Abstract - In order to ensure the safe and stable operation of power system, enrich the means of power grid analysis and control, and expand the application of deep intelligent learning methods in power grid systems, the application of deep learning intelligent machine learning method in frequency prediction of large power grid is explored. First, on the basis of deep learning, the frequency response mode of large power grid is analyzed and the key characteristic quantities that affect the frequency response mode are extracted. Second, the deep belief neural network (DBN, DNN) frequency prediction model is constructed. Also, the training and testing of the model are introduced. Finally, the input and output based on the DBN.CNN prediction model and the network structure design of the model are analyzed. The prediction performance of the model is evaluated. The results show that when the number of neurons in the hidden layer is 50, the model achieves the optimal prediction effect. Increasing the number of training samples helps to improve the modeling ability and prediction accuracy of the model. For frequency prediction problems, the number of training samples should be set to ≥400, and the number of hidden layers corresponding to the model should be 5. When the number of hidden layer neurons is 10, the prediction accuracy of the DBN/DNN network is poor. When the number of hidden layer neurons is 50, the model can achieve the best prediction effect. Overall, the DBN.DNN prediction model has good prediction performance. The RMSE of the forecast data is 0.0073Hz can basically meet the actual application requirements. Therefore, the frequency prediction method based on deep belief neural network has certain advantages in accuracy and efficiency.

Keywords - in-depth learning, power system, post-disturbance frequency, prediction
networks (ANN), decision tree and supporting vector machine (SVM) are restricted in feature extraction, which can't solve complex problems very well [3-8].

The power system is a very complex and highly nonlinear dynamic system, and its frequency calculation problem involves complex high-dimensional differential algebraic equations [9]. In addition, the traditional BP neural network is prone to local optimization in training, which hinders its in-depth study [10-12]. In recent years, the rapid development of deep learning provides a new direction and train of thought for frequency dynamic analysis of power system. As an intelligent machine learning method, compared with shallow learning, deep learning has more powerful feature extraction ability and can mine deep abstract features of data. Deep learning has an important application prospect in power system analysis and evaluation. It is especially suitable for the analysis of dynamic behaviors with weak transient characteristics such as frequency response modes.

In summary, based on the deep belief neural network (DBN. DNN) to predict and analyze the frequency response pattern after large disturbance, this paper designs the structure of deep learning network adapted to the analysis of large power grid and develops the related algorithms. Through the construction of the frequency prediction model based on DBN.CNN and the design of the model input and output as well as model network structure, the MATLAB2014a platform is used to compare and analyze the hidden layer and hidden layer neurons of the prediction model. Also, the prediction performance of the model is analyzed and discussed. It is aimed to conduct the rapid and accurate prediction and analysis of system disturbance frequency dynamics is of great theoretical and practical significance to enrich the power network analysis and control methods.

II. RESEARCH METHOD

Basic ideas for deep learning

Deep learning DL is a sub-domain of machine learning, also called deep machine learning and deep structure learning. It is an effective algorithm for high-level abstract modeling of data through multiple nonlinear transformations. Its core idea is to take the output of the upper layer of the stacked multi-layer structure as the input of the next layer. Complete the extraction of high-level features from the original data. In general, deep learning is a kind of machine learning algorithm which can effectively train deep neural networks, and a typical deep learning model is a very deep neural network. It is important to note that deep learning is an algorithm, not a network structure. Deep learning uses hierarchical structures to simulate the intricate relationships between data, Deep learning uses hierarchical structures to simulate the intricate relationships between data, which use multiple layers or stages of nonlinear information processing to extract, classify, and predict supervised or unsupervised features to interpret large amounts of text data such as images, sounds, and text. This so-called hierarchical structure is called deep structure. Deep learning networks can represent more complex functional relationships with fewer parameters than shallow networks. Based on a large number of training sample data and building models with multiple hidden layers, learning to obtain more useful features to improve the accuracy of model classification or prediction is the essence of deep learning. Therefore, the Deep Model is a means to improve the accuracy of model classification or prediction. Feature Learning is the purpose.

The problem of frequency prediction is basically a regression problem. The pure deep belief network (DBN) only carries on the feature learning in the training and the operation, cannot carry on the decision-making such as forecast, the classification. Therefore, it is necessary to extend DBN to Deep belief Neural Network (DBN-DNN). The structure of the deep belief neural network is as follows: taking the DBN-DNN with three hidden layers as an example, the network is composed of three restricted Boltzmann machine (RBM) units, in which the RBM-is composed of two layers, the upper layer is the hidden layer and the lower layer is the explicit layer. The output layer of the previous RBM is also the input layer of the next RBM cell, stacking in turn [9]. The basic structure composed of RBM 1, RBM 2 and RBM 3 is called DBN structure (deep belief network structure). Finally, a standard DBN-DNN structure is formed by adding a layer of output layer.
Construction of Frequency Prediction Model of Deep belief Neural Network

In this paper, the large disturbance frequency response database is used to study and predict. Firstly, the training data and test data of the original data samples are applied. The learning of the deep belief neural network model, so that the model can extract and model the latent main features of the original data. The test data is used to predict the performance of the trained prediction model. Figure 5.2 shows the construction process of the whole prediction model. The construction process of the prediction model can be divided into two stages: the training stage and the test stage.

The concrete steps are as follows:

(1) in order to eliminate the influence of different quantity levels on the training and testing of the network model and speed up the convergence of the training network, the data are normalized in this paper because of the obvious order of magnitude difference in the data samples. So that the data range is [0, 1], where $X_{\text{max}}$ and $X_{\text{min}}$ is the maximum and minimum values of the input data respectively, the size relationship between the data does not change after normalization. The general formula for data normalization is as follows:

$$X = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$
(2) Part of the training samples were taken as training data samples and the rest as test data samples. The training samples were subdivided into labeled training samples and unlabeled training samples. For labeled training samples and test samples, the data includes input vectors (eigenvectors before and after system disturbance) and output vectors (frequency response mode variables), and only input vectors are selected for unlabeled training samples.

(3) According to the input vector of all training samples, the model is pre-trained according to the bottom-up approach. For layer $k$, the parameter space is $\left( W^k, b^k \right)$ constructed from the data of layer $K-1$. The activation probability radical formula (2) and formula (3) of the neurons in each layer are calculated, and the base formula (4) is used to update the weights.

\[
p(h_j = 1|v, \theta) = \sigma \left( b_j + \sum_i \Delta w_{ij}v_i \right)
\]

\[
p(v_j = 1|h, \theta) = \sigma \left( a_i + \sum_j \Delta w_{ij}h_j \right)
\]

\[
\Delta w_{ij} = \Delta w_{ij} + p(h_j = 1|g^{(i)}, \theta) \cdot g^{(i)}_j - p(h_j = 1|g^{(i)}, \theta) \cdot g^{(i)}_j
\]

(4) Input vector and output vector data of labeled sample data set are used as input and output of DBN.DNN model after pre-training, respectively. According to the way from top to bottom, the gradient descent method is used to adjust the whole model, and the square reconstruction error is chosen as the cost function in this step.

(5) The test data are outputted by testing DBN.DNN after training with test sample data set and the test data are inverse normalized to get the complete frequency prediction value.

(6) Comparing the output vector in the test sample data and the frequency prediction in (5), the prediction performance of the DBN.DNN frequency prediction model is measured by using the indexes such as the prediction average accuracy and so on.

**Input and output of prediction models**

The input vectors of the prediction model of the depth belief neural network are the sets of some state variables of the system before and after disturbance. Because the frequency response modes of different disturbances are different, the characteristics of frequency prediction are also different. For example, the tangential line after short circuit is equivalent to two disturbances of the system, and the accumulation of unbalanced energy between the two disturbances has a great influence on the maximum frequency shift in the frequency response mode. And this does not exist in the cutting machine and load sudden disturbance. Therefore, only two single disturbances, cutting machine and load mutation, are taken into account in the extraction of input features in this chapter. For the IEEE39 node system, this chapter extends the feature variables studied in Chapter 3, and constructs an original feature sets of frequency prediction, as shown in Table 5-1. The table is the steady-state time of the initial time and the time when the disturbance occurs.

<table>
<thead>
<tr>
<th>Number</th>
<th>Input Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-10</td>
<td>Inertia practice constant values of generators</td>
</tr>
<tr>
<td>11</td>
<td>Coefficient of rotational standby capacity for generators</td>
</tr>
<tr>
<td>12</td>
<td>Maximum active power output value of generators</td>
</tr>
<tr>
<td>13</td>
<td>Frequency regulation factor KL of system load</td>
</tr>
<tr>
<td>14-16</td>
<td>ZIP ratio of system load</td>
</tr>
<tr>
<td>17-26</td>
<td>Active output of generators at $t_0$</td>
</tr>
<tr>
<td>27-36</td>
<td>Electromagnetic Power of generators at $t_f$</td>
</tr>
<tr>
<td>37</td>
<td>Power deficit caused by disturbance at $t_f$</td>
</tr>
</tbody>
</table>

The characteristic parameter 1-26 is the steady state characteristic of the system, which reflects the influence of the operation mode of the system on the frequency response after the disturbance. The characteristic value 27-37 is the characteristic quantity of the system immediately after the disturbance, which reflects the degree of damage to the system balance caused by the disturbance and the reallocation of power between the units after the disturbance. At the same time, it can indirectly reflect the influence of disturbance location on the frequency response mode. The above features indicate the characteristics of the system before and after
the disturbance and complement each other. In actual power grid, the above characteristics can be obtained by the measurement data of the wide area measurement system or by the operation personnel of the power network. Based on the previous research, the output vector of the prediction model of the deep belief neural network, that is, the dynamic frequency response pattern vector, take the average rate of change of the frequency after the time disturbance, the maximum frequency shift \( \Delta f_{\text{max}} \) and the time to reach the maximum shift frequency \( t_m \) as the output Eigenvectors of the prediction model [14-16].

**Design of the Network structure of Prediction Model**

Different depth belief neural network structures are designed. By determining the optimal structural parameters of the model by testing, the optimal performance of the model is achieved.

DBN-DNN is distributed through the stackable structure of RBM and the structural features of the original input data layer by layer. When the number of layers and neurons is too few, the modeling ability and the mining degree of abstract features of original data have no advantage compared with the traditional neural network, but the efficiency of learning be reduced when the number of layers and the number of neurons are too much. It is easy to cause over-fitting problem, and the training time of the model will be significantly increased. Therefore, it is usually necessary to construct the structure of DBN reasonably combined with practical problems. At present, there is no theoretical basis for the selection of optimal values of network parameters. Most of the measures taken in the study are to carry out many experiments and select the optimal parameters among them.

According to the structure of depth belief neural network, DBN-DNN model has stack structure, including input layer, hidden layer and output layer, input layer and one output layer, so it is necessary to determine the number of hidden layer. The number of each unit of input layer, hidden layer and output layer should also be set according to the specific problem. The number of elements in the input layer and the output layer is equal to the dimension of the input vector and the output vector respectively, and the number of hidden layers is not more than 5 layers. The number of input layer units and output layer units of DBN-DNN are designed to be 37 and 3. The number of hidden layer units is selected from the set \{10, 20, 30, 50, 100\}. After the parameters of the underlying structure are determined, a layer of hidden layer is first superimposed on it, After the parameters of the underlying structure are determined, a layer of hidden layer is first superimposed on it, and the number of units in the upper hidden layer is continuously increased from 10 until the performance of the existing model is no longer improved. In this case, a hidden layer is superimposed on the upper layer of the existing model, and the optimum number of the topmost unit in the superimposed structure is searched. If the performance of the superimposed structure is not improved compared with that the implicit layer is not increased, the layer is deleted, and the implicit layer is no longer added. Conversely, continue to add hidden layers to the upper layer until the number of hidden layers reaches the limit. Finally, the performance of the model is optimized, and the structural parameters of the model are the best.

### III RESULTS ANALYSIS AND DISCUSSION

**Simulation setup**

Taking the 39 bus system in New England as an example, this section validates and evaluates the effectiveness of the deep belief neural network model in the prediction of frequency response patterns after large disturbances. This section uses the New England 39 bus system as an example to verify and evaluate the effect of the deep-belief neural network model on the prediction of the frequency response pattern of the large-disturbance. The effectiveness of the deep-belief neural network in frequency prediction is also investigated. In order to evaluate the prediction performance of the model, this chapter uses indicators such as the average accuracy of prediction as the evaluation criteria, and the definition and calculation method of MA and other indicators have been given, and the description will not be repeated here. This chapter is based on the MATLAB2014a platform programming and simulation, and the depth learning algorithm involved comes from the Deep Learn Toolbox open source toolkit.

The data set generation method is used to generate 3000 samples. The randomly selected training set and test set account for 70% and 30% of the total sample set respectively. The specific data are shown in Table 5-2. In this section, 37 eigenvalues in Table 5-1 are selected and normalized as the input of the deep learning network.
model, and the normalized frequency response eigenvalues after large disturbance are taken as the output items of the model.

<table>
<thead>
<tr>
<th>Initial Data</th>
<th>Random sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Samples</td>
<td>Characteristic Value</td>
</tr>
<tr>
<td>3000</td>
<td>37/3</td>
</tr>
</tbody>
</table>

###TABLE II Frequency prediction raw data set

###Analysis of network structure parameters

(1) Comparison of different models of hidden layers

In view of the specific problems and data situation, it is necessary to determine the number of hidden layers of the network model first. In this section, the single variable method is adopted to set the number of neurons in each hidden layer to be 10. The momentum and learning rate of unsupervised learning are 0.5 and 0.7 respectively, the number of iterations is 100, the number of iterations of supervised learning is 200, the training sample is 2100, and the test sample is 900. The number of hidden layers varies from 1 to 5, and five kinds of DBN-DNN models are constructed. Compare MA and RMSE by simulation to determine the number of hidden layers of the model, as shown in Table 5-3

<table>
<thead>
<tr>
<th>Model</th>
<th>Hidden layer number</th>
<th>MA</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBN-DNN N 1</td>
<td>1</td>
<td>0.8951</td>
<td>0.1341</td>
</tr>
<tr>
<td>DBN-DNN N 2</td>
<td>2</td>
<td>0.9122</td>
<td>0.1132</td>
</tr>
<tr>
<td>DBN-DNN N 3</td>
<td>3</td>
<td>0.9445</td>
<td>0.1120</td>
</tr>
<tr>
<td>DBN-DNN N 4</td>
<td>4</td>
<td>0.9663</td>
<td>0.1045</td>
</tr>
<tr>
<td>DBN-DNN N 5</td>
<td>5</td>
<td>0.9885</td>
<td>0.1007</td>
</tr>
</tbody>
</table>

Further, in the simulation, the number of training samples is increased from 100 to 1000, the step size is set to 50, the analysis is carried out, and the analysis results are combined with Table 3. Firstly, when the number of hidden layers is 5 layers, the MA is slightly higher than that when the number of hidden layers is 1-4, but its RMSE is a little lower as a whole. When the number of hidden layers is 1 or 2, the MA of the model is lower than that of the model with hidden layer ≥ 3. This shows that when the number of hidden layers is not more than 5 layers, the DBN-DNN network with five hidden layers has the highest degree of fitting to the input data, and the prediction accuracy of the model is relatively optimal. When the number of hidden layers is 1 or 2, the fitting ability of DBN-DNN network to input data is obviously weak. Secondly, with the increase of the number of DBN-DNN training samples, the MA of the hidden layer is increasing when the number of the hidden layers is 5 layers. When the number of training samples is more than 400, the change is more significant. This shows that increasing the number of training samples is helpful to improve the modeling ability and prediction accuracy of the model.

In general, increasing the number of hidden layers can improve the modeling ability of DBN-DNN. At the same
time, when the number of training samples increased, the prediction performance of the model showed an upward trend as a whole. Therefore, for the frequency prediction problem, the number of training samples should be set to \( \geq 400 \), and the corresponding hidden layer number of the model should be 5.

(2) Comparison of different hidden layer neuron models

In order to investigate the effect of the number of neurons in the hidden layer on the prediction performance of the DBN-DNN network, the single variable method is adopted in this section, the number of hidden is fixed at 5. In order to investigate the effect of the number of neurons in the hidden layer on the prediction performance of the DBN-DNN network, the single variable method is adopted in this section, the number of layers fixed in the hidden layer is 5 layers, and the number of neurons in the hidden layer is set to 10, 20, 30, 50, 100, respectively, forming 5 kinds of DBN-DNN models. The momentum and learning rate of unsupervised learning are 0.5 and 0.7 respectively, the number of iterations is 100, the number of iterations of supervised learning is 200, the number of training samples is 2100, and the number of test samples is 900. The simulation compares MA and RMSE to determine the number of hidden layer neurons in the model, as shown in Table 5-4:

<table>
<thead>
<tr>
<th>Model</th>
<th>Hidden layer number</th>
<th>MA</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBN-DNN N-10</td>
<td>10</td>
<td>0.8521</td>
<td>0.1551</td>
</tr>
<tr>
<td>DBN-DNN N-20</td>
<td>20</td>
<td>0.9263</td>
<td>0.1113</td>
</tr>
<tr>
<td>DBN-DNN N-30</td>
<td>30</td>
<td>0.9776</td>
<td>0.1122</td>
</tr>
<tr>
<td>DBN-DNN N-50</td>
<td>50</td>
<td>0.9904</td>
<td>0.1105</td>
</tr>
<tr>
<td>DBN-DNN N-100</td>
<td>100</td>
<td>0.9831</td>
<td>0.1137</td>
</tr>
</tbody>
</table>

Furthermore, in the experiment, the number of training samples was set to increase from 100 to 1000 in step of 50. The results of the analysis combined with Table 4 show that with the increase of the number of neurons in the hidden layer, the total MA increases and the RMSE decreases obviously. Especially when the number of neurons in the hidden layer is only 10, the performance of both MA and RMSE is better when the number of neurons is 30 or 50. This shows that when the number of hidden layer neurons is 10, and the prediction accuracy of DBN / DNN network is poor.

Secondly, it should be noted that when the number of neurons in the hidden layer is fixed at 100, the overall performance of MA and RMSE is better when the number of neurons in the hidden layer is 50. This is mainly because the training sample input feature dimension is small in the prediction of the frequency of this paper, and there is a big difference between the input feature dimension and the number of neurons in the hidden layer when the number of neurons in the hidden layer is 100. The DBN-DNN network has a lot of useless data when fitting and extracting the features of the input data, which reduces the prediction accuracy of the model. This effect will weaken slightly as the number of training samples increases.
As can be seen from the above, the number of hidden layer neurons is not as much as possible, and the choice and determination must be made according to the actual problems. The number of hidden layer neurons and the problem of overfitting and underfitting will result in a decrease in the prediction performance of the DBN-DNN network, both resulting in a decline in the predictive performance of the DBN-DNN network. Moreover, with the increase of the number of neurons in the hidden layer, the learning time of the DBN-DNN network will continue to increase, which is not conducive to the prediction efficiency of the model. In the frequency prediction problem in this paper, when the number of neurons in the hidden layer is 50, the model can achieve the optimal prediction effect. This section does not discuss the circumstances in which the specific combinations of neurons in each layer achieve optimal results.

Model prediction performance analysis

In order to further verify the DBN-DNN prediction model, the prediction results of DBN-DNN model and other shallow neural network models are compared and analyzed in this section, aiming at the problem of system frequency prediction after disturbance. The validity and superiority of the DBN-DNN network model are verified.

Firstly, three models are constructed, including DBN-DNN model, DNN (Deep Neural Network) model and BPNN (BP Neural Network) model. According to the method of determining network parameters in section 5.3.3, the DBN-DNN model is set to seven layers, including one layer of input and output and five layers of hidden layer. The number of neurons at the bottom-up level is set in the order of 37-50-30-20-10-10-3, that is, the dimension of the predicted vector is 3. The number of iterations of unsupervised learning is 100, momentum is 0.5, learning rate is 0.7. The gradient descent method is used for supervised learning, and the Number of iterations is 200. The structure of DNN model is the same as that of DBN-DNN, but the unsupervised pre-training calculation is not carried out. The number of hidden layers in BPNN model is set to 1, the number of neurons in hidden layer is set to 20. DNN and BPNN are trained by gradient descent method, and the number of iterations is 200. Considering the randomness of the prediction performance of each model, in the total sample 3000, 70% of the total sample is randomly selected as the training sample, 30% of the sample is taken as the test sample, and 5 samples are sampled. The above three prediction methods are simulated and tested respectively.

The results are shown in Table 5.5. Each index in the table is the average of the five simulation test results.

<table>
<thead>
<tr>
<th>Predictive Model</th>
<th>Indexes</th>
<th>MA</th>
<th>MRE</th>
<th>RMS E</th>
<th>Time used/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBN-DNN</td>
<td></td>
<td>0.9774</td>
<td>0.0415</td>
<td>0.104</td>
<td>3310</td>
</tr>
<tr>
<td>DNN</td>
<td></td>
<td>0.9066</td>
<td>0.0866</td>
<td>0.121</td>
<td>11512</td>
</tr>
<tr>
<td>BPNN</td>
<td></td>
<td>0.9113</td>
<td>0.0755</td>
<td>0.120</td>
<td>7915</td>
</tr>
</tbody>
</table>

According to the evaluation index analysis of the model output in Table 5.5, the DBN-DNN prediction model shows superior prediction performance. Although DBN-DNN and DNN have the same structure, network parameters and supervised training process parameters, the prediction accuracy of the former is 7.09% higher than that of the latter. This shows that DBN-DNN provides a good initial value for network parameters through unsupervised pre-training stage and improves the prediction accuracy of the model. The initial value of the model parameters of DNN is provided by randomization method, and then the training is started directly on this basis, which leads to the low optimization degree of network parameters, and then reduces its prediction accuracy. In addition, the prediction accuracy of the BPNN model is 6.61% lower than that of the DBN-DNN model, which shows that compared with the shallow structure, the deep network architecture has a deeper degree of mining data features and better prediction performance. At the same time, in terms of MRE, RMSE and simulation time-consuming three evaluation indexes, DBN-DNN prediction model is superior to the other two prediction models, and a smaller mean square error RMSE indicates that the prediction error dispersion degree of the model is low. Therefore, it can be concluded that the prediction stability of the DBN-DNN model is strong. It can be seen that compared with DNN and
BPNN prediction model, DBN-DNN prediction model has better prediction performance. 900 test samples were selected and the maximum frequency shift of the frequency transient simulation results in each test sample was taken as the reference value. The absolute error AE of the maximum frequency offset predicted by using the input vector of the test sample as the input of the DBN-DNN prediction model is in the range of ±0.025Hz, and the RMSE of the calculated prediction data is 0.00073Hz, which can basically meet the needs of practical applications. Therefore, the prediction results verify the effectiveness and practicability of the proposed frequency prediction model.

IV. CONCLUSION

Based on the deep belief neural network (DBN-DNN), this paper presents a scheme for the prediction and analysis of post-disturbance frequency response patterns. First, the structure of deep learning network adapted to the analysis of large power grid is designed, and the related algorithms are developed. Secondly, the frequency prediction process and prediction model based on DBN-DNN are constructed, and the evaluation index of the model performance is established. Finally, the influence of the change of network structure parameters on the prediction performance of the model is analyzed by simulation. The results show that the increase of the number of hidden layers and single layer neurons can improve the modeling ability of DBN-DNN network. However, if the number of layers and neurons is too many, it will lead to the problem of over-fitting, and the modeling ability will be decreased. Compared with the prediction effect of the shallow learning model, the frequency prediction method based on the deep belief neural network has some advantages in both precision and efficiency.

Compared with the shallow structure, the deep network structure proposed in this paper can realize deeper mining of data features, and it has better prediction performance. It avoids the limitations of traditional machine learning methods in feature extraction. It can quickly and accurately analyze and predict the state of the power system. However, in the investigation of hidden layer neurons, due to various factors, the interaction of each layer of neurons has not been considered. It will be the direction of further improvement and development in the future.

ACKNOWLEDGMENT

Science Foundations of Nanjing Institute of Technology (No. YKJ201421, No. JCYJ201615, No. YKJ201420, No. YKJ201423)

REFERENCES


