Apply to the Micro Grid Data Limit Learning Algorithm

Wenxing Jin, Haiming Li, Guixue Cheng

Institute of Computer Science and Technology, Shanghai Electric Power University, Shanghai
Email: zjxulhm@163.com

Abstract

More and more microgrid projects are put into operation and completed, and the load data are becoming more and more multidimensional and massive. This requires effective classification of load data. Most of the traditional processing methods are based on neural network to classify the grid data. However, with the development of microgrid, the traditional neural network algorithm is having a hard time meeting the requirement of the classification and operation of massive microgrid data. In this paper, the back propagation neural network (BPNN) algorithm is parallelized based on the traditional reverse neural network algorithm. Multiple algorithms are applied for data learning, for example, the combined application of extreme learning algorithm and simulated annealing algorithm, artificial fish swarm algorithm and other evolutionary algorithms. The input variables in BPNN are optimized in the network training process. After adding the algorithm fitness evaluation function, the combined algorithm of improved back propagation neural network algorithm came out. It is most in line with the real-time data of power grid by means of root mean square error. This result could provide data support and theoretical basis for load management, microgrid optimization, energy storage management and electricity price modeling of microgrid.

Keywords

Extreme Learning Machine, Micro-Grid, Auxiliary Decision-Making, Merit-Based Mechanism, Power Load Classification

1. Introduction

In recent years, microgrids combined with new energy technologies have been increasingly applied to our daily lives, and power load information has also begun to show an increase in the proportion of new energy sources, diversified

power loads, huge amount of power information and inaccurate load classification. Therefore, it is even more necessary to use existing shallow data samples to efficiently and accurately classify different loads. In order to find the best power consumption mode for various users, and find the best operation point for microgrid operation. Power modeling for power response, energy storage management is essential.

At present, decision-making and calculation in the field of power load have experienced advances from original expert analysis and decision-making, to neural network applications, and then to improved artificial intelligence algorithms that combine neural networks with multiple clustering methods. The involved fields also began to involve high-voltage fault diagnosis, power load prediction and classification, power management decision-making and other fields. According to the power load classification literature, He Min used the Hadoop platform to classify the power load characteristics [1]; Qiu Wenyuan also used the power characteristic indicators to classify the power load [2]; Huang Qiyuan also analyzed the microgrid based on the fuzzy clustering algorithm. The load data is uniformly divided [3]. However, it should be noted that the above-mentioned power load classification methods all have the most applicable power energy fields. For the situation of variable load conditions caused by the use of new energy in the microgrid, the use of new energy is unstable. Generally, a single method cannot be used to systematically solve all microgrid operating conditions [4]. As a result, in many cases, the complete power algorithm also leads to inaccurate real-time models of the power grid, large errors, and the impact of power decision-making, thereby affecting the overall operation and decision-making of the microgrid.

In this paper, an improved extreme learning algorithm for microgrid mass data is proposed. It aims to use artificial intelligence to classify loads in parallel for different real-time power loads of microgrids, and to divide the shallow microgrid data structure. The feature combination forms a high-level data structure, and the clustered data is used for parallel computing and classification of extreme learning algorithms. Finally, the most suitable real-time microgrid data classification decision-making method is found through parallel multiple data in the form of fitness preference.

2. Overview of the Improved VELM Theoretical Model

ELM (Extreme Learning Machine) or "Extra-Limited Learning Machine" is a type of fast learning method based on single hidden layer feedforward neural networks (SLFNs) [5] based on feedforward neural networks in recent years. As with other neural network-based algorithms, it has the characteristics of randomly selecting the parameters of SLFNs hidden layer nodes and the corresponding nodes, which is suitable for unsupervised learning problems. In specific practical problems, because of the randomness of hidden layer nodes, the entire algorithm runs extremely fast [6]. For the massive data of the microgrid, it has

strong computational advantages and has the characteristics of good network generalization. In the research of microgrid, we often regard it as an extension of neural network algorithm, or as an improvement of reverse neural network.

In the original microgrid data processing, all the weights and other important parameters involved in the neural network algorithm model need to be set and modified manually [7]. There is a strong correlation between different parameters, and the requirements for manually set parameters are very high. Gradient descent algorithm, simulated annealing algorithm and artificial fish swarm algorithm are often used to determine the optimal solution [8]. However, if these algorithms are used alone in practice, they often fall into the situation of falling into the local optimal solution, and the learning speed is often unable to keep up. These are two problems that can not be solved by the traditional algorithm, which greatly limits the universality of the algorithm [9].

In order to solve these problems and solve the problem of using feedforward neural network to accurately learn multiple heterogeneous input vectors in a short time, Huang and other researchers proposed elm algorithm. Elm algorithm randomly generates hidden layer nodes and corresponding node parameters of feedforward neural network in the operation calculation of feedforward network. The original solving process is transformed from a nonlinear model to a linear model, and the output weight is generated by using the least square method. Compared with other gradient descent algorithms, these methods can effectively improve the computation speed without interfering with the whole algorithm operation process. It has huge advantages when dealing with massive data.

Based on the ELM algorithm, this paper proposes an optimized decision extreme learning machine (VLEM) algorithm, which combines the ELM algorithm with the simulated annealing algorithm and the artificial fish school algorithm to ensure the speed and adaptability of the operation. After data is analyzed, fitness evaluation function is applied to improve the entire algorithm structure and computing performance.

3. ELM Algorithm

In the ELM algorithm, the overall microgrid data is now divided into N parts (A1, T1) according to the collection category, where A1 contains all the data of a category in the microdot network data (such as voltage collected at different times or different current collected at all times) [10]. After the data classification, the ELM model with the activation function of X hidden layer nodes according to the feedforward neural network can be written as the showing in formula 1.

$$y(a_j) = \sum_{i=1}^{\bar{L}} \beta_i g(W_i \cdot A_j + b_i) = o_j, \quad j = 1, \cdots, N \tag{1}$$

The values of each variable are as the showing in formula 2, 3 and 4

$$A_i = [a_{i1}, a_{i2}, \cdots, a_{in}]^{\mathrm{T}} \in R^n \tag{2}$$

$$t_i = [t_{i1}, t_{i2}, \cdots, t_{in}]^{\mathrm{T}} \in R^m \tag{3}$$

$$W_i = [w_{i1}, w_{i2}, \cdots, w_{in}]^{\mathrm{T}} \in R^k \tag{4}$$

In the formula, β_i is the output weight, b_i is the offset of the hidden layer unit of the feedforward neural network, $g(a)$ is the activation function of the neural network, W_i is the input weight, A_j is the transposition of the input data vector, $W_i * A_j$ is the internal product of W_i and A_j.

The goal of the algorithm is to make the error of the training samples of the feedforward neural network as close to 0 as possible. As the showing in formula (5)

$$\sum_{j=1}^{\bar{L}} \beta_i \left\| o_j - t_j \right\| = 0 \tag{5}$$

with the existence of β_i, W_i, b_i, then:

$$\sum_{i=1}^{\bar{L}} \beta_i g \left(W_i \cdot A_j + b_i \right) = t_j, j = 1, \cdots, N \tag{6}$$

The matrix form of the above formula:

$$H\beta = T \tag{7}$$

H is the output matrix of the hidden node, β is the output weight matrix, and T is the expected output matrix, as the showing in formula 8 and 9

$$H(W_1, \cdots, W_L, b_1, \cdots, b_L, A_1, \cdots, A_L)$$
$$= \begin{bmatrix} g(W_1 \cdot A_1 + b_1) & \cdots & g(W_L \cdot A_1 + b_L) \\ \vdots & & \vdots \\ g(W_1 \cdot A_L + b_1) & \cdots & g(W_L \cdot A_L + b_L) \end{bmatrix}_{L \times L} \tag{8}$$

$$\beta = \begin{bmatrix} \beta_1^{\mathrm{T}} \\ \vdots \\ \beta_L^{\mathrm{T}} \end{bmatrix}_{l \times m}, \quad T = \begin{bmatrix} T_1^{\mathrm{T}} \\ \vdots \\ T_L^{\mathrm{T}} \end{bmatrix}_{l \times m} \tag{9}$$

ELM decision method:

In the process of ELM algorithm, it is necessary to get the $\tilde{\beta}_i$, $\tilde{b}_i$, $\tilde{W}_i$ of feedforward neural network. And causes:

$$\left\| H\left(\tilde{W}_i, \tilde{b}_i\right) - T \right\| = \min_{w, b, \beta} \left\| H\left(\tilde{W}_i, \tilde{b}_i\right) - T \right\| \tag{10}$$

This is equivalent to the minimum loss function:
The minimum loss function:

$$E = \sum_{i=1}^{L} \left(\sum_{i=1}^{\bar{L}} \beta_i g \left(W_i \cdot A_j + b_i \right) - t_j \right), j = 1, \cdots, N \tag{11}$$

In order to improve the stability and generalization ability of the network, it is often necessary to modify the output results. Given the correction coefficient η, the least square solution of the output weight is:

$$\tilde{B} = H^{\mathrm{T}} \left(HH^{\mathrm{T}} + \eta L \right)^{-1} T \tag{12}$$

The corresponding output function of ELM is:

$$y(a) = h(a)\tilde{B} \tag{13}$$

Compared with the deformation algorithm (such as gradient descent algorithm) of feed-forward neural network, the ELM algorithm determines the output matrix H uniquely once the input weight W_i and the bias b_i are randomly determined. It has a certain degree of advancement [11].

4. The Program to Optimize VELM

4.1. Add Data Characteristics

The initial algorithm input belongs to the shallow data itself, and the data complexity is poor. If there is a slight error in the data monitoring, it will be gradually enlarged in the model, which will affect the operation of the overall model. Therefore, it is necessary to perform feature engineering first to remove all data. Data is selected for feature selection. First, data binning technology is used to preprocess the data to reduce the impact of minor observation errors. Then use one-hot encoding technology to turn the category data into features of the same length. Through feature scaling and standardization operations, the values measured on different scales of data features are adjusted to a conceptual common scale, and independent variables or data feature ranges are standardized [12].

Then we need to build the model with the feature information obtained from the feature engineering. The model is a multi-objective evaluation model combining the voltage, current, electricity, temperature and their corresponding change rate. We build a multi-objective evaluation model based on Pareto optimal solution, NSGA-II algorithm model, which is a fast non dominated multi-objective evaluation model with elite retention strategy, to evaluate the efficiency and quality of the whole system. After data processing, elm algorithm, simulated annealing algorithm and artificial fish swarm algorithm are used to optimize the feature parameters. Each algorithm can get an evaluation value of a system. After running a variety of algorithms, analyze each evaluation value, select the most appropriate algorithm to run the system again to optimize the parameters of the model, get the optimal parameters of the system, and update the parameters.

4.2. VELM Method and Its Fitness Optimization

Based on the analysis of the above algorithm, when A_j and the correction coefficient η was put into the original ELM algorithm, the final result is always affected. We improve the original LEM algorithm and add fitness selection rules. After the calculation data is calculated through the combination of simulated annealing and artificial fish school algorithm, we add the optimal fitness selection, and select the most suitable algorithm for combining with VLEM at the current stage [13].

VELM determination method:

In the improved VELM algorithm, in order to reduce the impact of input va-

riable A and correction coefficient η on the output results, the fitness evaluation optimization mechanism is added as the showing in formula (14) and (15)

$$C_{\text{Adaptability evaluation function}} = \frac{1}{1 + E(y, y_d)} \tag{14}$$

$$E(y, y_d) = \sqrt{\frac{1}{N} \sum_{j=1}^{L} \left[y(c) - y_d(c) \right]^2} \tag{15}$$

$E(y, y_d)$ is the root mean square error, $y(c)$ is the predicted output of the model, and $y_d(c)$ is the expected output.

4.3. Combination of VELM Algorithm and Simulated Annealing Algorithm

Based on the combination of simulated annealing algorithm and ELM algorithm:

The generation and acceptance of new solution of simulated annealing algorithm can be divided into four steps as follows:

Step 1: A new solution in the solution space is generated from the current solution by a generating function. That is, a new parameter value to be optimized (new current value I1) is generated from the value of the current parameter to be optimized (such as current I0) through a simple transformation. In order to facilitate subsequent calculation and acceptance and reduce the time-consuming algorithm, select the current new parameter value, a method in which a new solution can be generated by simply transforming the solution.

Step 2: Combine the VLEM algorithm to calculate the objective function difference corresponding to the new solution. That is, the evaluation function value S' of the multi-objective optimization model corresponding to the parameter to be optimized (such as current I1) is calculated. Because the objective function difference is only generated by the transformation part, the calculation of the objective function difference is calculated by the increment $\Delta T = C(S') - C(S)$.

Step 3: Determine whether the new solution is accepted. The judgment is based on an acceptance criterion. The acceptance criterion here is Metropolis criterion: If $\Delta T < 0$, S' should be accepted as the new current solution S. Otherwise S' should be accepted as the new current solution with probability $\exp(-\Delta T / T) S$.

Step 4: When the new solution is determined to be accepted, it should replace the current solution. That is to say, the new parameter value (current I1) is regarded as the current optimal parameter value (current value), which only needs to be realized in the current solution corresponding to the transformation part when the new solution is generated, and the objective function value can be modified at the same time. At this point, the current solution has achieved one iteration. The next round of trials can be started on this basis. When the new solution is judged to be discarded, the next round of trials is continued on the basis of the original current solution. At the end of the experiment, the global optimal parameter value (current value) is obtained.

The simulated annealing algorithm has nothing to do with the initial value, and the solution obtained by the algorithm has nothing to do with the initial solution state S (which is the starting point of the algorithm iteration); the simulated annealing algorithm has asymptotic convergence. Due to the parallelism of the simulated annealing algorithm, characteristic parameters such as current, voltage, power, temperature and their corresponding rate of change can be simulated and optimized to obtain the corresponding optimal value.

Step 5: Comparison of different algorithms based on fitness-based extreme learning device (VELM). Use extreme learning machine based on voting mechanism to classify all data, then use different optimization algorithms to test, and compare this algorithm with other algorithms that use basic methods [14].

4.4. Combination of VLEM Algorithm and Artificial Fish Swarm Algorithm

Research in this section is based on the combination of artificial fish swarm algorithm and VELM algorithm.

The generation and acceptance of new solutions of artificial fish swarm algorithm can be divided into four steps as follows:

Step 1: The artificial fish school algorithm and VLEM algorithm are combined to initialize the population. Combining data processing such as voltage and current, randomly distribute the values in the experimental area, and set four behavior modes: foraging, rear-end collision, clustering and random behavior. Among them, the step size is also randomly set by the VLEM algorithm. Thereby reducing computing time and increasing algorithm generalization.

Step 2: Combine the VLEM algorithm to calculate the objective function difference corresponding to the new solution. That is, calculate the evaluation function value S' of the multi-objective optimization model corresponding to the parameter to be optimized (such as the current I1), compare it with the actual value, calculate its relative error, and determine whether to use the local optimal solution trap.

Step 3: Determine whether the new solution is accepted. Judgment should rely on the accuracy required by the experiment. If the accuracy does not meet the standard, one should return to step 1: distribute the values randomly again, and recalculate the optimal solution result. If the accuracy meets the requirements, the new solution is used as the current solution.

Step 4: Comparison of different algorithms based on fitness-based extreme learning device (VELM). One should use extreme learning machine based on voting mechanism to classify all data, then use different optimization algorithms to test, and compare this algorithm with other algorithms that use basic methods [15].

Step diagram as the showing in Figure 1.

5. Experimental Analysis

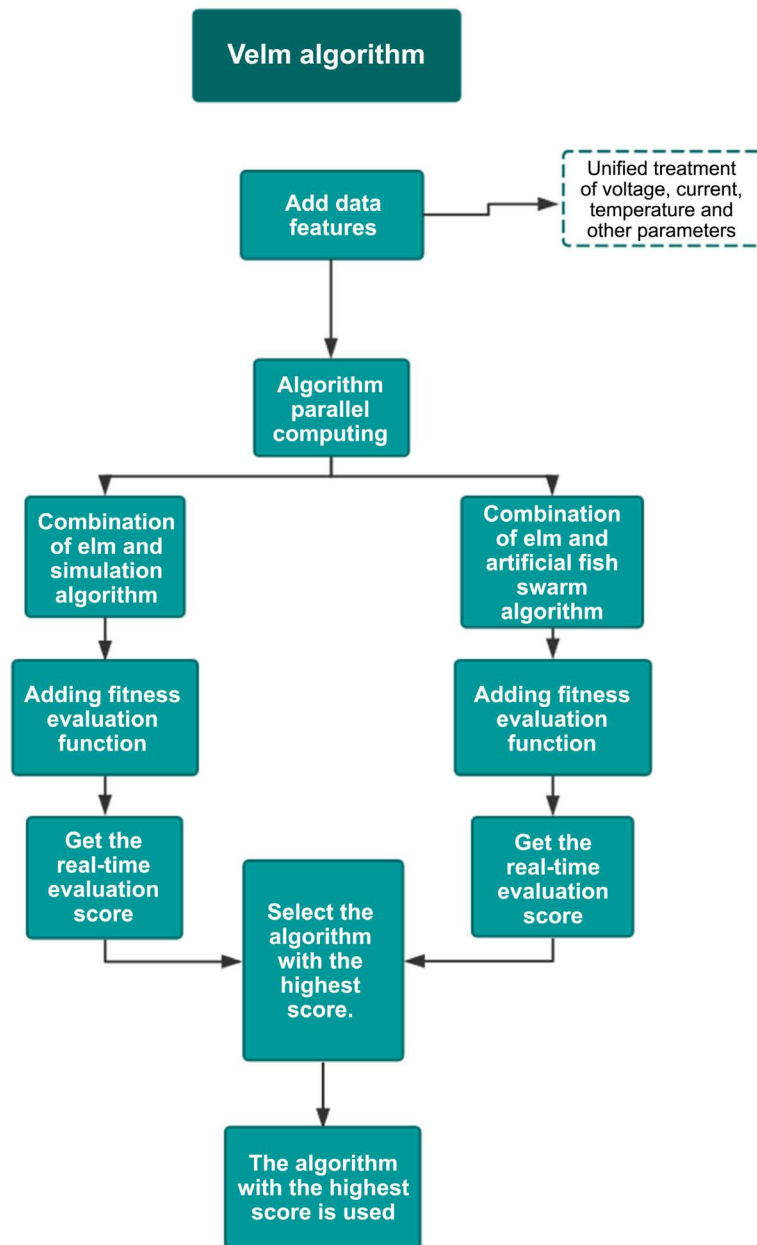In order to test our improved new algorithm, we used the power, voltage and

Velm algorithm

Add data features

Unified treatment of voltage, current, temperature and other parameters

Algorithm parallel computing

Combination of elm and simulation algorithm

Combination of elm and artificial fish swarm algorithm

Adding fitness evaluation function

Adding fitness evaluation function

Get the real-time evaluation score

Select the algorithm with the highest score.

Get the real-time evaluation score

The algorithm with the highest score is used

Figure 1. VLEM algorithm flow chart.

current data of a universal microgrid. 40 days later, we got the results. In order to ensure the consistency of different shallow data to the results, this article first used data binning technology to pre-process the data to reduce the impact of minor observation errors. Then it used one-hot encoding technology to turn the category data into features of the same length. Finally, the three sets of data are combined to form a deep learning sample, which were divided into four parts.

Then, in the neural network operation, the deep learning algorithm that combines VELM with simulated annealing algorithm and artificial fish school algorithm is used to run the data, and the result comparison chart of the data value

and the real value under the VLEM algorithm is obtained. The results of the four samples are compared as the showing in **Figures 2-5**.
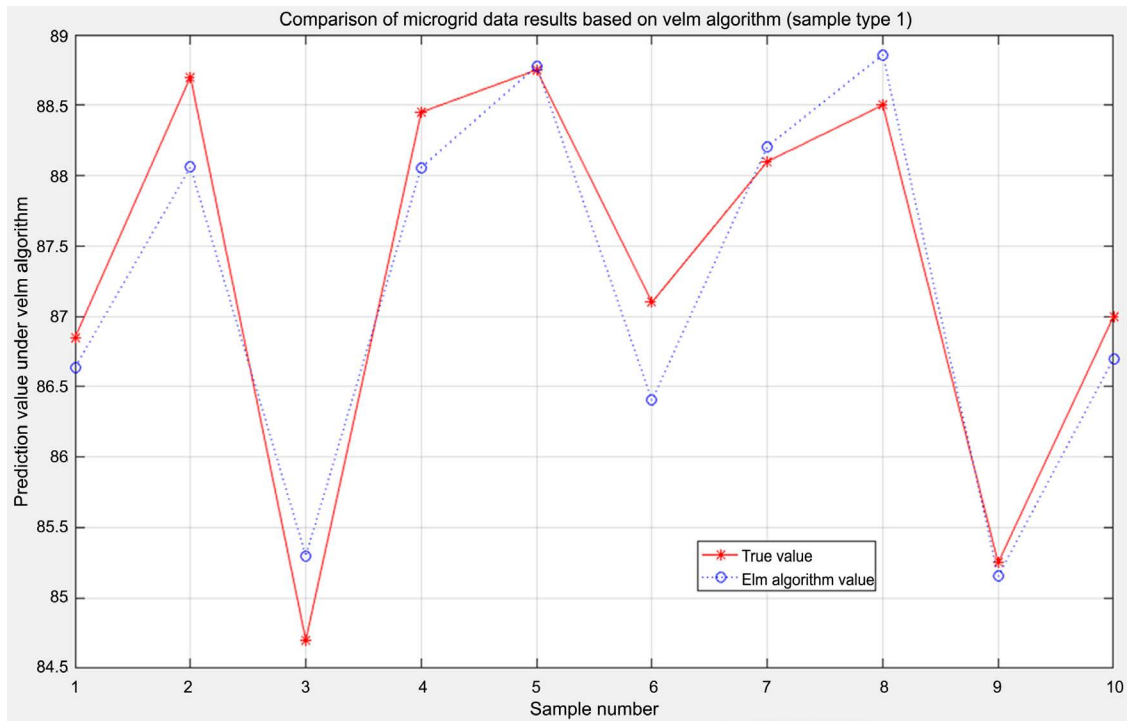
Figure 2. VLEM algorithm data diagram of sample 1.
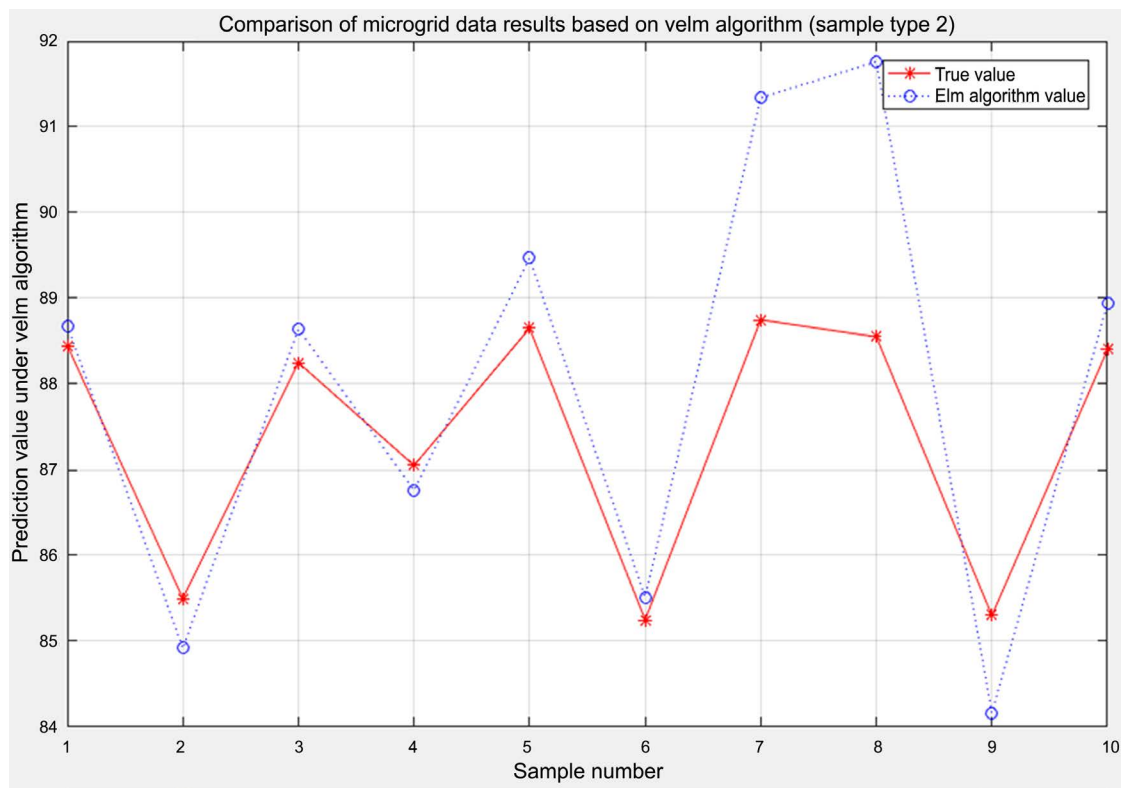
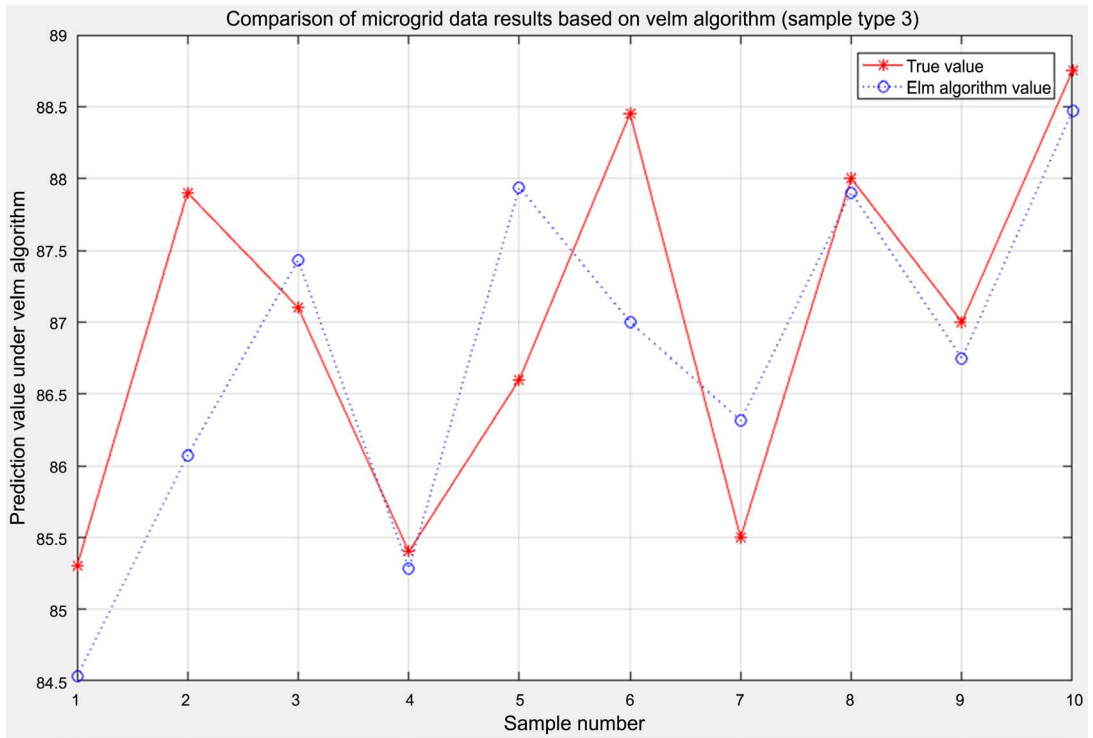**Figure 3.** VLEM algorithm data diagram of sample 2.

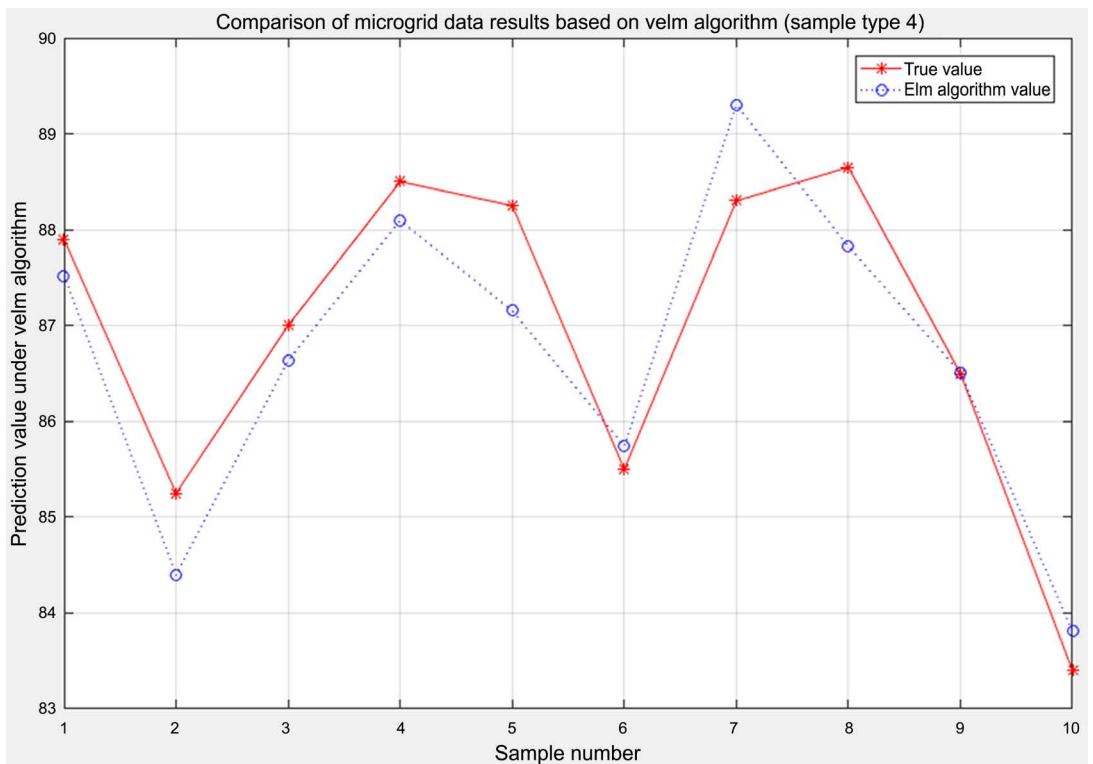**Figure 4.** VLEM algorithm data diagram of sample 3.

Figure 5. VLEM algorithm data diagram of sample 4.

The figures show that there are some differences between the predicted value and the real value under the operation of VLEM algorithm. In order to further

improve the algorithm, adaptive optimization selection mechanism is added at the end of VLEM algorithm to further improve it. The improvement are as the showing in **Figure 6** and **Figure 7**:
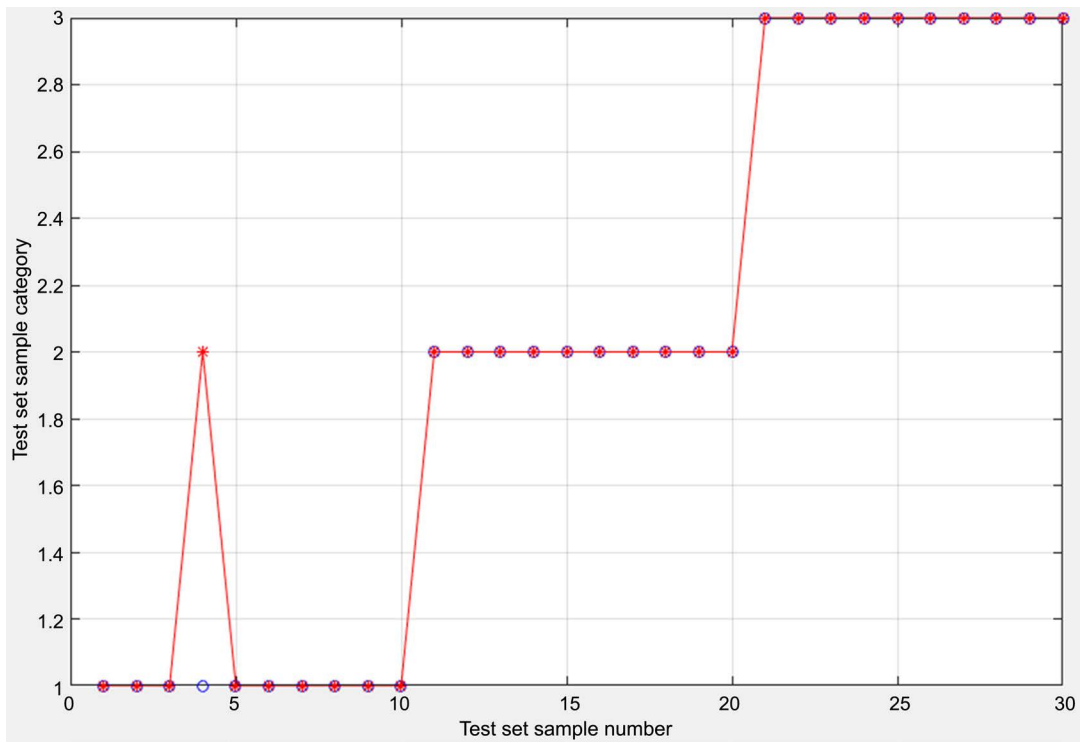
Figure 6. Results of the first test.
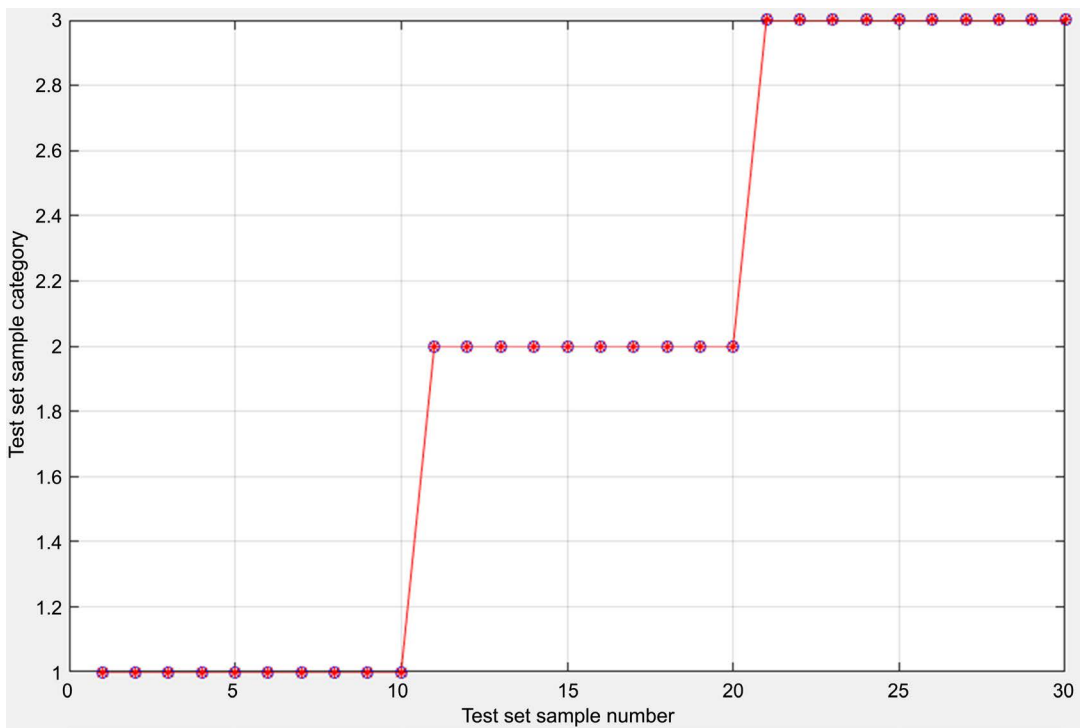
Figure 7. Results of the second test.

(The overall sample were tested twice. Because the implicit function and nodes are randomly determined, even though it is the same sample, the results of the two tests are still slightly different. In the first test, there is a sample point prediction distortion, but the overall accuracy rate has been significantly improved.)

The above figures show that the same set of data, under the VLEM algorithm, one of the first 30 results failed to predict. In the second prediction of the same set of data, all thirty results were successfully predicted. The overall accuracy rate was 98.3%. The operating speed and accuracy have been significantly improved compared with the original ELM algorithm.

6. Conclusion

This paper proposes a new VLEM algorithm for computing massive data in microgrids. The core idea is to randomly determine the implicit function and node, and merge it with the simulated annealing algorithm and artificial fish school algorithm to make adaptive selection. Experiments on the comprehensive data set show that the extreme learning algorithm can better identify the complex patterns in the data set, and is more effective than the existing extreme learning algorithm. And the load data classification of the massive data of the microgrid was carried out, and the task of load classification was effectively completed, indicating that the VELM algorithm has broad application prospects in the field of electric power big data. Future studies on VELM will continue to overcome the shortcomings of this algorithm and apply it to different fields.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] He, M. (2018) Research on Power Load Data Classification Algorithm Based on Hadoop Platform. Xi'an University of Science and Technology, Xi'an.

[2] Qiu, W.Y., Wang, D.H. and Zhang, X.Y. (2015) Study on Classification of Power Load by Characteristic Index. *Electrical Application*, **34**, 512-516.

[3] Huang, Q.Y., Wang, Z.J., Zhu, J., Du, B., Wang, D.W., Lv, J.D. and Wang, H.Q. (2015) Load Classification Analysis of Power System Based on Fuzzy Clustering. *Journal of Electric Power*, **30**, 200-205.

[4] Xu Xuanhao. (2018) Research on Classification and Application of Power Load Form for Demand Side Management. Guangdong University of Technology, Guangzhou.

[5] Zhao, T. (2018) Research on Power Load and Photovoltaic Power Spatial and Temporal Distribution Prediction Method under Big Data Environment of Intelligent Distribution Network. Shanghai Jiaotong University, Shanghai.

[6] Zhang, Q., Zhang, L., Hu, S.J., Fan, P.L. and Ma, Y.W. (2017) Research on Statistical Method of Electric Load of Offshore Nuclear Power Platform. *Shipbuilding Technology*, **37**, 49-50, 55.

[7] Su, H.Z. (2017) Research on Power Load Classification Based on K-Means Cluster-

ing Algorithm. *Electronic World*, No. 23, 78-79.

[8] Yang, X.T. (2017) Research on Spatial Load Forecasting Method Considering Cell Development Degree. Northeast Electric Power University, Jilin.

[9] Peng, B. (2017) Research on Power Load Clustering Analysis and Optimal Management. Zhejiang University, Hangzhou.

[10] Liu, Q.Y. (2016) Analysis on the Importance of Each Influencing Factor in SLF to Prediction Accuracy. *Heilongjiang Science and Technology Information*, No. 29, 71.

[11] Tang, X.M. (2016) Research on Comprehensive Load Modeling Based on Multi-Source Data. Wuhan University, Wuhan.

[12] Wang, Q.J. (2016) Analysis of Power Load Composition and Research on Orderly Power Utilization Countermeasures in Ningbo. North China Electric Power University (Beijing), Beijing.

[13] Yang, Z.N. (2016) Research on Decision Making Method of Orderly Power Utilization. Southeast University, Nanjing.

[14] Wang, X.D., Huang, Z.H., Wu, J., Xue, Y.S., Cui, S. and Li, W.L. (2020) Application of Artificial Intelligence Technology in Power System State Estimation. *Automation and Instrumentation*, No. 3, 179-183.

[15] Liu, Y.P., Bu, Y.N., He, P. and Tian, Y. (2020) Strain Detection and Identification of Transformer Winding Based on S-Transform and Elm. *High Voltage Apparatus*, **56**, 9-17.