A Reliable Ensemble Classification Algorithm by Genetic Neural Network based on Multiple Regression

Xishang Dong, Meili Sun, Ting Zhang, Qiaolian Liu, Weikuan Jia

Abstract—**For complex pattern classification, the neural network has unique advantages. However, when encountering complex data, one single neural network is usually difficult to achieve satisfactory classification accuracy. In this study, two different types of neural network have been selected to design sub-classifier. RBF or Elman neural networks have some defects, which restrict the operating efficiency of classification system. By constructing the genetic neural network, the new optimization way of hybrid encoding and simultaneous evolving is adopted. While being ensemble learning, integrated classification system is modeled by multiple regression (MR), and the output of each sub-classifier is used to establish regression model. OLS theory is used to optimize regression equation, and then the weights of each sub-classifier can be determined. A novel ensemble classification algorithm is established through two consecutive optimizations. In order to verify the reliability of new algorithm, three experiments are arranged. The experimental results show that the generalization ability, operating efficiency and classification precision of new algorithm have been greatly improved.**

Index Terms—**Ensemble learning, Genetic neural network, Multiple regression, Classification, Sub-classifier design**

I. Introduction

PATTERN classification has almost permeated every aspect of our life, in many fields such as biology recognition, economic prediction, weather forecast, industrial and agricultural production has been widely used. With the rapid development of information technology, the continuous advancement of scientific research and the requirements of practical problems, the results of pending problems are becoming more and more precise, which means that pattern classification methods are facing new challenges. As data becomes increasingly complex, classification issues are also becoming increasingly serious. Such as some small samples, incomplete learning will affect the classification precision; complex large sample, repeating learning or over-fitting will lead to reducing the classification accuracy; and so on. The essence of the pattern recognition can be defined as an input and output classification system, while the neural network (NN) can approximate any nonlinear system with arbitrary precision, so neural network can exactly show the advantages to deal with these classification problems. Neural network [1-3] is based on computer network system to simulate the biological neural network with intelligent computing, which shows powerful functions in dealing with nonlinear and large-scale computing.

Giving the complex data, neural networks have achieved a large number of results with relatively satisfactory classification results. However, in practical application, it is a challenge for a single conventional neural network to deal with classification of complex samples. The traditional neural network always exists some inherent defects, and the classification ability of single neural network is relatively low. In order to improve the classification precision for complex data, two optimizing strategies will be adopted. Firstly, a genetic algorithm (GA) is used to optimize the traditional neural network. In terms of network structure and connection weights, hybrid coding and simultaneous evolution optimization methods are adopted. Secondly, ensemble learning is introduced into classification algorithm, and the method of multiple classifiers integrating is adopted.

Hansen and Salamont first proposed the concept of neural network ensemble learning [4], by the experiment, they proved that the performance of an integrated series of neural network is better than the best single neural networks. And the generalization ability and recognition accuracy of meanwhile the integrated system of multiple classifiers has been significantly improved. Soon afterwards, the ensemble learning was raised to a climax. Many scholars [5-9] proposed a series of concepts and algorithms, which have been applied in many fields and have achieved good results. The study of neural network integration systems has two main aspects: one is to design the sub-classifiers, namely generating or selecting the individual neural network; the other is integrating the output of each single neural network,

Manuscript received October 6, 2022; revised revised January 23, 2024.

This work was supported by Natural Science Foundation of Shandong Province in China (No.: ZR2020MF076).

X.S. Dong is an associate professor of School of Information Science and Engineering, Zaozhuang University, Zaozhuang 277160, China (dxs@uzz.edu.cn)

M.L. Sun is a postgraduate student of School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China (sunmeili95@126.com)

T. Zhang is a lecturer of School of Information Science and Engineering, Zaozhuang University, Zaozhuang 277160, China (Corresponding author, phone: +86-632-3785947; fax: +86-632-3785947; e-mail: zhangting0607@126.com)

Q.L. Liu is a lecturer of School of Information Science and Engineering, Zaozhuang University, Zaozhuang 277160, China (Corresponding author, phone: +86-632-3785947; fax: +86-632-3785947; e-mail: llgx1207@126.com)

W.K. Jia is an associate professor of School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China (Corresponding author, phone: +86-531-86181755; fax: +86-531-86181750; e-mail: jwk_1982@163.com)

namely to design an integrator of sub-classifier. The integrated system of two sub-classifiers is selected and used to illustrate the ensemble learning.

Neural networks have been around for over 70 years and hundreds of network models have been proposed. Different network models have their advantages to deal with different problems. However, as far as the neural network itself is concerned, it is not perfect [10]. For this purpose, we design two sub-classifiers for ensemble learning. To design a sub-classifier, we select two different types of neural network: the feed-forward model radial basis function (RBF) neural network [11, 12] and the feedback model Elman regression neural network [13, 14]. The most outstanding advantage of RBF neural networks is that they use linear learning algorithms to do the work previously done by non-linear learning algorithms, while maintaining the high accuracy of non-linear algorithms with best approximation and global optimality. RBF neural network and its improved algorithms have been widely used in the traditional control, recognition, and classification problem [15-17]. However, the learning ability of RBF neural networks is low, and it is difficult to determine the center and width, which ultimately leads to low classification accuracy. Elman regression neural network can be seen as the optimized based on the back propagation (BP) neural network[18-19], which has the ability of adapt to the time-varying characteristics and has strong global stability. Elman neural network has been widely applied in many fields[20], however, it inevitably inherits the inherent defects of BP neural network, such as being easily trapped into local minimum, fixed learning rate, and difficult to determine the number of hidden layer neurons[21], all of these defects will lead to the unsatisfactory operating efficiency and classification precision.

Aiming at the above shortages of traditional neural network, the genetic algorithm [22] is introduced into optimizing neural network algorithm. As the representative of evolutionary computing, genetic algorithm does not need the knowledge based on the problem, thus it shows advantages of solving the complicated problem under the robustness, nonlinear and parallelism situation. It has a broad space that uses genetic algorithm to optimize neural network. Improving the generalization ability, operating efficiency, and classification precision of neural networks is one of the focus research in the intelligent algorithm [23-25]. There are two problems with using genetic algorithm to optimize RBF neural network sub-classifier [26-28]. On the one hand, the optimal number of hidden layer neurons and their center and width will solve the problem of uncertain hidden layer structure; On the other hand, optimizing the hidden layer to output layer weights can solve the defects of low ability of weight learning. Genetic algorithm is used to optimize Elman neural network sub-classifier [29-31] to optimize the number of hidden layer neurons, which will reduce the time consumption of constructing the network. Optimizing connection weights and thresholds will improve the efficiency of online learning and address the drawbacks of lacking a fixed learning rate and the tendency to get stuck in local minima. Above of all, given the GA-NN optimization algorithm, some progress has been achieved. The previous studies mostly devoted to single optimization, for example, connection weights or network architecture achieved good results. However, on the whole, the single optimizing way is slightly weakened, such as connection weights learning is the basis of the fixed network structure; the determination of network structure can hardly do without connection weights learning. Therefore, in this study, for connection weights and network structure (neurons in the hidden layer), we attempt a new optimizing way that combines hybrid encoding and simultaneous evolving, which means acquiring the optimal structure while completing the learning of connection weights concurrently. This optimization strategy will improve the operating efficiency and classification precision of neural network to get the optimal sub-classifier.

The core problem of the ensemble optimizing is to design the integrator for sub-classifier, which determines the final recognition precision of classification system. The key of ensemble learning is how to determine the weights of each sub-classifier, which has attracted many scholars' research interest [32-39] as the hot spot. Zhou proposed simple average or weighted average [40], which is generally believed that the weighted average method has better generalization ability than the simple average, however, it also should prevent over-weighted allocation problem. Jia proposed the integrated classification system of three sub-classifiers for small sample classification problem [41], and to determine their weights by ordinary least squares (OLS). In this study, we attempt a simple and effective method, the integrated system by multiple regression to establish the regression equation between the sub-classifier result and the actual output, by solving the regression equation to determine the weight of each sub-classifier and get the optimal integration system.

To address the recognition problem of complex data, considering the nature of neural networks and ensemble learning, we have established a new ensemble-optimized classification algorithm based on multiple regression using genetic neural networks. Firstly, we have adopted a novel optimization approach, namely hybrid encoding and simultaneous evolution. By constructing a genetic neural network, we enhance the generalization ability and operational efficiency compared to traditional neural networks. The genetic neural network regard as sub-classifier (individual neural network). Here, two sub-classifiers are designed: one is that using new optimized RBF neural network based on genetic algorithm (GA-RBF) to constitute sub-classifier I; the other is that by new GA-Elman as sub-classifier II. With this, the optimal sub-classifiers have been constructed, and the firstly optimization is realized. Secondly, two sub-classifiers are used to establish binary regression equation, and then the theory of ordinary least square is used to solve regression equation, to further determine the weights of two sub-classifiers, so far, the ensemble learning is completed, namely, the secondly optimization is realized. Through these two successive optimizations, a reliable classification system for complex data was established with the aim of improving the operational efficiency and recognition accuracy of the classification system to improve the classification accuracy of complex data.

Two sub-classifiers are used to illustrate neural network ensemble algorithm. In order to validate the effectiveness of ensemble classification algorithm, three groups of experiments are arranged. two groups of experimental

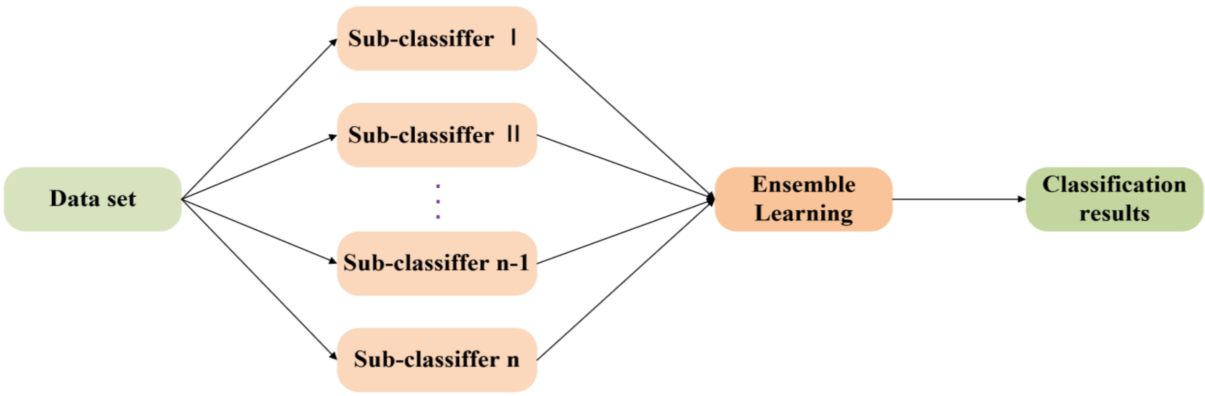

Fig. 1. Thought of ensemble classification system.

data from UCI standard data sets, another data set from practical agricultural production. The experimental results indicate that the ensemble classification algorithm improves the classification precision of complex data. According to the analysis of experimental results, the new classification algorithm has unique superiority, which is worthy of further promotion.

II. Thought of Ensemble Classification Algorithm

In the face of complex data, it is often difficult for a single classifier to get the ideal classification results. In addition, the single classifier restricted by their own defects, classification effect will be worse. For these reasons, we attempt to construct an ensemble classification system to improve the classification precision. The follow chart of new system is shown in figure 1.

From figure 1, firstly, complex data is recognized by different sub-classifiers, and then results of sub-classifiers are ensemble learning, finally, the relatively accurate classification result of complex data is output. In this system, two main aspects are followed with interesting, which are the design of sub-classifiers and the method of ensemble learning. In this classification system, neural network is used to design sub-classifier, namely neural network ensemble system. Aiming at the defects of traditional neural network, genetic algorithm is used to optimize them. For the network structure and connection weights of neural network, a novel optimizing way of hybrid encoding and simultaneous evolving is adopted, which can avoid the deficiency in the process of individually optimizing. In the process of ensemble learning, OLS is used to design the integrator to determine the weights of the different sub-classifiers. For instance, two neural network sub-classifiers are selected, and some relevant studies and discussions are made on the multiple regression-based neural network ensemble learning algorithm.

III. Sub-classifier Design (Individual Neural Network)

In order to eliminate the effects of the above-mentioned defects and obtain better sub-classifiers, we introduce genetic algorithms into neural networks and try to construct genetic neural networks for designing sub-classifiers. In this study, the bran-new evolutionary approach is used, in which hybrid coding and evolution are performed simultaneously.

RBF and Elman neural networks are selected, respectively belong to different types of network models. RBF neural network is a feed forward network model; Elman regression neural network is a kind of feedback network. They are used to design sub-classifiers, to a certain extent, which can play a complementary role.

A. *Construct bran-new genetic neural network*

Genetic algorithms are used to optimize neural networks, which can be regarded as an adaptive system without human intervention, which can automatically adjust the structure and connection weight of the neural network. Genetic algorithm is suitable for the large-scale, non-differentiable, multi-mode space search. It does not need error function gradient information, which has a unique advantage under the situation of obtaining all these details difficultly. The genetic algorithm is used to optimize the connection weights of the neural network. By incorporating some penalty terms in the error function, irrespective of their differentiability, the genetic algorithm aims to enhance the universality and reduce the complexity of the neural network. Consequently, it holds great potential for optimizing connection weights. In terms of optimization of neural network structure, it can be in the face of different tasks to optimize the different topology structure. The basic flow chart of genetic neural network is shown in figure 2.

During the process of design genetic neural network, the biggest problem is how to design a reasonable coding scheme and express the network structure and connection weights; how to define the fitness function and construct genetic operators. This study optimizes two sub-classifiers, adopts the way of the structure and weights hybrid encoding and simultaneous evolving. While designing two sub-classifiers, coding scheme and fitness function will be changed according to the different neural network, here, the same genetic operators are adopted.

Constructing genetic operators of genetic neural network are as follows:

Selection operator uses the "roulette wheel" selection, which means the probability of each individual in the next generation is equal to the ratio between its fitness and the sum of the whole population's fitness, the higher fitness value, and the greater probability getting into next generation. Crossover operator, using "single point crossover", which swaps partially recombinant parental genes with a certain probability. Mutation operator, which uses the binary-code "reverse" mutate and real-encoded "Gauss" mutate.

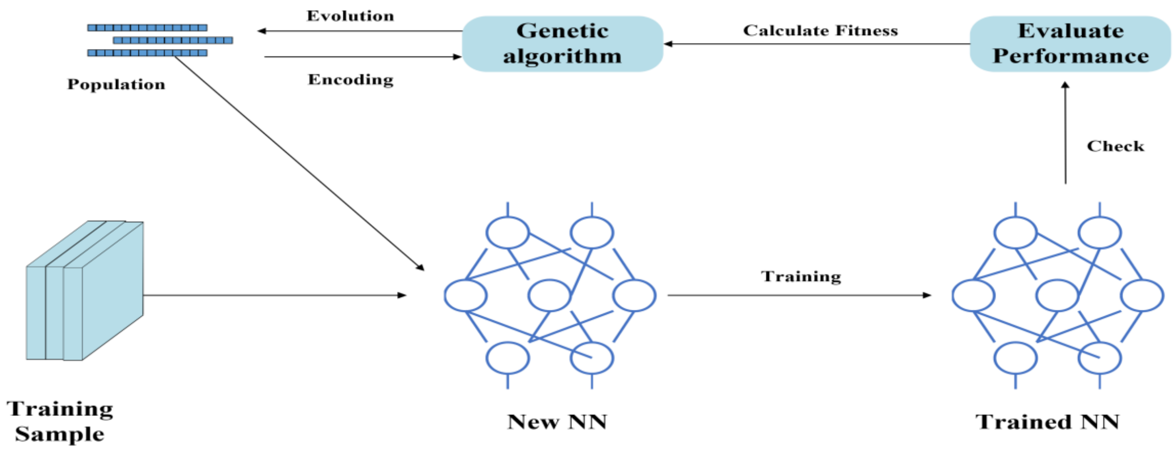

Fig. 2. Flow chart of genetic neural network algorithm.

B. GA-RBF sub-classifier I

RBF neural network can adjust its hidden layer adaptively according to the specific problems in the training process. Hidden layer neurons can be distributed by the capacity, category and distribution of the training sample, which can dynamically determine the number of neurons and its center and width with the faster learning speed. The radial basis function is used as the "base point" of the hidden neurons to form the hidden space. In fact, it is a local distribution of symmetric centers about a nonlinear function. Therefore, once the centers of radial basis functions are determined, they can be directly mapped (without connections) as an input vector to the hidden space. Indeed, the mapping from the hidden space to the output space is linear. This implies that the output of each unit is a sum of linear weights, which are adjustable parameters in the network.

GA-RBF neural network algorithm

It uses genetic algorithms to optimize the neuron centers of the hidden layers of RBF neural networks and their widths and connection weights without prior knowledge. It is insensitive to the initial parameters and does not easily fall into local minima. Previous GA-optimized RBF neural network algorithm is mostly optimizing weights or network structure separately. We provide a new way called hybrid encoding and simultaneous evolution to optimize both the structure and connection weights of the RBF neural network.

GA-RBF algorithm parameters settings:

The main content of the GA-RBF neural network algorithm is chromosome encoding, fitness function defining and genetic operator constructing, which is also one of the biggest challenges of designing genetic optimization.

1) Encoding scheme

Set a maximum number of hidden layer neurons of the RBF neural network is s, number of output layer neurons is m, thus one chromosome coding of the GA algorithm as:

$$c_1 c_2 \cdots c_s w_{11} w_{21} \cdots w_{s1} w_{12} w_{22} \cdots w_{s2} \cdots w_{1m} w_{2m} \cdots w_{sm} \theta_1 \theta_2 \quad (1)$$

The number of hidden layer neurons adopts binary encoding, c_i with a value of 0 or 1. c_i is 1 indicates that the neurons exist; the opposite is 0; s is the upper limit. The connection weights between neurons in hidden layer and output layer neuron use real number coding w_{ij} means that from i^{th} the hidden layer neurons to the j^{th} output layer connection weights of neuron. Output neurons threshold also adopt the way of real number coding θ_j. θ_j is threshold value of j^{th} output layer neurons.

2) Fitness calculating

The original dataset is divided into a training dataset and a test dataset, where the network training error and the network are used to determine the size of the chromosome. The fitness function is:

$$F = \frac{1}{e} \quad (2)$$

Among them, e is for error. It is generally believed that the great error in the network, accordingly, the fitness was low.

Steps of GA-RBF Algorithm

Thus, the basic steps of the GA-RBF optimization algorithm are as follows:

Step 1 Setting RBF neural network, according to the largest number of hidden layer neurons to get the center of the basis function and its width;

Step2 Setting the parameters of the GA algorithm, the population size is Ps, crossover rate is Pc and mutation rate is Pm, selection, crossover and mutation operator, the number of iterations $G = 0$, error objective function is $Emin$, maximum number of iterations is for $Gmax$;

Step 3 Initializing population P at random, coding each individual network according to (5) scheme;

Step4 To decode the code of step 3 and obtain N different networks, then calculate the connection weight from the hidden node to the output node and the fitness of each chromosome, until the number of iterations $G = 0$, $G = G_{max}$, according to the fitness to calculate replicate probability;

Step5 Setting the initial count $k = 1$, to generate offspring with the above genetic operators, using replication probability to select individual two pairs, using the crossover probability to cross of two parents, create two individuals, using mutation probability to mutate some bit of individual;

Step 6 Decoding the two sub-individuals, then calculated the weights between hidden layers and output layer neurons and calculate the fitness of each individual, to compare two sub-individuals with the father individuals, and retain the best;

Step 7 Setting $k = k + 1$, and if $k > Ps$, then turn to step 8, otherwise turn to step 5

Step 8 Setting $G = G + 1$, and if $G = G_{max}$, terminate the algorithm, or turn to step 4.

C. GA-Elman sub-classifier II

Elman regression neural network is a kind of feedback model, which adds an undertake layer into the hidden layer based on the BP model as delay operator. The output of hidden layer through delay and storage by undertake layer, to achieve the memory, so that the system has the ability to adapt to the time-varying dynamic characteristics, and has strong global stability.

GA-Elman neural network

GA-Elman algorithm, for connection weights and structure, also adopts the method of hybrid encoding and evolving simultaneously. The significance of the GA-Elman algorithm lies in optimizing the connection weights and thresholds, thereby improving the training speed and convergence properties. The GA-Elman algorithm also saves the running time and improve the operating efficiency of the neural network. Optimizing the number of hidden layer neurons is for determining the optimal network structure, reducing the time of constructing the neural network structure factitiously and improving the ability of solving problem. The structure of the concealed layer has been determined, as has the taking over layer.

1) Encoding scheme

In the GA-Elman design process, the biggest problem is how to design a reasonable encoding scheme for connection weights, output threshold and the number of hidden layer neurons. So, the way of hybrid encoding and evolving simultaneously is adopted.

Suppose Elman neural network with n inputs and m outputs. Real number encoding is adopted to code the number of hidden layer neurons and the upper limit is defined as p. In addition, a binary code is embedded to the hidden layer neurons for the control of genes that means genes can be generated by random function. When the value of the control gene is 0, the neuron corresponding to the hidden layer to the output layer does not work, otherwise a value of 1 for the control gene corresponds to the neuron from the hidden layer to the output layer. The encoding length is:

$$L = p + n \times p + p \times m + p + m = P \times (n + m + 2) + m \quad (3)$$

Given that the hidden layer of the Elman neural network corresponds to the context layer, the same encoding scheme is employed. Connection weights are mainly adopts real number coding, which represents each weight value is directly with a representation of an intuitive real number, on the other hand, overcome the disadvantages of the original binary code. This kind of encoding scheme is based on the confirmation of the neural network structure, its encoding length is:

$$L = n \times p + p \times m + p + m = P \times (n + m + 1) + m \quad (4)$$

2) Fitness calculating

Here using the neural network training error to determine chromosome fitness with the corresponding network, the fitness function adopts formula (2).

Steps of GA-Elman algorithm

The basic steps of GA-Elman optimization algorithm as follows:

Step1 Initializing the population P at random with size of N and the Elman neural network weights and threshold is unified, and carry on the real number coding;

Step 2 The number of hidden layer neurons code by real number encoding, and set the upper limit, then add a binary encoding of hidden layer neurons to control genes;

Step 3 Decoding the step 1 & 2 coding and get N different neural networks;

Step 4 The network in the step 3 was trained by the given training samples, to judge whether it meets the accuracy requirement or the training is to stop, otherwise be turn to Step 5;

Step 5 According to the results of the objective function and training determine the individual fitness, and choose several individuals with biggest fitness value, which directly inherit to the next generation;

Step 6 Using crossover and mutation operators for the current generation will produce the next generation of group, here$P = newP$, $G = G + 1$;

Step 7 Estimating the value of $G = G_{max}$and the algorithm is suspended, or turned to Step 3.

IV. CONSTRUCT GENETIC NEURAL NETWORK ENSEMBLE SYSTEM

Faced with the classification problems of complex data, generally, a single classifier is difficult to achieve the ideal recognition accuracy and the single classifier has its own deficiencies. Integrating multiple classifier system will can improve the generalization ability and recognition accuracy obviously. The main idea of ensemble learning is to multiple sub-classifiers to solve the same problem, which get the results of system by integrating each sub-classifier input. The purpose is to improve the classification precision and generalization ability of learning algorithm effectively.

A. Thought of genetic neural network ensemble learning

Neural network ensemble learning is integrating a series of single neural network to construct ensemble classification system. Obviously, the performance of the integrated system is better than any single neural network. The main purpose of neural network ensemble learning is to improve the recognition accuracy of the classification system. The weights determination of each sub-classifier is the key to ensemble algorithm. The main task of the neural network ensemble algorithm is based on the characteristics of each sub-classifier is to determine its weights and reduce the classification error of integrated system.

GA-RBF and GA-Elman genetic neural networks are regarded as sub-classifier, where the traditional RBF and Elman neural networks are optimized by genetic algorithm. To a certain extent, GA-RBF and GA-Elman genetic neural networks can overcome its own disadvantages and improve their operating efficiency. However, the classification accuracy is still unsatisfactory, so we focus on designing the integrator for two sub-classifiers of ensemble learning. The integration of genetic neural networks based on the previously established genetic neural networks is used to further model the neural network ensemble learning.

B. Genetic neural network ensemble modelling

The working principle of the integrated system is that given the problem to be addressed, it is trained by GA-RBF neural network (sub-classifier I) and GA-Elman neural network (sub-classifier II). Each sub-classifier can obtain different recognition results.

The output of the sub-classifiers is considered as the input of the integrated system. The integrator then performs weighted learning and provides weights for each sub-classifier. Finally, output the precision recognition results. The follow chart is show as figure 3.

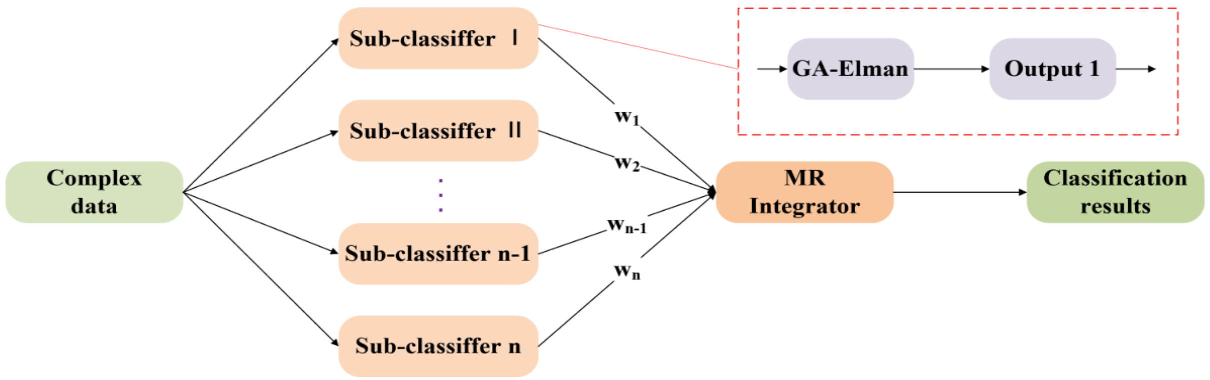

Fig. 3. Optimized classification algorithm assembled with genetic neural network.

The figure 3 illustrates the basic process of genetic neural network integration system. In this study, two sub-classifiers are investigated as examples, sub-classifier I and sub-classifier II using GA-RBF and GA-Elman neural network models, respectively. The sample data to be processed is identified by each sub-classifier. The recognize result of sub-classifier I is A and the recognize result of sub-classifier II is B. And the two sub-classifiers are independent between each other, so the results A and B are also independent between each other.

The weights of two sub-classifiers are defined w_1, w_2 respectively, the output of the integrated system is Y, which is summed by weighted A and B. And then the ensemble learning algorithm model is established.

$$Y = Aw_1 + Bw_2 \tag{5}$$

When $w_1 = 1, w_2 = 0$, it means that only sub-classifier I works and the output result of the system is the recognition result of sub-classifier I. In the similar way, when $w_1 = 0$, $w_2 = 1$, it means that only sub-classifier II works, and the output result of the system is the recognition result of sub-classifier II. The optimal integration system, that is, find out the optimal weights w_1, w_2, which makes the model (5) the output of the recognize results to achieve the most optimal state.

C. Determined weights based on multiple regression

The study of ensemble optimization can be seen as building an integrated model of optimization (15). That -means determine the optimal weights of two sub-classifiers. The multiple regression (MR) is used to create a binary regression equation for the outputs of the two sub-classifiers. The ordinary least squares (OLS) principle is used to determine the optimal estimate weights w_1 and w_2.

The training samples are used to train the integrated system, and the two sub-classifiers are input separately to obtain different recognition results. Suppose the i^{th} training sample, the recognize result of the sub classifier I is a_i, the recognize results of sub-classifier II is b_i.The output of the integrated system is y_i, and the real value of the samples is y_i^o $(i = 1,2,\cdots,n)$. By actual value $Y^o$of training sample, we define the binary regression equation as:

$$Y = \gamma + \alpha A + \beta B + \varepsilon,\quad \varepsilon \sim N(0, \sigma^2) \tag{6}$$

Using the method of maximum likelihood estimation to estimate the unknown parameters α, β and γ, which needs the following equations to obtain the minimum residual error value Q.

$$Q = \sum_{i=1}^{n} (y_i^o - y_i)^2 \tag{7}$$

It means:

$$min\ Q = \sum_{i=1}^{n} (y_i^0 - \gamma - \alpha a_i - \beta b_i)^2 \tag{8}$$

If Q is the minimum, then α, β, γ need to satisfy the following equation set:

$$\begin{cases} \frac{\partial Q}{\partial \alpha} = \sum_{i=1}^{n} (y_i^0 - \alpha a_i - \beta b_i - \gamma)\, a_i = 0 \\ \frac{\partial Q}{\partial \beta} = \sum_{i=1}^{n} (y_i^0 - \alpha a_i - \beta b_i - \gamma)\, b_i = 0 \\ \frac{\partial Q}{\partial \gamma} = \sum_{i=1}^{n} (y_i^0 - \alpha a_i - \beta b_i - \gamma) = 0 \end{cases} \tag{9}$$

Set:

$$X = \begin{bmatrix} 1 & a_1 & b_1 \\ 1 & a_2 & b_2 \\ \vdots & \vdots & \vdots \\ 1 & a_n & b_n \end{bmatrix}_{n\times 3} \qquad \eta = \begin{bmatrix} \gamma \\ \alpha \\ \beta \end{bmatrix}_{3\times 1} \tag{10}$$

The equation (19) will be translated into its matrix form:

$$X'X\eta = X'Y^o \tag{11}$$

Obtain the least squares estimate of coefficient matrix η:

$$\eta = (X'X)^{-1}X'Y^o \tag{12}$$

By the coefficient matrix η, the regression coefficient values are α, β, and weights w_1, w_2 can be further obtained:

$$w_1 = \frac{\alpha}{\alpha+\beta},\quad w_1 = \frac{\beta}{\alpha+\beta} \tag{13}$$

When w_1, w_2 values for formula (13), it makes the output value Y is closer to the actual output value Y^o, so make residual error Q get the minimum.

$$min\ Q = \sum_{i=1}^{n} (y_i^o - y_i)^2 = \sum_{i=1}^{n} (a_i w_1 + b_i w_2 - y_i^o) \tag{14}$$

Model (14) achieves the optimal ensemble learning in this moment, further establish ensemble optimization algorithm.

V. Genetic Neural Network Ensemble Learning Algorithm

When facing complex pattern classification problems, genetic neural networks are used to design sub-classifiers in order to optimize the neural network structure, improve the success rate of neural network training, and accelerate the convergence of neural networks, thus improving the operation efficiency and classification accuracy. The key point of the sub-classifier optimized by genetic algorithm is improving the operating efficiency, while the recognition accuracy is still difficult to achieve satisfactory results. To get ideal classification result, each sub-classifier will be integrated by ensemble learning. For recognize results of sub-classifier, multiple regression model is established. It performs optimization by OLS algorithm and then determines the weights of each sub-classifier.

From the above, an ensemble optimized classification algorithm by genetic neural network based on multiple regression is established.

A. Parameter setting

Data preprocessing

Data preprocessing removes the incommensurability by different data index distribution and numerical differences in itself, which ensure the quality of the use of data from the source. Using standardization transformation, the data is processed based on its distribution of N(0,1). The formula for standardization transformation is as follows:

$$x'_{ij} = \left(x_{ij} - \overline{x_j}\right) \Big/ \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_{ij} - \overline{x_j}\right)^2} \tag{15}$$

Parameters of RBF neural network

Set the center of the RBF network, choose the center of the basis function empirically, as long as the distribution of training sample can represent the problem, s centers of uniform distribution can be selected according to the experience, its spacing distance is d, choose the width of the Gaussian basis function is:

$$\sigma = d/\sqrt{2s} \tag{16}$$

Basis functions is selected by K-means clustering method, and regard the cluster centers as each center of functions.

Parameter of genetic algorithm

Set the population size is Ps, crossover probability is Pc, mutation probability is Pm, target error function is $Emin$, the largest number of iterations is $Gmax$.

B. Basic steps of new algorithm

All of above, the genetic neural network ensemble algorithm based on multiple regression is established, and its basic steps are as follow.

Step 1 To standardize the original data by formula (15), and the standardized data will be seen as the new samples;

Step 2 Set the RBF neural network, which is according to the GA-RBF neural network model established in 3.2.3 for training;

Step 3 Whether the GA-RBF neural network model of $Emin$ or $Gmax$ reach to set data, then terminate the algorithm and get the output A of the sub-classifier I, turn to Step 6.

Step 4 Set the Elman neural network, according to the GA-Elman neural network model established in 3.3.3 for training;

Step5 Whether the GA-Elman neural network model is $Emin$ or $Gmax$ reach to set data, then terminate the algorithm and get the output$B$of the sub-classifier II, turn to step 6.

Step 6 Set up the bivariate regression model (6) by the output of the two sub-classifiers;

Step 7 Transform equation (6) into equations set (9) by maximum likelihood estimation;

Step 8 Solve the set of equations (9) using OLS theory to obtain formula (12), which represents the regression coefficient model (6);

Step 9 Calculate the weights of the two sub-classifiers using formula (13);

Step 10 To get the optimal solution of integrated model (5) and terminate the algorithm.

VI. Cases Analysis

To verify the reliability of the new algorithm, three experiments are arranged. Two experiments are tested by UCI standard data, the ionosphere data subset of radar [42] and the waveform generator set [43] is selected. Another experiment is tested by agricultural production data [44]. In the process of experiments, each group of experiments is conducted in two steps: firstly, test the overall performance of each algorithm; secondly, select 5 simulation samples and compare the simulation results with the actual values.

A. Marked every classification algorithm

In order to evaluate the performance of integrated classification algorithm, we will compare with the results of each algorithm.

Traditional RBF neural network is denoted by RBF;

Traditional Elman neural network is denoted by Elman;

The optimized GA-RBF neural network model is denoted by sub-classifier I;

The optimized GA-Elman neural network model is denoted by sub-classifier II;

The optimized classification algorithm by genetic neural network ensemble based on multiple regression is denoted by assembling system.

When $w_1 = 1, w_2 = 0$, sub-classifier I is used to identify the dataset; When $w_1 = 0, w_2 = 1$, sub-classifier II is used to identify the dataset.

B. Preparation of experiment

The algorithms are evaluated in terms of four aspects: training success rate, training steps, total squared error and recognition accuracy. The training success rate is the number of times each algorithm is successfully trained in the training case, here 50 trainings for each algorithm. Convergence step, i.e., the optimal model is obtained after 50 training sessions for each algorithm. The sum of squared errors is used to measure the closeness between the actual values and the estimated values. It represents the sum of the squared differences between the predicted values and the actual values. Accuracy is the recognize accuracy of the algorithm. Under the circumstances of same simulation accuracy, the smaller the error sum of squares shows that the precision of the algorithm is higher;

Genetic neural network, the maximum number of iterations of genetic algorithm is 300 generations, the target of error is MSE 1e-5, the population size S is 30, the crossover rate P_c is 0.9, mutation rate P_m is 0.01.

Evolutionary strategy, Elite reserved strategy (directly retain four most adaptive individuals into next generation); Linear transformation ratio strategy (original fitness makes linear transformation ratio firstly, and then to select operation).

GA-RBF neural network, the maximum number of iterations of the LMS method is 2000 and the learning rate is 0.1. GA-Elman neural network, the maximum number of iterations of LM algorithm is 2000 and the learning rate is 0.1.

The experiments run on Intel Core2 Duo CPU E7300 2.66 GHz, RAM 1.99GB.

C. Experiment

We conduct experiments on two sets of data from the UCI standard databases and one set of experimental data from agricultural production. The experiments are arranged as following.

Experiment I

The algorithm is tested using a subset of radar ionospheric data, which has 351 samples and 34 features. Each sample is used to predict the quality of the radar, i.e., there are 34 inputs and 1 output of the data. 200 samples are selected as training samples and the remaining 151 samples are used as simulation samples.

We use the traditional RBF neural network, the traditional Elman neural network, GA-RBF neural network (sub-classifier I) and GA-Elman neural network (sub-classifier II), genetic neural network ensemble algorithm to test respectively. The simulation results of ionosphere radar data set list in table 1.

In order to better illustrate the classification ability of the new algorithm, we list the simulate results of last simulating samples. Compared with the simulation results and the actual value, the test results are listed in table 2.

Experiment II

The waveform generator database dataset consists of 5000 samples, where each sample has 40 features to describe the data, including 40 inputs and 1 output. We selected 500 samples as training samples and used 200 samples for simulation testing. After testing with 5 different algorithms, the simulation results of the waveform generator dataset are shown in table 3.

TABLE I
COMPARISON OF EVERY CLASSIFIER'S PERFORMANCE FOR IONOSPHERE DATA SUBSET OF RADAR

Network model	Success rate, %	Weights	Convergence steps	Sum of square error	Accuracy, %
RBF	68.00	—	85	26.3162	68.21
Elman	82.00	—	934	15.0965	77.48
Sub-classifier I	100	0.3633	300+35	18.3284	81.46
Sub-classifier II	100	0.6367	300+167	7.3705	93.38
Assembling system	—	—	—	3.1526	98.68

TABLE 2
COMPARISON OF EVERY CLASSIFIER'S ACCURACY FOR TEST I

Network model	Weights	1st	2nd	3rd	4th	5th	SSE
actual value	—	0	1	0	1	0	—
RBF	—	-0.4163	0.6917	0.2231	0.4179	-0.2553	0.7221
Elman	—	0.4255	0.7966	-0.2065	1.3505	0.2176	0.4353
Sub-classifier I	0.3633	-0.3219	0.7818	-0.2116	0.5132	-0.1279	0.4493
Sub-classifier II	0.6367	0.2137	0.8972	-0.1037	1.2881	0.1384	0.1691
Assembling system	—	0.0191	0.8553	-0.1429	1.0066	0.0417	0.0435

TABLE 3
COMPARISON OF EVERY CLASSIFIER'S PERFORMANCE FOR WAVEFORM GENERATOR SET

Network model	Success rate, %	Weights	Converge-nce steps	Sum of square error	Accuracy%
RBF	64.00	—	85	40.5057	69.50
Elman	82.00	—	1163	15.6860	78.00
Sub-classifier I	100	0.3809	300+55	30.4631	83.50
Sub-classifier II	100	0.6191	300+213	10.0362	92.50
Assembling system	—	—	—	4.4693	99.00

TABLE 4
COMPARISON OF EVERY CLASSIFIER'S PERFORMANCE FOR WAVEFORM GENERATOR SET

Network model	Success rate, %	Weights	Convergence steps	Sum of square error	Accuracy %
RBF	66.00	—	92	3.8717	70.80
Elman	84.00	—	1077	2.0593	78.70
Sub-classifier I	100	0.3393	300+57	1.8326	83.30
Sub-classifier II	100	0.6607	300+223	1.2197	92.90
Assembling system	—	—	—	0.2962	98.80

Experiment III

This experiment uses a set of wheat aphid (a wheat pest) prediction data from actual agricultural production. The wheat midge will influence the growth of wheat, and even lead to reduce the production. To predict the occurrence situation of wheat midge in advance, it can make targeted prevention, and has a great significance for agricultural production. The dataset consists of 60 years of agricultural production data, which is considered as 60 samples. Each sample includes 14 agricultural-related meteorological conditions and the severity of pest occurrence. Since the growth of wheat aphids is closely related to meteorological conditions, it is possible to utilize these conditions to predict the severity of aphid infestation, which is categorized into five levels. In this experiment, the first 40 samples as the training sample, and the rest of 20 samples as testing sample. Respectively, the simulation results are listed in table 4.

D. Results analysis

According to the experimental results in tables 1 to 4, it can be seen from three experimental results that the four performance indicators of the new optimized genetic neural network are significantly improved compared to traditional neural networks. Among the training success rate and convergence steps are improved, which means that operating efficiency has got improvement. However, classification accuracy and sum of squares error still have some space to improve, namely the classification precision is not ideal.

In terms of operational efficiency, the GA-RBF algorithm performs slightly better than the GA-Elman model among the two sub-classifiers. However, in terms of classification accuracy, the GA-RBF algorithm shows poorer performance. Therefore, in the ensemble learning process, the weight assigned to the GA-Elman sub-classifier is slightly higher than that of the GA-RBF sub-classifier. The classification accuracy of the new ensemble algorithm is higher than that of any sub-classifier, and the sum of squares error is significantly lower than that of any sub-classifier. These results show that the classification precision has been greatly improved, which meets the ideal requirements. It is important to note that the classification of new ensemble algorithm cannot to reach 100%, which may be caused by some outlier samples in complex data.

E. Discussion

To sum up, aiming at the classification problem of complex data, neural network is used to design classifier. By genetic and ensemble two consecutive optimizing, a reliable ensemble classification algorithm by genetic neural network based on multiple regression is proposed. Genetic optimizing improves the operating efficiency, and ensemble optimizing improves the classification precision. The results of three group experiments demonstrate that the effectiveness of new algorithm is best.

In this paper, we perform ensemble learning by two sub-classifiers whose weights can be determined by the optimal solution of the binary regression model, and obtain the optimal ensemble classification system. If more complex problems are encountered, multiple sub-classifiers may need to be integrated. How to determine the number of sub-classifiers, design sub-classifiers, and determine the weight of each sub-classifier is a challenge for ensemble classification system.

VII. CONCLUSION

New genetic neural network as classifiers can be improved on operating efficiency. Using genetic algorithm to optimize the structures and connection weights of neural network can overcome some inherent defects, which includes accelerating the convergence, improving the success rate, reducing the time of designing neural network structure. The results show that the optimizing way of hybrid encoding and simultaneous evolving to improve the performance of neural network is obvious. However, the two sub-classifiers fail to achieve satisfactory results in terms of classification accuracy and need to be further optimized.

To obtain more desirable classification accuracy, an ensemble optimization of two sub-classifiers, i.e., an ensemble classification system, is constructed in this study. The key problem of ensemble learning is how to determine the weight of each sub-classifier effectively. By establishing the multiple regression model, OLS principle is used to solve multiple regression equation and to get regression coefficient, and then the weight of each sub-classifier is determined. For both sub-classifiers, the weights of GA-Elman neural network are larger than GA-RBF, indicating the high classification accuracy of the sub-classifiers. The experimental results show that the classification accuracy of the sub-classifiers is significantly higher than that of any individual classifier.

REFERENCES

[1] Jain LC, Seera M, Lim CP, et al. A review of online learning in supervised neural networks. Neural Computing & Applications, 2014, 25(3-4): 491-509.

[2] Kim TH. Pattern Recognition Using Artificial Neural Network: A Review. Communications in Computer and Information Science, 2010, 76: 138-148.

[3] Mccllochw S, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 1943, 10(5): 115-133.

[4] Hansen L K, Salamon P. Neural network ensembles. IEEE Trans Pattern Analysis and Machine Intelligence, 1990, 12(10): 993-1001.

[5] Kourentzes N, Barrow DK, Crone SF. Neural network ensemble operators for time series forecasting. Expert Systems with Applications, 2014, 41(9): 4235-4244.

[6] Pai DR, McFall KS, Subramanian GH. Software effort estimation using a neural network ensemble. Journal of Computer Information Systems, 2013, 53(4): 49-58.

[7] Lee H, Kim E, Pedrycz W. A new selective neural network ensemble with negative correlation. Applied Intelligence. 2012, 37(4): 488-498.

[8] Jia W, Sun M, Lian J, et al. Feature dimensionality reduction: A review. Complex & Intelligent Systems, 2022, 8(3): 2663-2693.

[9] Tian J, Li MQ, Chen FZ, et al. Coevolutionary learning of neural network ensemble for complex classification tasks. Pattern Recognition, 2012, 45(4): 1373-138.

[10] Ding SF, Jia WK, Su CY, et al. Neural Network Research Progress and Applications in Forecast. Lecture Notes in Computer Science, 2008, 5264: 783–793.

[11] Moody J, Dkaren C. Fast learning in networks locally-tuned processing units. Neural Computation, 1989, l: 281-294.

[12] Liu GH, Xiao XH, Mei CL, et al. A review of learning algorithm for radius basis function neural network. The 24th Chinese Control and Decision Conference, 2012, 1112-1117.

[13] Elman J L. Finding structure in time. Cognitive Science, 1990, 14(2): 179-211.

[14] Sheela KG, Deepa SN. Review on Methods to Fix Number of Hidden Neurons in Neural Networks. Mathematical Problems in Engineering, 2013, 425740: 1-11.

[15] Etemad SA, Arya A. Classification and translation of style and affect in human motion using RBF neural networks. Neurocomputing, 2014, 129(SI): 585-595.

[16] Martinez LFJ, Torres AJA, Martinez PS, et al. Multilevel RBF to resolve classification problems with large training sets: new pseudo-exact procedure. Soft Computing, 2014, 18(11): 2245-2252.

[17] Nakagawa M, Kondo T, Kudo T, et al. Three-dimensional medical image recognition of cancer of the liver by a revised radial basis function (RBF) neural network algorithm. Artificial Life and Robotics, 2009, 14(2): 118-122.

[18] Qian J, Long H, Long Y & Zhao C. Improved NSGA-III algorithm and BP fuel-cost prediction network for many-objective optimal power flow problems. IAENG International Journal of Applied Mathematics, vol. 51, no. 2, pp. 307-320, 2021.

[19] Xu X, Wang L, Sun Z & Wu L. Aseptic packaging of wheat bun and the prediction of shelf life using a BP neural network. IAENG International Journal of Computer Science, vol. 48, no.4, pp1129-1134, 2021

[20] Jia WK, Zhao DA, Shen T, et al. Study on optimized Elman neural network classification algorithm based on PLS and CA. Computational Intelligence and Neuroscience, 2014, 724317: 1-13.

[21] Lu JJ, Chen H. Researching development on BP neural networks. Control Engineering of China, 2006, 13(5): 449-451. (in Chinese)

[22] Zhao M, Zhao Y, Ma C & Wu X. Robust Optimization of Mixed-Load School Bus Route Based on Multi-Objective Genetic Algorithm. Engineering Letters, vol. 30, no. 4, pp. 1543-1550. 2022.

[23] Ding SF, Li H, Su CY, et al. Evolutionary artificial neural networks: a review. Artificial Intelligence Review, 2013, 39(3):251-260.

[24] Yao X. Evolving artificial neural networks. Proceedings of the IEEE, 1999, 87(9): 1423-1447.

[25] Piotrowski AP. Differential evolution algorithms applied to neural network training suffer from stagnation. Applied Soft Computing, 2014, 21: 382-406.

[26] Harpham C, Dawson C W, Brown M R. A review of genetic algorithms applied to training radial basis function networks. Neural Computing and Applications, 2004, 13(3): 193-201.

[27] Ding S F, Xu L, Su C Y, et al. An optimizing method of RBF neural network based on genetic algorithm. Neural Computing and Applications, 2012, 21(2): 333-336.

[28] Chang WY. An RBF neural network combined with OLS algorithm and genetic algorithm for short-term wind power forecasting. Journal of Applied Mathematics, 2013, 971389:

[29] Ding SF, Zhang YN, Chen JR, et al. Research on using genetic algorithms to optimize Elman neural networks. Neural Computing and Applications. 2013, 23(2): 293-297.

[30] Chen XY, Song R, Shen C, et al. Application of a genetic algorithm Elman network in temperature drift modeling for a fiber-optic gyroscope. Applied Optics, 2014, 53(26): 6043-6050.

[31] Rohitash C, Zhang MJ. Cooperative coevolution of Elman recurrent neural networks for chaotic time series prediction. Neurocomputing, 2012, 86(1): 116-123.

[32] Zhou ZH, Chen SF. Neural network ensemble. Chinese Journal of Computers, 2002, 25(1): 1-8. (in Chinese)

[33] Kourentzes N, Barrow DK, Crone SF. Neural network ensemble operators for time series forecasting. Expert Systems with Applications, 2014, 41(9): 4235-4244.

[34] Yang J, Zeng XQ, Zhong, SM, et al. Effective neural network ensemble approach for improving generalization performance. IEEE Transactions on Neural Networks and Learning Systems, 2013, 24(6): 878-887.

[35] Lee H, Kim E, Pedrycz W. A new selective neural network ensemble with negative correlation. Applied Intelligence, 2012, 37(4): 488-498.

[36] Chen WC, Tseng LY, Wu CS. A unified evolutionary training scheme for single and ensemble of feedforward neural network. Neurocomputing, 2014, 143(SI): 347-361.

[37] Ao SI, Palade V. Ensemble of Elman neural networks and support vector machines for reverse engineering of gene regulatory networks. Applied Soft Computing, 2011, 11(2): 1718-1726.

[38] Zhang CX, Zhang JS. A survey of selective ensemble learning algorithms. Chinese Journal of Computers, 2002, 34(8): 1399-1410. (in Chinese)

[39] Reynolds RG, Ali M, Yao X, et al. Evolving artificial neural network ensembles. IEEE Computational Intelligence Magazine, 2008, 3(1): 31-42.

[40] Zhou ZH, Wu JX, Tang W. Ensemble neural networks: Many could be better than all. Artificial Intelligence, 2002, 137(1-2): 239-263.

[41] Jia WK, Zhao DA, Tang YY, et al. An optimized classification algorithm by neural network ensemble based on PLS and OLS. Mathematical Problems in Engineering, 2014, ID: 395263

[42] Sigillito V, Wing S, Hutton L, et al. Ionosphere. UCI Machine Learning Repository. https://doi.org/10.24432/C5W01B. 1989.

[43] Breiman L, Stone CJ. Waveform Database Generator (Version 2). UCI Machine Learning Repository. https://doi.org/10.24432/C56014. 1988.

[44] YM Zhang. The application of artificial neural network in the forecasting of wheat midge. Northwest A&F University. 2003.

Date of Modification: January 23, 2024.
Brief description of the changes: Correct the name of the first author from Xishan Dong to Xishang Dong.