used to model fault diagnosis and prediction in flexible production of power systems. The training effect, prediction performance, feature extraction and fault diagnosis of this paper's model in flexible production are evaluated and analysed through simulation experiments. The Loss value of this paper's model converges to about 0.05 after 15 rounds of training, and has a good fitting effect on the training and test sets. The RMSE, MAE and R^2 of the model in this paper are 0.613, 0.371 and 0.988, respectively, which show good prediction performance. And the prediction results in the measurement system of power generation in flexible production are also more close to the real results. In addition, the DBN model incorporating ISSA-VMD feature fusion can completely separate the five fault signals, and the overall fault identification accuracy reaches 98.53% for the fault test set selected in this paper, which has strong diagnostic effect. This study provides more scientific and effective technical support for metrological verification in flexible production.

Keywords: Deep confidence neural network, ISSA-VMD feature fusion, fault diagnosis and prediction, flexible production

1. Introduction

Metrological calibration refers to the process of testing, calibrating and verifying instruments, measuring equipment and metrological systems of production processes to ensure that they meet predetermined standards of accuracy and reliability. The purpose of metrological verification is to ensure the reliability and consistency of the metrological data, to ensure the accuracy of the

✉ Corresponding author.
E-mail address: 19720319454@163.com (Y. Zhao).

Received 05 March 2024; Revised 15 May 2024; Accepted 20 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-171

measurement results, and to provide the necessary guarantee for scientific research, engineering design and quality control [1-4]. The principle of metrological verification mainly includes the accuracy, sensitivity and repeatability of instruments. The accuracy of the instrument refers to the size of the deviation between the measurement result and the actual value of the instrument, i.e. the error of the instrument. The accuracy of the measurement results depends on the accuracy of the instrument, so the instrument must be made to achieve the required accuracy [5-8]. With the development of industrial modernisation, intelligence, automation, large-scale and complexity have gradually become the main features of flexible production, which, due to the high degree of integration and complexity of the structure, leads to a high failure rate and elevated risks in terms of safety and economy [9-12].

Flexible production is a flexible, adjustable and efficient production method, which plays an important role in modern manufacturing. The core concept of flexible production is to adjust the production capacity and production speed of the production line according to the market demand and customer requirements, in order to maximise customer satisfaction and improve the competitiveness of the enterprise [13-15], therefore, the fault diagnosis and prediction techniques of advanced measurement systems in flexible production are particularly important. Fault diagnosis and prediction techniques utilise real-time and historical state information of subsystems and components [16-17] to reduce the time and cost of the system operation and maintenance process with efficient and cost-effective diagnostic and predictive measures, improve system performance, safety, reliability and maintainability, and are a core technology for achieving both contingent maintenance and autonomous assurance [18-20].

In this paper, a flexible production simulation system for power systems is constructed, and a deep confidence neural network is used as the basic model for fault diagnosis and prediction of advanced measurement systems in flexible production. The power flexible production fault diagnosis and prediction system is modelled by determining the top-level feature classifiers, selecting appropriate model parameters, and incorporating the fault feature extraction model with ISSA-VMD feature fusion. Finally, combined with simulation experiments, the circuit faults in the system are identified, diagnosed, and the prediction performance of this paper's model is evaluated using RMSE, MAE, and R^2 assessment metrics.

2. Fault diagnosis and prediction models in flexible production of power systems

2.1. Power system flexible production simulation system

This paper proposes a power system “flexible production simulation”, which is a power system production simulation that takes into account load forecasting errors, the randomness, volatility and intermittency of uncertain power supply, the random failures of units and other uncertainty factors, as well as energy storage. In addition to obtaining the economic and reliability indicators of conventional power system stochastic production simulation, it can also be used to assess the impacts brought by the uncertain power source access to the power system.

At present, the power system large-scale access to the uncertainty of the power supply mainly wind power and photovoltaic power generation, so this paper proposes the power system “flexible production simulation” study of the power supply form is mainly conventional power (thermal, nuclear and hydropower), wind power, photovoltaic as the representative of the uncertain power supply as well as to cooperate with the large-scale uncertainty of the power supply effectively grid-connected energy storage (mainly pumped storage). Energy storage (mainly pumped storage). The steps of the “flexible production simulation” of the power system including the above power sources are as follows:

1) Select the horizontal annual planning scheme for which the power system “flexible production simulation” is required.

2) Conduct stochastic production simulation of the selected horizontal annual planning scheme by the serial operation frequency method, and obtain the conventional economic and reliability indexes, as well as the capacity credibility of the uncertain power source and the avoidable costs.
3) According to the multi-objective optimization theory, the maintenance plan of power system generating units is arranged by considering the system standby rate and maintenance risk degree.
4) Based on the system maintenance plan, develop the system unit combination program, if the unit combination program does not satisfy the system constraints, go back to re-conduct the unit maintenance plan development.
5) Conduct peak analysis based on the unit combination scheme to obtain the level of power system abandoned energy and electricity based on the peak level.
6) Based on the unit combination scheme for frequency regulation analysis, to get the level of power system abandoned energy and electricity based on the frequency regulation level.
7) Comprehensively derive the total power system power loss with uncertainty power source and energy storage.
8) Analyze and evaluate other characteristics of the power system as needed.
9) The end of the “flexible production simulation” of the power system.

2.2. Deep Confidence Neural Network Model

DBN (Deep Belief Network) is a very important structure in deep learning algorithms and is a probabilistic generative model [21-22]. Structurally, DBN has a deep neural network with multiple hidden layers, which is a combination of unsupervised feature learning and supervised parameter tuning. Unsupervised greedy learning is carried out through stacked restricted Boltzmann machines to achieve high-level, abstract feature extraction from raw data, after which the parameters are back-tuned by a back-propagation neural network (BP) to obtain accurate feature matrices for supervised learning of the data.

2.2.1. DBN model training based on restricted Boltzmann machines:
1) Restricted Boltzmann Machine. Restricted Boltzmann Machine (RBM) is an energy-based model that learns probability distributions from input data, and its basic composition includes two layers, visible and hidden, with no directionally full connections between neurons in the visible and hidden layers through weights and no connections between neurons within the layers [23-24].

Where n is the number of neural units in the visible layer; $v=[v_1,v_2,\ldots,v_n]$ is the state of neurons in the visible layer; m is the number of neural units in the hidden layer; $h=[h_1,h_2,\ldots,h_m]$ is the state of neurons in the hidden layer; and w_{nm} is the connection weight between the visible layer and the hidden layer. Assuming that there are n visible units in the visible layer and m hidden units in the hidden layer, the energy function between the visible layer neural units and the hidden layer neural units (v,h) is when both the visible and hidden unit states are binary (0 or 1):

$$E\left(v,h|\theta\right)=-\sum_{i=1}^{n}a_iv_i-\sum_{j=1}^{m}b_jh_j-\sum_{i=1}^{n}\sum_{j=1}^{m}v_iw_{ij}h_j \tag{1}$$

where v_i is the state of the i nd neuron in the visible layer; h_j is the state of the j th neuron in the hidden layer; $\theta=\left(w_{ij},a_i,b_j\right)$ is the value of the parameter when training the RBM; w_{ij} is the connection weights between the i th neural unit in layer v and the j th neural unit in layer h; and a_i and b_j are the bias values of the i th unit in layer v and the i th unit in layer h, respectively. $\left(v,h\right)$ for the joint probability distribution:

$$P\left(v,h|\theta\right)=\frac{e^{-E\left(v,h|\theta\right)}}{Z\left(\theta\right)} \tag{2}$$

Where, $Z(\theta)=\sum_{v,h} e^{-E(v,h|\theta)}$ is the normalisation factor, which represents the energy summation of all possible states of the set of neurons in the visible and hidden layers.

When the states of the visible layer neurons are determined, the activation probability of the hidden layer neurons can be obtained, in this case, the activation probability of the i nd neuron of the hidden layer is:

$$P\left(h_j=1|v\right)=\sigma\left(b_j+\sum_{i=1}^{n} v_i W_{ij}\right) \tag{3}$$

Similarly, the activation probability of the i st neuron of the visible layer is obtained when the state of the hidden layer neurons is determined:

$$P\left(v_i=1|h\right)=\sigma\left(a_i+\sum_{j=1}^{m} W_{ij} h_j\right) \tag{4}$$

Where $\sigma(x)$ is the activation function, which exists to improve the nonlinear mapping ability of the model. The sigmoid is chosen as the activation function because it can map the value of $(-\infty,+\infty)$ to the interval $[0,1]$:

$$\sigma(x)=\frac{1}{1+e^{-x}} \tag{5}$$

Summing all the states of the hidden layer node set yields the marginal distribution of the visible layer node set, which represents the probability that the visible layer node set is under some state distribution:

$$P(v|\theta)=\frac{1}{Z(\theta)}\sum_{h} e^{-E(v,h|\theta)} \tag{6}$$

The parameter $\theta=\left(w_{ii},a_i,b_j\right)$ is learnt by solving the great log-likelihood function of the training dataset to find the most suitable value of θ.

The likelihood function is solved by using gradient descent method to solve θ. The logarithm of the likelihood function is greatly simplified for ease of computation:

$$\theta^{*}=\arg\max L(\theta)=\arg\max\sum_{t=1}^{T}\log P\left(v^{(t)}|\theta\right) \tag{7}$$

where T is the number of training samples and the partial derivative of parameter θ during the solution is:

$$\frac{\varphi L(\theta)}{\varphi\theta}=\sum_{t=1}^{T}\left[\left\langle\frac{\varphi\left(E\left(v,h|\theta\right)\right)}{\varphi\theta}\right\rangle_{P(h|v,\theta)}-\left\langle\frac{\varphi\left(E\left(v,h|\theta\right)\right)}{\varphi\theta}\right\rangle_{P(v,h|\theta)}\right] \tag{8}$$

Where, $\langle\cdot\rangle_P$ denotes the probability distribution P mathematical expectation; $P\left(h|v,\theta\right)$ denotes the probability distribution of the hidden layer when the neurons in the visible layer are the input training samples.

Combining the above equation, the update criterion for the model parameters is obtained:

$$\begin{aligned}\Delta W_{ij} &= \varepsilon\left(\left\langle v_i h_j\right\rangle_{data} - \left\langle v_i h_j\right\rangle_{model}\right)\\ \Delta b_j &= \varepsilon\left(\left\langle h_j\right\rangle_{data} - \left\langle h_j\right\rangle_{model}\right)\\ \Delta a_i &= \varepsilon\left(\left\langle v_i\right\rangle_{data} - \left\langle v_i\right\rangle_{model}\right)\end{aligned} \tag{9}$$

where ε is the learning rate; $\langle\cdot\rangle_{data}$ is the mathematical expectation defined by the input training dataset; and $\langle\cdot\rangle_{model}$ is the mathematical expectation defined by the model.

The Contrastive Divergence, (CD) algorithm [25], which can obtain a better valuation after a small amount of training, is based on the idea of reconstructing the RBM structure, which mainly consists of reconstructing a hidden layer to form a visible layer first, and then building a corresponding hidden layer based on the visible layer formed by the reconstruction.

According to Eq. (3) and Eq. (4) conditional probability distributions, the computation of the model expectation in Eq. (9) is converted into the mathematical expectation defined by the reconstructed model by two Gibbs sampling. The update law for the entire network model parameters in Eq. (9) becomes:

$$\begin{aligned}\Delta W_{ij} &= \varepsilon\left(\left\langle v_i h_j\right\rangle_{data} - \left\langle v_i h_j\right\rangle_{recon}\right)\\ \Delta b_j &= \varepsilon\left(\left\langle h_j\right\rangle_{data} - \left\langle h_j\right\rangle_{recon}\right)\\ \Delta a_i &= \varepsilon\left(\left\langle v_i\right\rangle_{data} - \left\langle v_i\right\rangle_{recon}\right)\end{aligned} \tag{10}$$

where $\langle\cdot\rangle_{recon}$ is the mathematical expectation defined by the reconstructed model.

2) DBN structure and training process. The DBN model structure is shown in Fig. 1, and its parameter training is mainly divided into two stages:

(1) Pre-training phase. This phase adopts greedy learning, using unlabelled data to train each layer of RBM bottom-up unsupervised, using CD-1 algorithm to reconstruct each layer of RBM according to Eq. (3) and Eq. (4), and updating the model parameters θ according to Eq. (10), and completing the training of one layer of RBM when reaching the set number of iterations or the threshold to ensure that the feature information is mapped optimally within the layer.

(2) Tuning phase. When the pre-training is finished, it is necessary to add an inverse BP network to the tail layer of the DBN network, connect the classifiers, use the cross-entropy as the loss function, and make use of the error between the labelled data and the output of the classifiers to carry out supervised fine-tuning of the parameter θ matrices obtained from the DBN network in the pre-training stage from the top down, so as to make the output of the DBN model maximize the presentation of the original inputs.

Finally, the top classifier is used to classify the output.

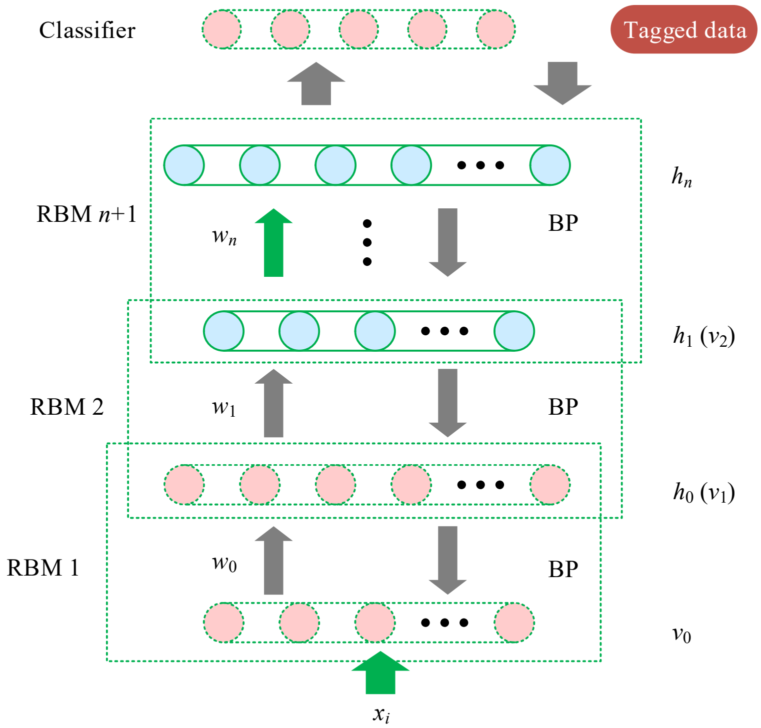

Fig. 1. DBN model structure

2.2.2. Top-level classifier determination. In this paper, Softmax function is chosen to implement the DBN output feature classification.Softmax regression model is an extension of Logistic regression model in solving multiclassification problems.Multiclassification carried out by Softmax regression, classes are mutually exclusive, i.e., an input can only be classified into one class.

Its output can take more than one, if the classification problem has k category, i.e., category label y can take k different values. Suppose there are m samples, for the training set $\{(x_1, y_1),(x_2, y_2),(x_m, y_m)\}$, the category label is $y_m \in \{1, 2, \cdots k\}$; for a given input sample i, its feature vector $x_i = [x_{il}, x_{21}, \ldots, x_{ni}]^T$, n is the sample dimension; using the hypothesis function $h_0(x_i)$ to estimate x belonging to the probability of occurrence of each category value $P(y_i = j|x_i)$, $j = 1, 2, \cdots k$, the class with the largest probability value is the category to which the input sample is ultimately identified as belonging to, i.e., the category label of the output:

$$h_\theta(x_i) = \begin{bmatrix} P(y_i = 1|x_i, \theta) \\ P(y_i = 2|x_i, \theta) \\ \vdots \\ P(y_i = k|x_i, \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{\theta_1^{\mathrm{T}} x_1}} \begin{vmatrix} e^{\theta_1^{\mathrm{T}} x_1} \\ e^{\theta_2^{\mathrm{T}} x_1} \\ \vdots \\ e^{\theta_k^{\mathrm{T}} x_1} \end{vmatrix} \tag{11}$$

where $h_0(x_1)$ outputs a k-dimensional column vector; $\sum_{j=1}^{k} e^{\theta_j^{\mathrm{T}} x_1}$ is a normalisation of the probability distribution of the classification result, which ensures that the sum of the probability values of the samples belonging to different classes is equal to 1; and $\theta_1, \theta_2, \cdots, \theta_k$ is the parameter vector of the Softmax classifier. The probability value that sample x_i belongs to class k is:

$$P(y_i = j|x_i, \theta) = \frac{e^{\theta_j^{\mathrm{T}} x_1}}{\sum_{j=1}^{k} e^{\theta_j^{\mathrm{T}} x_1}} \tag{12}$$

The Softmax loss function is to transform the true weight vector into a probability distribution, often using cross entropy as the loss function to ensure that the output of the model is as close as possible to the desired output, the loss function is:

$$J(\theta)=-\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k}1\{y_i=j\}\log\frac{e^{\theta_j^T x_1}}{\sum_{j=1}^{k}e^{\theta_j^T x_1}}\right] \tag{13}$$

where $1\{\cdot\}$ is the indicative function, which takes the value of 1 when $y_1=j$ and 0 otherwise.

The minimisation process of cross-entropy is the process of improving the accuracy of Softmax classification, i.e., finding the appropriate parameter θ to minimise the value of Eq. (13), which is generally used to solve $\min J(\theta)$ the unconstrained optimisation problem by gradient descent method. The partial derivative of each parameter θ in Eq. (13) during gradient descent is:

$$\frac{\varphi J(\theta)}{\varphi\theta_j}=-\frac{1}{m}\sum_{i=1}^{m}\left\{x_i\left[1\{y_i=j\}-p\left(y_i=j\middle|x_i,\theta\right)\right]\right\} \tag{14}$$

The gradient descent process requires the following updates at each iteration:

$$\theta_j^{t+1}=\theta_j^t-\alpha\frac{\varphi J(\theta)}{\varphi\theta_j}\ j=1,2,\cdots k \tag{15}$$

where α is the learning rate; t is the number of iterations.

In order to penalise excessive parameter values, a weight decay term needs to be added to equation (13):

$$J(\theta)=-\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k}1\{y_i=j\}\log\frac{e^{\theta_j^T x_1}}{\sum_{j=1}^{k}e^{\theta_j^T x_1}}\right]+\frac{\lambda}{2}\sum_{i=1}^{m}\sum_{j=1}^{k}\theta_{ij}^2 \tag{16}$$

where θ_{ij} is the i rd row and i th column element of θ; after adding the weight decay term, for any $\lambda>0$.

2.2.3. DBN model parameter selection. The goodness of DBN network fault identification is not only dependent on the algorithm, but also requires the regulation of network parameters.

1) Input and output layer nodes: the input layer size is the data length of a single sample, while the output layer nodes depend on the sample type.

2) Number of hidden layers: there is no standard for selecting the number of hidden layer nodes. But it can be referred to as equation (17):

$$S=\sqrt{m+n}+a \tag{17}$$

3) Weights and bias: In the unsupervised learning process of DBN, it is necessary to initialise the weights of each node and set the value of bias. The initialisation can be set according to equation (18):

$$\begin{aligned} w&=0.1\times randn(n,m)\\ a&=0.1\times randn(1,n)\\ b&=0.1\times randn(1,m) \end{aligned} \tag{18}$$

In the above equation, where $randn(\ )$ function is used to generate standard normally distributed random numbers, and $randn(n,m)$ means to generate n rows and m columns of normally distributed matrix.

4) Number of iterations: the number of iterations is the number of optimisation calculations in the training process, which makes the approximation ability of the network to meet the requirements.

5) Learning rate: the learning rate affects the speed of the training process. When the learning rate is larger, the learning is faster, but it may lead to oscillations in the model training process and jump out of the optimal solution; when the learning rate is small, it may not converge in the iteration cycle.

2.3. Research on fault diagnosis and prediction based on DBN neural network

2.3.1. Feature extraction model based on ISSA-VMD feature fusion. Aiming at the problem that the early fault features of the bearing are weak and easy to be flooded by noise and difficult to be extracted, this paper proposes a fault feature extraction model based on ISSA-VMD feature fusion, and the flow of ISSA-VMD feature extraction model is shown in Fig. 2.

Firstly, the bearing signal is decomposed using the VMD optimised by ISSA parameters to obtain k IMF components containing fault information. Secondly, the correlation coefficient between each component and the original signal is calculated, and the first 6 components with the richest fault information are selected and the feature indexes energy entropy and sample entropy are calculated to form a high-dimensional multi-feature set. Finally, the t-SNE algorithm is used to perform feature fusion on the high-dimensional multi-feature set to obtain the low-dimensional deep-level feature set, to achieve the fault feature extraction and visualisation.

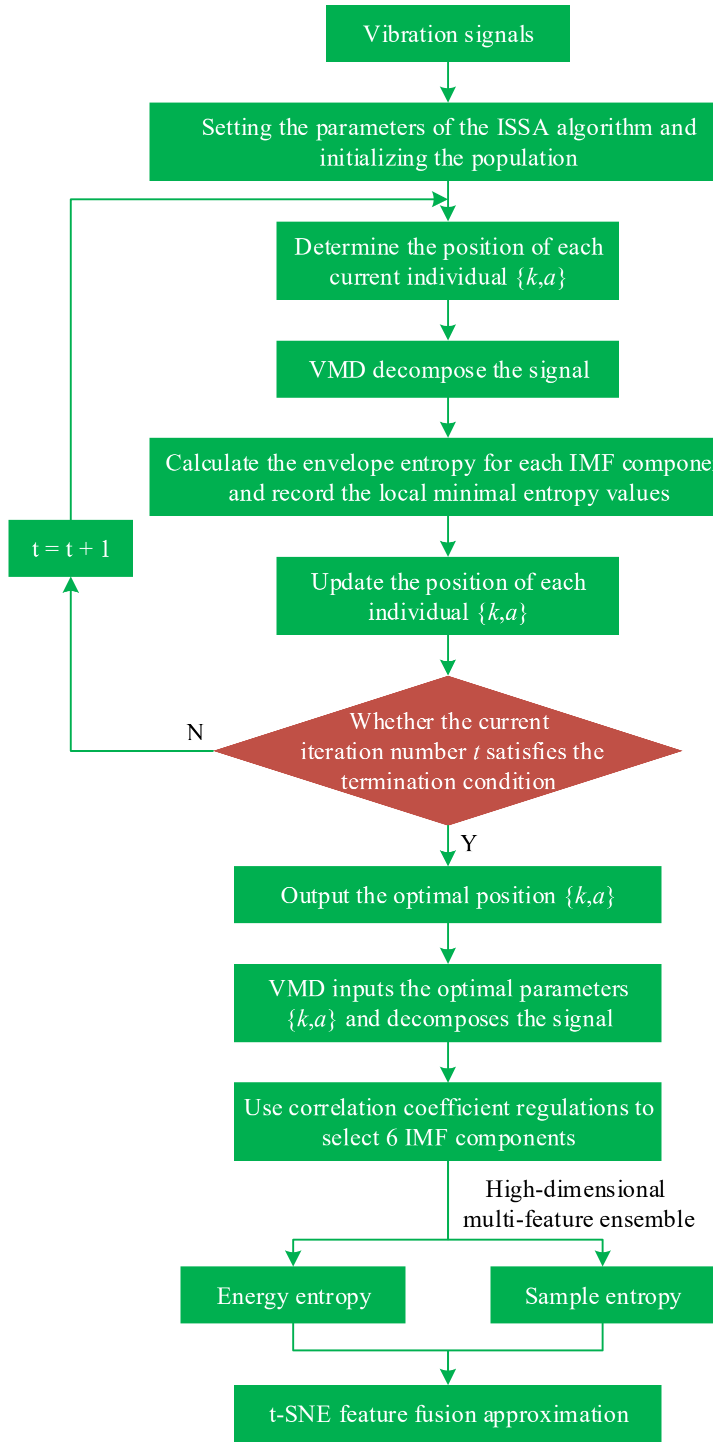

Fig. 2. ISSA-VMD Feature fusion feature extraction model

2.3.2. Analogue circuit troubleshooting process. The key factor affecting fault diagnosis is feature extraction, here we propose a feature extraction method based on time domain signal analysis, the features selected for time domain feature extraction include sample extreme deviation, mean, standard deviation, skewness, cragness, and entropy as a characterisation of the analogue circuit fault features, and then the resulting feature vector is used as an input to the DBN network model. The flow of the analogue circuit fault methodology is shown in Fig. 3.

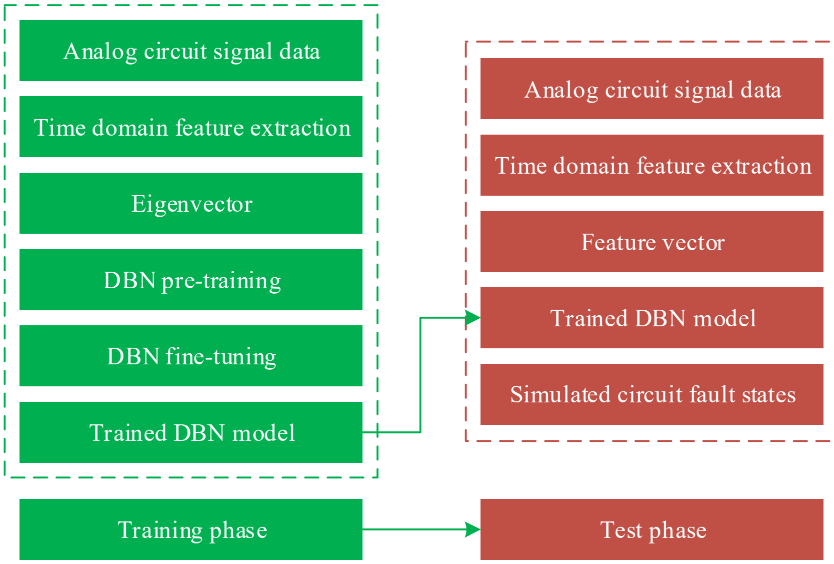

Fig. 3. Circuit fault diagnosis flow chart

2.3.3. Modelling of Fault Diagnosis and Prediction System for Power Flexible Production. When a large amount of flexible production process data is obtained, data noise needs to be filtered. The study of flexible production process modelling and optimization based on evolutionary deep confidence learning enables the model to have better generalisation and fitting ability, fully reflecting the distribution of process extremes, so that near-optimal values can be obtained when finding the optimum.

The evolutionary algorithm can help the deep confidence network to find the optimal structure and better initialisation weight values in the first phase of the algorithm. When the first stage is completed, all the connection weight values and activation function values are optimised, all the hidden layers will be connected, and then the back propagation algorithm is used for fine-tuning, which can significantly reduce the computational cost and is more suitable for multidimensional data processing.

Facing the flexible production process monitoring of the power system, the establishment of equipment fault diagnosis and prediction model can be carried out for the unit-level CNC equipment, which can make full use of the collected equipment data for fault classification identification and prediction identification. The deep confidence network model architecture is continuously adjusted through evolutionary algorithms to ultimately obtain the optimal structure, and then the labelled data is used for pre-training to obtain the model, and then the equipment data is collected online for fault diagnosis and prediction. The modelling process of fault diagnosis and prediction based on deep confidence is shown in Figure 4.

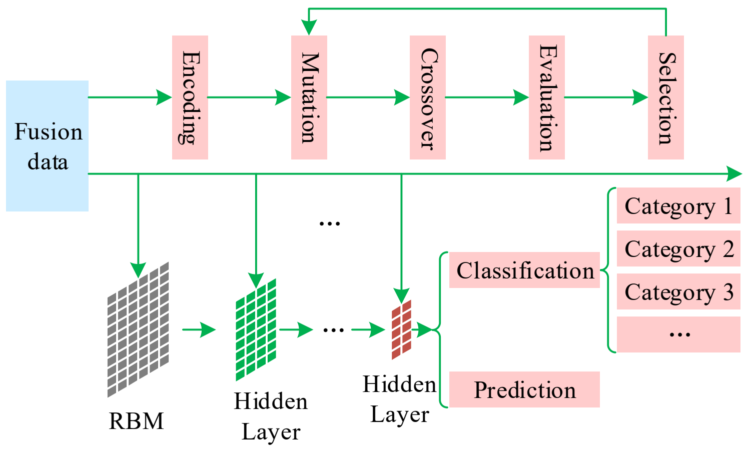

Fig. 4. Based on deep confidence fault diagnosis and prediction modeling process

2.3.4. Indicators for assessing the performance of time series forecasting. The structure of the basic DBN forecasting model is shown in Figure 5. There are many metrics to assess the performance of time series forecasting, and in this paper, we will choose the following three quantifiable metrics to assess the results of forecasting.

1) Root Mean Square Error (RMSE):

$$RMSE = \sum_{i=1}^{l} \sqrt{\frac{\left(y_i - \tilde{y}_i\right)^2}{l}} \tag{19}$$

2) Mean Absolute Error (MAE):

$$MAE = \frac{1}{l} \sum_{i=1}^{l} \left| y_i - \tilde{y}_i \right| \tag{20}$$

3) Coefficient of determination R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{l} \left(y_i - \tilde{y}_i\right)^2}{\sum_{i=1}^{l} \left(y_i - \bar{y}_i\right)^2} \tag{21}$$

Of the three indicators, the length of the prediction data is l, y_i represents the actual value of the i rd prediction data, and $\tilde{y}_i$ represents the predicted value of the i th prediction data. Since RMSE reflects the average deviation between the predicted value and the actual value, it is sensitive to too large or too small errors in the prediction, and its value is greater than or equal to 0. When the value is 0, it indicates the highest accuracy of the prediction. MAE is the mean absolute error, and the smaller its value, the smaller the difference between the predicted value and the actual value. The value of the coefficient of determination R^2 is also greater than or equal to 0. When the value of R^2 is equal to 1, it means that there is no error between the predicted value and the actual value, and the prediction accuracy is the highest.

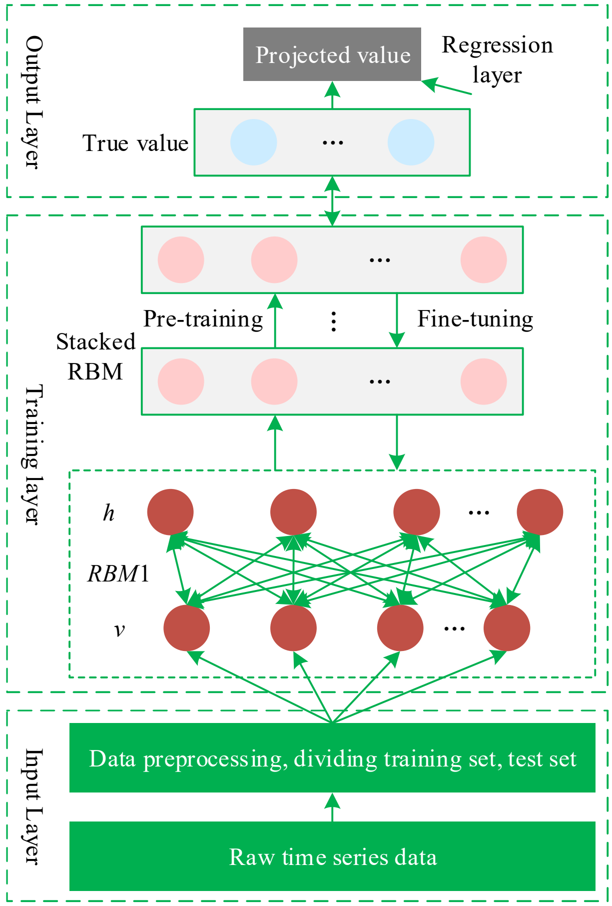

Fig. 5. DBN basic prediction model

3. Diagnostic and predictive analysis of faults in flexible production of power systems

3.1. Effectiveness of training based on flexible production in power systems

In this paper, based on the "flexible production simulation" of the power system, we verify the prediction performance of the deep confidence network model on the photovoltaic power generation. The pre-processed historical PV power and corresponding historical meteorological data from December 31, 2022 to December 31, 2023 are intercepted as the dataset, of which 50% is used as the training set and the remaining 50% as the test set. The DBN model is trained and tested for effectiveness several times, and the DBN model with optimal performance is selected to predict the power values at 84 time nodes on the following day.

Figure 6 shows the loss variation curve (objective function is Loss) when the ultra-short-term prediction model is trained. From the figure, it can be seen that the loss function is decreasing for both training and testing, and converges to about 0.05 after 15 rounds of training. It shows that the training of the model in this paper is effective and has a good fit on both the training and test sets.

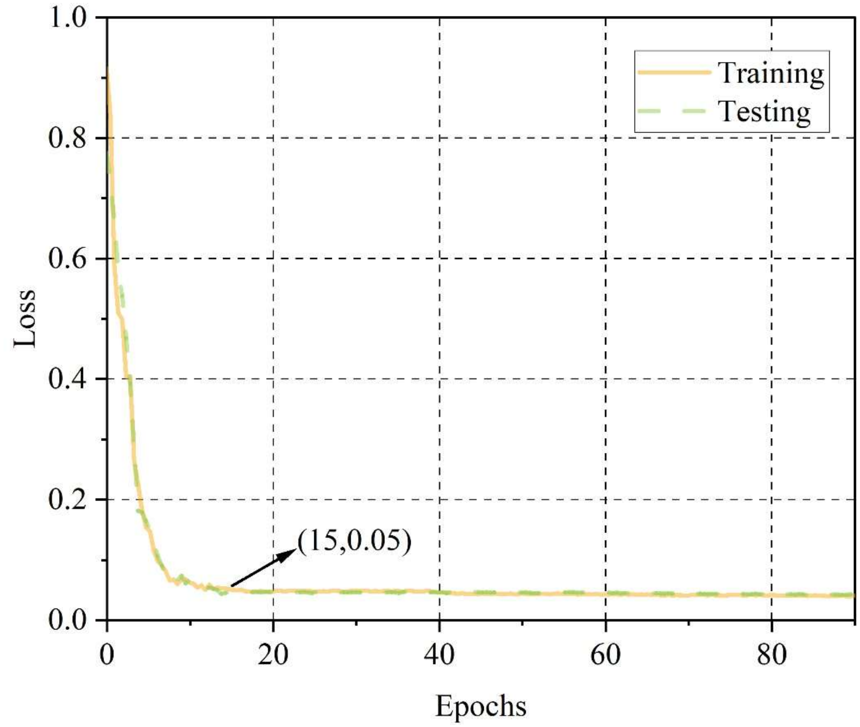

Fig. 6. The loss curve of the super short-term prediction model training

3.2. Comparison of ultra-short-term forecasting performance of flexible production in power generation

In this section, based on the experimental results in the previous section, the model in this paper is compared with the current commonly used prediction models (LSTM, CNN, BPN) for the comparison experiments of predicting the PV power in the next 2 days.

Figure 7 demonstrates the results of the comparison of the performance indexes of the four prediction models LSTM, CNN, BPN and DBN. From the data in Fig. 7, it can be seen that this paper's model has the smallest RMSE and MAE and the largest R^2 among all the prediction models. The RMSE and MAE of this paper's model are reduced by 32.4% and 42.1%, respectively, and R^2 is improved by 20.4% compared to the worst performing model (BPN). This validates the performance of the model proposed in this paper.

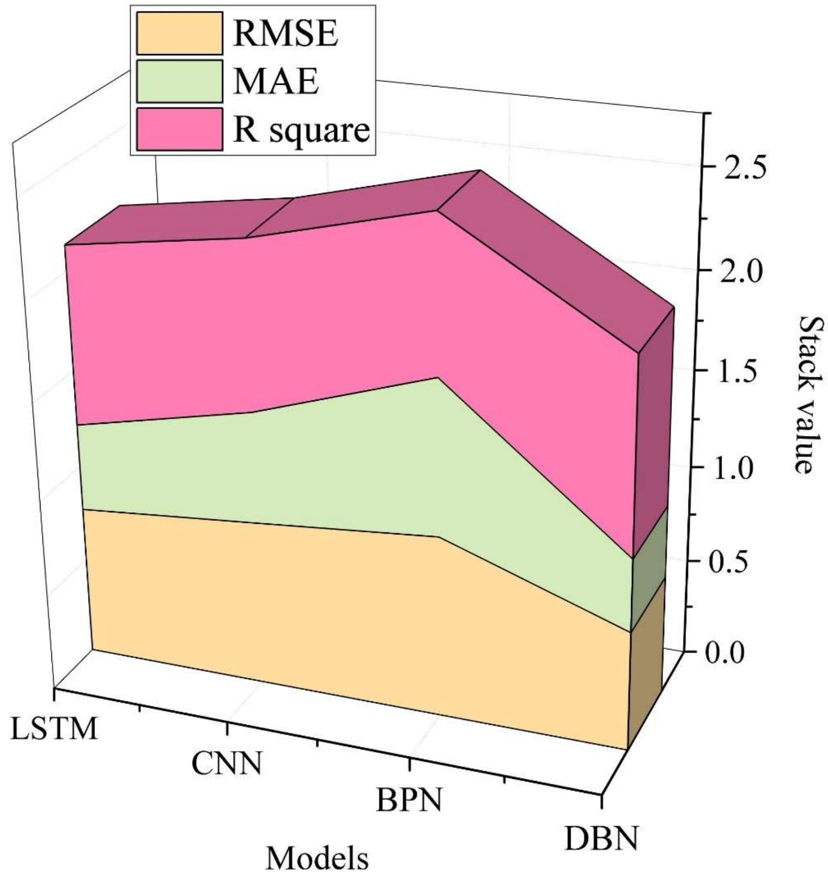

Fig. 7. The performance index of four prediction models is compared

Observe the results of the ultra-short-term PV power prediction using the DBN model with a prediction step of 1, i.e., predicting the PV power after 10 min. Figure 8 shows the prediction folds of PV power from 23 September 2023 to 25 September. The red folds in the figure represent the real PV power, and the rest of the colours represent the prediction folds of different models, respectively. According to the data in the figure, it can be seen that the model proposed in this paper has a better prediction fitting effect and higher prediction accuracy than the comparison model.

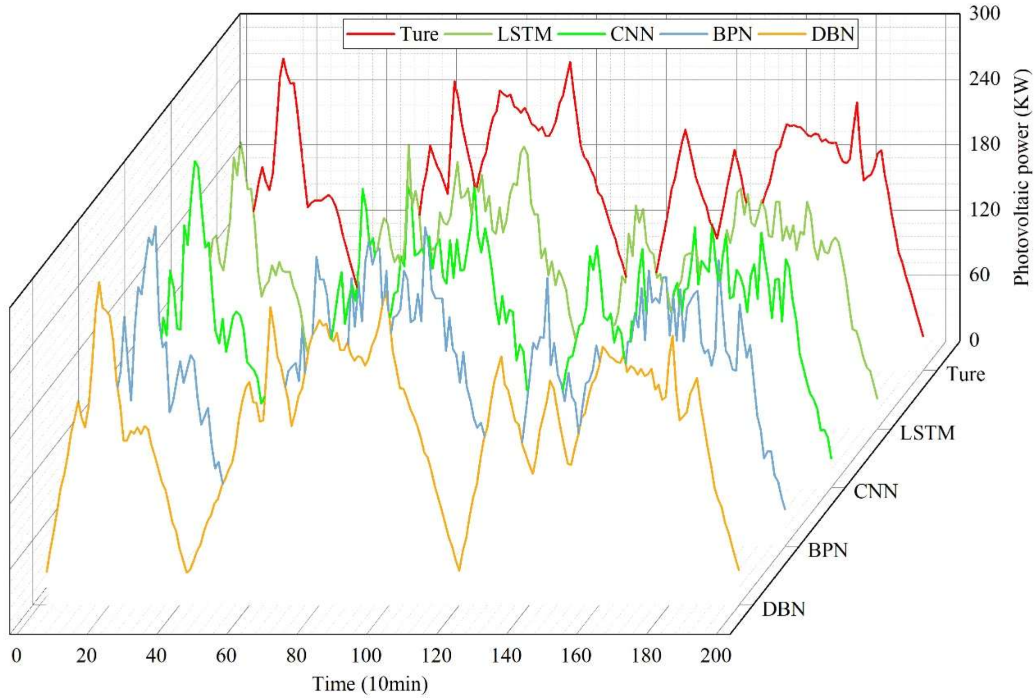

Fig. 8. Prediction of photovoltaic power

3.3. Feature Extraction Visual Scatter Analysis

The experiments in this section incorporate the feature extraction model of ISSA-VMD feature fusion into the DBN structure to verify the layer-by-layer feature extraction capability of the DBN model in fault diagnosis. Since the dimension of the input data determines the number of nodes in the input layer, and the number of fault categories determines the number of nodes in the output layer, the input nodes are set to 500, and the output nodes are set to 5. Since the constructed fault simulation dataset and the deep confidence network have a high number of feature dimensions extracted at each layer, which is not conducive to the visual observation, for the convenience of visualisation, only the first 3 principal components of the fault features are shown. Figures 9 and 10 give the visualisation results of the features extracted from the original data and the third hidden layer of the DBN model based on ISSA-VMD feature fusion, respectively.

As can be seen from the visualised scatter plot of the original features in Fig. 9, the five fault signals are tightly clustered together. When extracted layer by layer from the third implicit layer of the DBN model, it can be seen from Fig. 10 that the number of differentiated signal types gradually increases, and finally the five faults are completely differentiated. Combining the results in the two figures, it can be concluded that signals of the same type are gradually clustered together and signals of different types are gradually distinguished. It is obvious that the features extracted from the third hidden layer of the DBN model can be input into the classifier to obtain higher fault recognition accuracy.

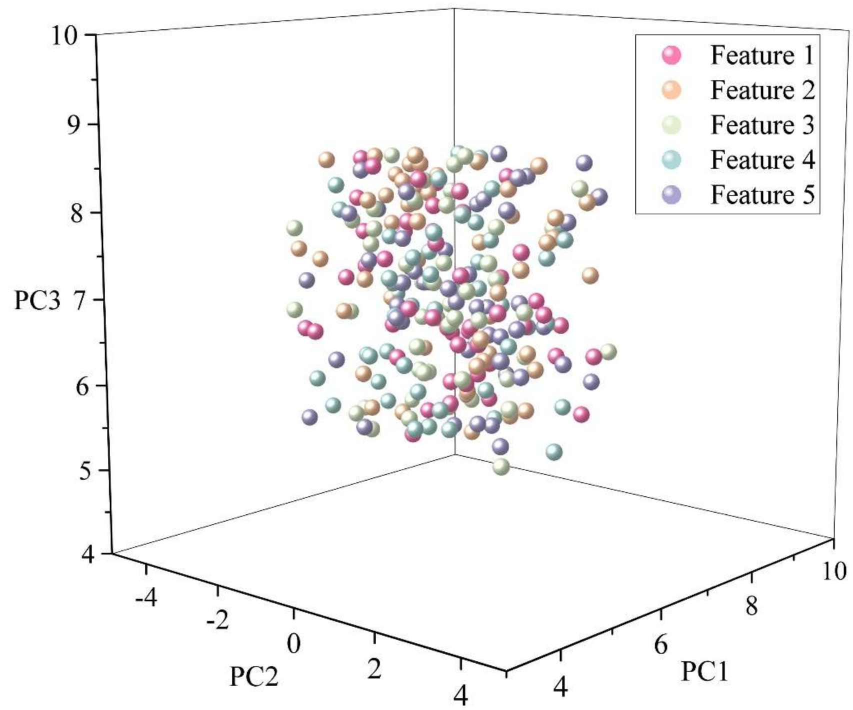

Fig. 9. Original feature 3D-PCA mapping

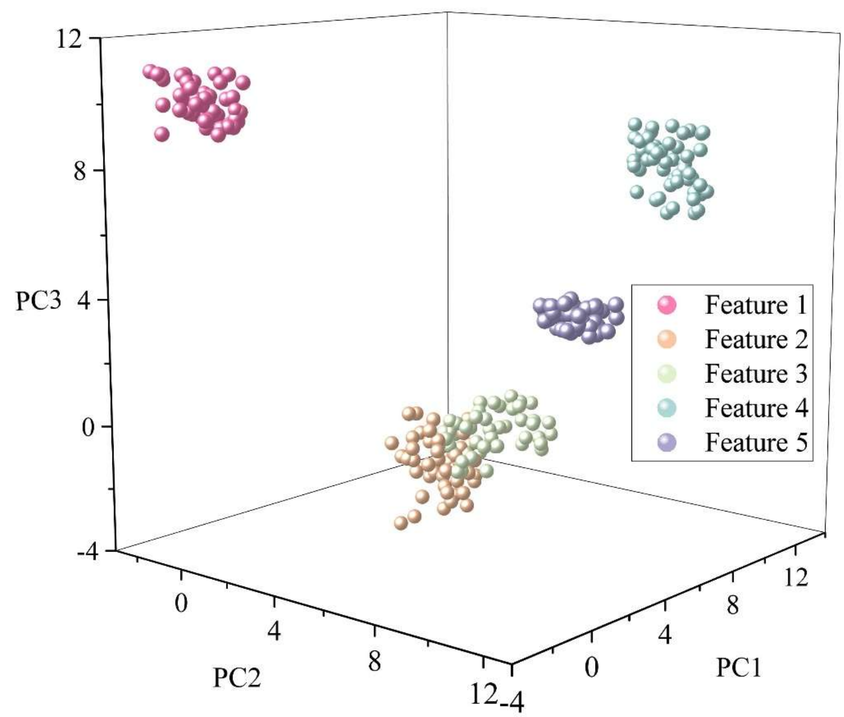

Fig. 10. The third implicit layer features 3D-PCA mapping

3.4. *Fault diagnosis analysis in power based flexible production*

In order to measure the effectiveness of fault diagnosis using the DBN model, this section focuses on evaluating the effectiveness of fault diagnosis in terms of the accuracy of fault classification. In this paper, the overall accuracy rate ACR and the single fault identification accuracy rate SAR are used as the evaluation indexes of the fault diagnosis method, which are defined as follows, respectively:

1) ACR. The number of correct fault classifications divided by the number of all samples:

$$ACR = \frac{TP + TN}{TP + TN + FP + FN} \tag{22}$$

2) SAR. The number of single fault identification accuracies depends on the number of fault categories and indicates the number of samples that are identified correctly out of all samples of a particular fault category:

$$SAR_1 = \frac{TP}{TP + FN} \tag{23}$$

$$SAR_2 = \frac{TN}{FP + TN} \tag{24}$$

The fault identification confusion matrix for classifying the fault data set consisting of the five simulated signals above using the DBN model is shown in Figure 11 below. The figure 150 represents the number of samples in each category of the test set. Based on the results in the confusion matrix, combined with the overall accuracy ACR of the above formula and the single fault recognition rate SAR, the fault classification results of the deep confidence network are derived as shown in Table 1.

Combining the figure and table, it can be seen that the recognition rate of two kinds of faults are up to 100%, respectively, Fault 2 and Fault 5, and the remaining three kinds of faults are also up to more than 95%, and the overall fault recognition accuracy rate is also up to 98.53%, which to a certain extent can show that the DBN has a strong fault diagnostic ability.

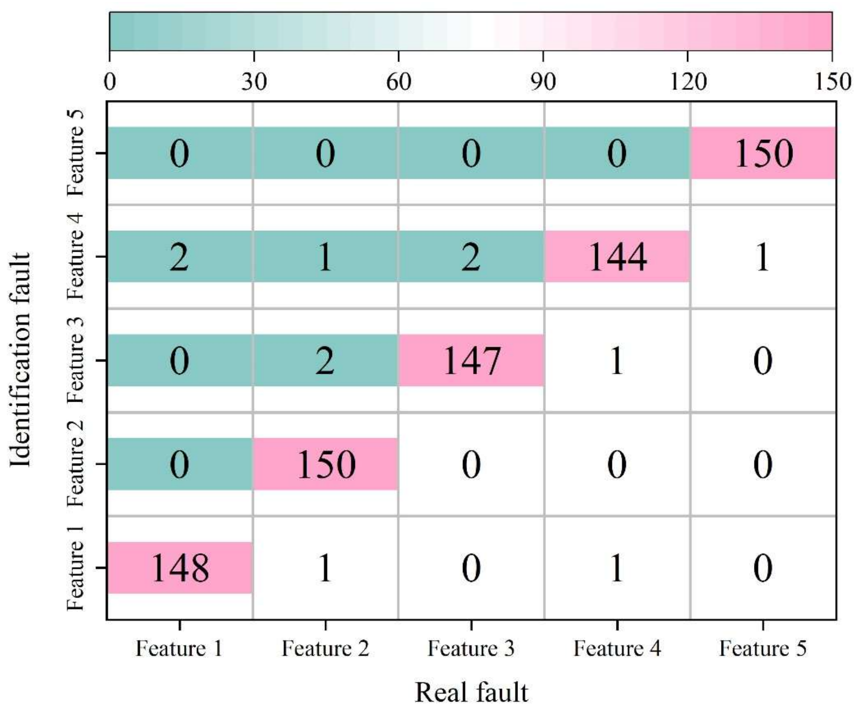

Fig. 11. Fault identification confusion matrix

Table 1. The fault classification of the deep confidence network

Fault type	1	2	3	4	5
SAR	98.67%	100%	98%	96%	100%
ACR	98.53%				

4. Conclusion

In this paper, the flexible production in power system is taken as the object, and the fault diagnosis and prediction in the flexible production of power system is modelled by deep confidence neural network model and ISSA-VMD feature fusion fault extraction model. And simulation experiments are used to evaluate the prediction performance and fault diagnosis effect of the model in this paper.

The DBN model proposed in this paper shows superior results in model training, with the training and testing Loss converging to about 0.05 by the 15th round of training. And the RMSE and MAE of the model in this paper are reduced by 15.9%~32.4% and 7.6%~42.1%, respectively, and the R^2 is improved by 7.3%~20.4% compared with the comparison model. And the model in this paper has a better fitting effect with the real generation PV power folds. In addition, the DBN model incorporating ISSA-VMD feature fusion can completely separate the five fault signals after the feature extraction of the third hidden layer, and has a higher fault identification accuracy. The model in this paper has the lowest fault identification rate of 96% for the five simulated signals, and the identification rate of Fault 2 and Fault 5 is as high as 100%, and the overall fault identification accuracy also reaches 98.53%.

Funding

This research was supported by the State Grid Jiangxi Electric Power Co., Ltd. Technology Project: Research on Key Technologies for Flexible Production of Metrological Calibration for Advanced Measurement Systems (Project No: 52185223000H).

References

[1] Meškuotienė, A., Dobilienė, J., Raudienė, E., & Gaidamovičiūtė, L. (2022). A review of metrological supervision: towards the common understanding of metrological traceability in legal and industrial metrology. Mapan, 37(3), 693-701.

[2] Kučera, Ľ., Patin, B., Gajdošík, T., Palenčár, R., Palenčár, J., & Ujlaky, M. (2020). Application of metrological approaches in the design of calibration equipment for verification of float level gauges. Measurement Science Review, 20(5), 230-235.

[3] Zhirnova, E. A., Malakhova, Y. G., Kulakova, N. S., & Kuzmich, R. I. (2020). Estimation Of Costs For Metrological Support Of Production. European Proceedings of Social and Behavioural Sciences.

[4] Savio, E., De Chiffre, L., Carmignato, S., & Meinertz, J. (2016). Economic benefits of metrology in manufacturing. CIRP Annals, 65(1), 495-498.

[5] OLESKIV, O., & MYKYTYN, I. (2015). Metrological verification of cyberphysical systems. Advances in IT and Electrical Engineering, 23(34-3), 87-92.

[6] Nani, V. M. (2018). A new algorithm for predictive metrological verification of measuring apparatus. International Journal of Precision Engineering and Manufacturing, 19, 167-172.

[7] Markovič, J., Mihok, J., Ďuriš, S., & Schreier, Z. (2016). Legal metrology and system for calibration and verification of the radar level sensors. Quality Innovation Prosperity, 20(1), 95-103.

[8] Huang, Y. L., Bai, H., Feng, J. W., Fan, H. Z., & Ma, J. F. (2014). Research on the metrological verification system for an equipment. Advanced Materials Research, 945, 2175-2179.

[9] Leng, S. (2020). Scheduling method of green flexible production line for enterprise products based on task priority. International Journal of Product Development, 24(2-3), 130-145.

[10] Jakubovskis, A. (2017). Flexible production resources and capacity utilization rates: A robust optimization perspective. International Journal of Production Economics, 189, 77-85.

[11] Schlegel, A., Langer, T., & Putz, M. (2017). Developing and harnessing the potential of SMEs for eco-efficient flexible production. Procedia Manufacturing, 9, 41-48.

[12] Radziwon, A., Bilberg, A., Bogers, M., & Madsen, E. S. (2014). The smart factory: exploring adaptive and flexible manufacturing solutions. Procedia engineering, 69, 1184-1190.

[13] Zhuravel, K. (2019). Concept of flexible management at enterprise in digitalization and lean production conditions. Humanities studies, (1), 98-107.

[14] Yadav, D., Singh, R., Kumar, A., & Sarkar, B. (2022). Reduction of Pollution through Sustainable and Flexible Production by Controlling By-Products. Journal of Environmental Informatics, 40(2).

[15] Harcuba, O., & Vrba, P. (2015, September). Ontologies for flexible production systems. In 2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA) (pp. 1-8). IEEE.

[16] Zhang, F., Chen, M., Zhu, Y., Zhang, K., & Li, Q. (2023). A review of fault diagnosis, status prediction, and evaluation technology for wind turbines. Energies, 16(3), 1125.

[17] Liu, C., Cichon, A., Królczyk, G., & Li, Z. (2021). Technology development and commercial applications of industrial fault diagnosis system: a review. The International Journal of Advanced Manufacturing Technology, 1-33.

[18] Heda, Z. (2019). Fault diagnosis and life prediction of mechanical equipment based on artificial intelligence. Journal of Intelligent & Fuzzy Systems, 37(3), 3535-3544.

[19] Zhang, X., Rane, K. P., Kakaravada, I., & Shabaz, M. (2021). Research on vibration monitoring and fault diagnosis of rotating machinery based on internet of things technology. Nonlinear Engineering, 10(1), 245-254.

[20] Wang, X., Guo, J., Jiang, L., Fu, J., & Li, B. (2016, August). Intelligent fault diagnosis and prediction technologies for condition based maintenance of track circuit. In 2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT) (pp. 276-283). IEEE.

[21] Amol Babaso Majgave & Nitin L. Gavankar. (2024). Automatic phishing website detection and prevention model using transformer deep belief network. Computers & Security104071-104071.

[22] Pushpa Athisaya Sakila Rani & N. Suresh Singh. (2024). Improved Detection and Classification of Multiple Tasks in Paddy Crops Using Optimized Deep Belief Networks with Bidirectional Long-Short Term Memory. SN Computer Science(7),831-831.

[23] J. Garí,E. Romero & F. Mazzanti. (2024). Learning restricted Boltzmann machines with pattern induced weights. Neurocomputing128469-128469.

[24] Seongil Im,JinGyeong Hwang,Jae Seung Jeong,Hyejin Lee,Min Hyuk Park,Jeong Ho Cho... & Suyoun Lee. (2024). Stochastic artificial neuron based on Ovonic Threshold Switch (OTS) and its applications for Restricted Boltzmann Machine (RBM). Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena115195-115195.

[25] Xinlei Wang,Jielei Chu,Hua Yu,Zhiguo Gong & Tianrui Li. (2024). Self-supervised Gaussian Restricted Boltzmann Machine via joint contrastive representation and contrastive divergence. Knowledge-Based Systems112121-112121.