A comparative study on artificial neural networks for environmental quality assessment

Yijun Liu*, Sheng He, Yao Wang, Xiumei Wang

Key laboratory of cloud computing & intelligent information processing of Changzhou City, Jiangsu University of Technology, 213001, China

Received 6 March 2014, www.tsi.lv

Abstract

The aim of this study is to use neural network tools as an environmental decision support in assessing environmental quality. A three-layer feedforward neural network using three learning approaches of BP, LM and GA-BP has been applied in non-linear modelling for the problem of environmental quality assessment. The case study shows that the well designed and trained neural networks are effective and form a useful tool for the prediction of environmental quality. Furthermore, the LM network has the fastest convergence speed and the GA-BP network outperforms the other two networks in both predictive and final classification accuracies of environmental quality.

Keywords: neural network model, hybrid ga-bp algorithm, environmental quality assessment

1 Introduction

Environment is the basic premise of human survival and development. Environmental quality is defined as a set of properties and characteristics of the environment, either generalized or local, as they impinge on human beings and other organisms [1]. It is a measure of the environmental condition relative to the requirements of human or other species. Today environmental protection has been a basic national policy in China. It is a difficult task for the government to solve serious environmental problems and make sustainable development strategies. As an essential reference for environmental protection policies and an effective tool for supervision and management, environmental quality assessment, a quantitative description of environmental quality is one of the most challenging areas in environment studies. It is significantly important to make environmental quality assessment more accurate and scientific.

Environment quality assessment can be considered as a pattern recognition task since it deals with a set of input environmental variables which contain information about the environment of a region, mapping to real values indicating environmental quality or a set of discrete and mutually exclusive classes indicating quality degree such as light pollution, heavy pollution, etc. Environmental quality ratings are typically costly to be obtained, since they require investing large amount of time and human resources to perform deep analysis of the environmental status based on various aspects of soil, atmosphere, water, etc. Previous studies focus on mathematical models created by conventional statistical methods [2]. However, in traditional statistical methods researchers are usually required to impose particular structures to different models, such as the linearity in the multiple regression analysis, and to construct the model by estimating parameters to fit the data or observation [3]. Closely related to social, economic, management and other fields, environmental quality assessment is a complex multiple-objective, multiple-level and multiple-factor engineering problem, and hence does not adhere to common function forms. In recent years, an artificial intelligence technique, namely artificial neural network, abbr. ANN, has attracted many attentions and has been widely used in many real-world applications such as medical diagnosis [4, 5], image classification [6], signal change detection [7], handwriting recognition [8], etc. By using neural networks the structure of the models can be obtained from data automatically. To capture the complexity and the process dynamics of complicated environmental quality system, the neural network method is used to make a comprehensive analysis for environmental quality.

The rest of this paper is organized as follows. In Section 2, the paper defines the problem of environmental quality assessment mathematically and suggests the novel method of neural networks for this problem. In Section 3, three neural network training approaches of backpropagation (abbr. BP), Levenberg-Marquardt (abbr. LM) and backpropagation optimized by the genetic algorithm (abbr. GA-BP) used in this work are introduced. In section 4, a case study is given to show the effectiveness of neural networks in assessing environmental quality. Firstly, the neural network topology is designed and presented according to actual situation. Then the modelling for environmental quality with using the designed network architecture is described. And lastly, the test results of three kinds of neural

* ***Corresponding author*** e-mail: yijunliu@vip.sina.com

networks are analysed and compared. In section 5, conclusions are drawn and the issue for future works is indicated.

2 Problem definitions

Environmental quality assessment can be defined as a pattern recognition task of regression. Let $P=\{p_1, p_2, \ldots, p_t\}$ be a set of monitoring points, $X=\{x_1, x_2, \ldots, x_n\}$ be a set of attributes. For attribute value vector $(x_{1i}, x_{2i}, \ldots, x_{ni})$ of each monitoring point p_i, there is a corresponding assessment score y_i of p_i, $1 \le i \le t$. Assume there is a mapping f as such $y_i=f(x_{1i}, x_{2i}, \ldots, x_{ni})$ from attribute values to assessment score. For a monitoring point to be assessed the input of the mapping f is its attribute value vector and the output is the assessment score.

Artificial neural networks can be used as an arbitrary function approximation mechanism that learns from observed data, and hence are applicable to environmental quality assessment. Although neural networks have been criticized for their poor interpretability, they have strong ability of non-linear mapping, high tolerance to noisy data and superior robustness if the learning algorithm and the cost function are appropriately selected for modelling.

The neural network method imitates the way by which the brain processes information [4]. Given an input vector $X=(x_1, x_2, \ldots, x_n)$, the network produces an output vector $Y=(y_1, y_2, \ldots, y_m)$, where n indicates the number of inputs and m the number of output units. Typically a feedforward neural network is organized into several layers of nodes. The first layer is the input layer and the last layer is the output. In the input layer every node corresponds to an attribute variable and hence the number of nodes equals the number of variables. The input and output layers are usually separated by one or more hidden layers. The nodes in adjacent layers are fully connected. There is a weight associated with each connection. The weight from unit i to unit j is denoted as w_{ij}. For neural network training, learning rules are used to update the weight and to minimize the error function. The learning process repeats until termination condition is met. A degree of nonlinearity is introduced to the model by the activation function to prevent the output from reaching very large values that paralyze neural network models and inhibit training.

A well-trained neural network is capable of exploiting the underlying non-linear relationships that determine the environmental rating of a region. In this paper, we propose a three-layer network based on three learning algorithms of BP, LM and GA-BP for the problem of environmental quality assessment.

3 Neural network approaches

Allowing the modelling of non-linear relationships, neural networks are useful tools for the analysis of large data sets of non-congeneric compounds with unknown or varying modes of action. The standard backpropagation algorithm is perhaps the most widely used algorithm for supervised training of multi-layered feedforward neural networks. However, the BP algorithm has two significant disadvantages of slow convergence speed and easiness of falling into the local minimum point [9]. Therefore improved training algorithms involving the LM algorithm and the GA-BP algorithm are proposed.

3.1 BACKPROPAGATION ALGORITHM

The BP algorithm is a learning method which minimizes the error of the neural network output compared to the required output. In this algorithm, the performance index $F(w)$ to be minimized is defined as the sum of squared errors between the target outputs and the network's simulated outputs, namely:

$$F(w)=e^T e\,, \tag{1}$$

where $w = [w_1, w_2, \ldots, w_n]$ consists of all weights of the network, e is the error vector comprising the errors for all training samples. The steps involved in training a neural network using the BP algorithm are as follows [8]:

step 1. Initialize each w_i to some small random value.

step 2. Do the step 3 until the termination condition is met.

step 3. For each training sample $<(x_1, x_2, \ldots, x_n), t>$, where t is the target output, do

1) Input the instance $(x_1, x_2, \ldots, x_n)$ to the network and compute the network outputs o_k.

2) For each output unit k, $\delta_k= o_k(1- o_k)(t_k- o_k)$.

3) For each hidden unit h, $\delta_h= o_h(1- o_h)\sum_k w_{h,k}\delta_k$

4) For each network weight, do $w_{i,j}=w_{i,j}+\Delta w_{i,j}$, where $\Delta w_{i,j} = \eta\delta_j x_{i,j}$, η is the learning rate defined by users. The algorithm is assumed to have converged when the norm of the gradient is less than some predetermined value, or when the error has been reduced to some error goal.

3.2 LEVENBERG-MARQUARDT ALGORITHM

As a kind of Gradient-based training algorithms, the BP algorithm is not efficient because the gradient vanishes at the solution. The neural networks are allowed to learn more subtle features of a complicated mapping by using Hessian-based algorithms, whose training process converges quickly as the solution is approached due to the fact that the Hessian does not vanish at the solution [10]. To benefit from the advantages of Hessian based training, the Levenberg-Marquardt algorithm is used to train the neural network for environmental quality analysis.

When training with the LM algorithm, Δw, the increment of weights can be obtained as follows:

$$\Delta w = \left[J^T(w)J(w) + \mu I\right]^{-1} J^T(w)e\,, \tag{2}$$

where $J(w)$ is the Jacobian matrix, $J^T(w)J(w)$ is referred as the Hessian matrix, I is the identity matrix, and μ is the learning rate which is to be updated depending on the outcome. In particular, μ is multiplied by the decay rate β ($0<\beta<1$) whenever $F(w)$ in Equation (1) decreases, whereas μ is divided by β whenever $F(w)$ increases in a new step.

The standard LM training process is illustrated in the following steps.

step 1. Initialize the weights and parameter μ.

step 2. Compute the sum of the squared errors over all inputs $F(w)$.

step 3. Until the termination condition is met, do

1) Solve Equation (2) to obtain Δw, the increment of weights.

2) Use $w+\Delta w$ as the trial w, and judge

if trial $F(w+\Delta w)<F(w)$ in step 2 then $w=w+\Delta w$, $\mu=\mu\beta$

Go back to step 2

else

$\mu=\mu/\beta$

Go back to step 3.1.

In the LM algorithm the parameter μ is adjusted automatically at each iteration in order to secure convergence. This algorithm becomes Gauss-Newton method for μ=0 and becomes steepest decent or the error backpropagation algorithm when μ is very large.

3.3 HYBRID GA-BP ALGORITHM

From mathematical point of view, the BP learning is a non-linear optimization problem and there exists the local minimum point inevitably. The genetic algorithm, abbr. GA, is a highly parallel, stochastic and adaptive optimization technique based on biological genetic evolutionary mechanisms [11]. GA has good global search performance since its search has always been throughout the solution space, while the BP learning algorithm is more effective in local search. Therefore, the hybrid training algorithm combining BP with GA can achieve the goal of network optimization successfully.

GA can be used to optimize the network topology and weights [12]. In this work it is used to optimize the network weights before the BP learning. The optimization problem is described as follows: $min(e)=f(w_1, w_2, \ldots, w_n)$, where $w_1, w_2, \ldots, w_n$ are all network weights satisfying $-1<w_i<1$, $1\le i\le n$, and n is the number of total weights.

GA simulates the process of natural evolution, performing operations similar to natural selection, crossover and mutation to obtain the final optimization result after repeated iterations. The weight optimization process of the neural network using GA is described in the following steps.

step 1. Start with a randomly generated population comprising N chromosomes, and each chromosome encodes a set of network weights. Set the maximum generation number *gen*, the crossover probality P_c and the mutation probability P_m.

step 2. If the maximum generation number *gen* has been reached, go to step 6.

step 3. Calculate the fitness of each individual chromosome by the fitness formula: $f_i = 1/E_i$, where f_i presents the fitness of chromosome i and E_i presents the MSE, i.e. mean square error of the neural network corresponding to the chromosome i.

step 4. Repeat the following steps until the new population size reaches N.

1) Use the roulette wheel strategy to select a pair of chromosomes from the current population for single-point crossover and mutation with probability of P_c and P_m respectively. For the mutation operator, add a random value between (-1, 1) to each mutation point.

2) Put the new pair of chromosomes in the new population.

step 5. Take the new N chromosomes as the new population. Go to step 2.

step 6. Decode the chromosome with the highest fitness to obtain initial weights of the neural network.

After global optimization with GA, the BP algorithm is used for the local search until the termination criterion is satisfied.

4 A case study

4.1 DESIGN OF NEURAL NETWORK TOPOLOGY

As mentioned in section 2, environmental quality assessment is a pattern recognition problem. The objective of the neural network is to give an accurate comprehensive quality assessment according to environmental input vector. The critical step in building a robust neural network is to create an architecture, which should be as simple as possible and has a fast capacity for learning the data set. The network topology are typically best determined empirically.

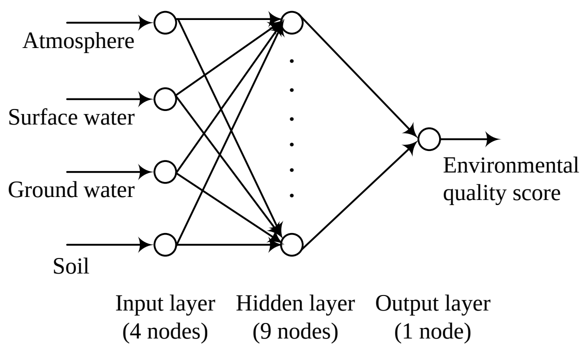

FIGURE 1 Neural network topology for environmental quality assessment

In this study, the selected environmental variables are atmosphere (x_1), surface water (x_2), ground water (x_3) and soil (x_4). Therefore the input layer consists of four nodes representing components of the four-dimensional input vector $X=(x_1, x_2, x_3, x_4)$. The number of nodes in the hidden layer usually is not less than the number of input nodes. Based on Kolmogorov theory, $2N+1$ hidden nodes should be used for one hidden layer, where N is the number of input nodes. Considering the three-layered topology with one hidden layer, we will have nine hidden nodes because of four input nodes. The output layer contains a single node giving the score of environmental quality. The architecture of the 4-9-1 feedforward neural network used in this analysis is shown in Figure 1.

4.2 TRAINING OF NEURAL NETWORK

Three learning algorithms of BP, LM and GA-BP described in section 3 together with the neural network architecture presented Figure 1 are used to create, train and test the neural network for environmental quality assessment. The neural networks are implemented using MATLAB 7 (MathWorks, USA) software with its neural network toolbox.

Table 1 tabulates the samples which are from 20 different monitoring points in a certain region of China [13, 14]. Depending on the actual condition of China, environmental quality is divided into four classes of A, B, C and D [13]. An "A" represents high environmental quality without pollution, "B" represents good quality with light pollution, "C" represents ordinary quality with moderate pollution, and "D" signifies a severe pollution. There is a score scope per output pattern, that is, A-[0, 1], B-[1, 2.5], C-[2.5, 5] and D-[5, 10]. To help speed up the learning process, the input values for each attribute have been scaled so as to always fall within a specified range [0.1, 1] by the following Equation: $F_j = \frac{x_j - x_{j\min}}{x_{j\max} - x_{j\min}} \times 0.9 + 0.1$, where F_j is the normalization of the attribute value x_j, $x_{j\min}$ is the minimum value and $x_{j\max}$ is the maximum value of the j^{th} attribute.

In this analysis, the first ten samples in Table 1 have been taken for the training set and the rest for the test set. The transfer function between the input layer and the hidden layer is the tangent Sigmoid function *tansig*, and the one between the hidden layer and the output layer is the linear function *purelin*. The performance curves of the BP, LM and GA-BP networks are shown as Figures 2-5. From Figure 2 we can see that the BP network converges to the preset precision after 15142 epochs, while from Figure 3 the LM network only needs 3 epochs, greatly faster than the BP network. For the GA-BP network we set the initial population size N to 30, the maximum generation number *gen* to 800, the crossing probability P_c to 0.95 and the mutation probability P_m to 0.09. Figure 4 shows the curves of MSE, i.e. Mean Square Error and the fitness in the first process of weight optimization by GA.

TABLE 1 Normalized data set in environmental quality assessment

No.	Atmosphere	Surface water	Ground water	Soil	Assessment score	Environmental quality rating
1	0.1010	0.1000	0.2171	0.1130	0.9	A
2	0.1000	0.1024	0.2220	0.1130	0.9	A
3	0.2173	0.2156	0.2951	0.1000	1.8	B
4	0.2006	0.1094	0.2366	0.1097	1.6	B
5	0.2168	0.1496	0.1000	0.1173	1.7	B
6	0.2284	0.2864	0.9000	0.1336	2.2	B
7	0.6198	0.1755	0.4220	0.1227	4.4	C
8	0.3554	0.2864	0.5732	0.1195	2.9	C
9	0.6082	0.1330	0.5878	0.4735	5.2	D
10	0.5329	0.5224	0.3439	0.9000	5.9	D
11	0.4322	0.2628	0.2463	0.1217	3.3	C
12	0.2067	0.1590	0.2951	0.1152	1.7	B
13	0.9000	0.5224	0.3537	0.1314	6.6	D
14	0.5880	0.3336	0.2902	0.1758	4.5	C
15	0.2775	0.1496	0.2951	0.1173	2.1	B
16	0.5243	0.1850	0.7049	0.1693	4.1	C
17	0.2896	0.1519	0.3098	0.1227	2.2	B
18	0.5637	0.9000	0.7146	0.3349	5.4	D
19	0.3073	0.1684	0.8951	0.1520	2.7	C
20	0.5030	0.7442	0.5878	0.1671	4.5	C

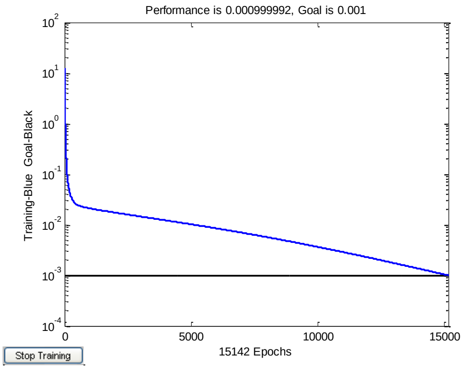

FIGURE 2 Gradient per epoch with the BP algorithm

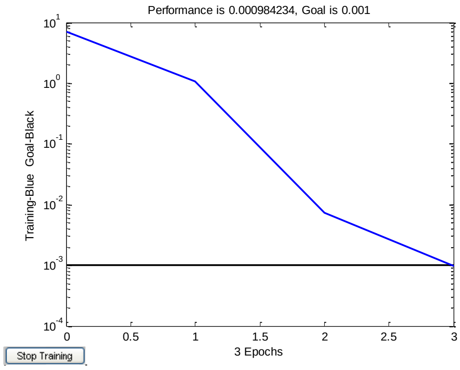

FIGURE 3 Gradient per epoch with the LM algorithm

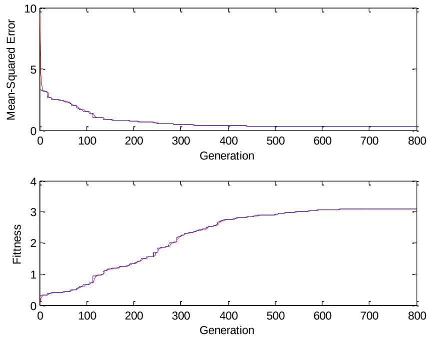

FIGURE 4 MSE and fitness curves of GA optimization

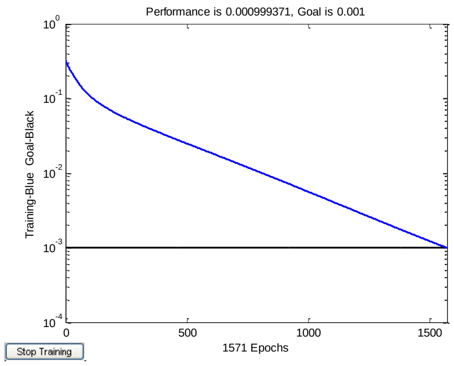

FIGURE 5 Gradient per epoch with the GA-BP algorithm

It can be seen that the fitness of chromosomes tends to stabilize after 500 generations. By this time the error of the GA-BP network reaches to 0.3. In the second process of network training the BP algorithm is used with the error goal of 0.001 and the learning rate of 0.01. The GA-BP network needs 1571 epochs to converge as shown in Figure 5 and it is much faster than the BP network but still slower than the LM network.

4.3 TEST RESULTS AND COMPARISONS

Table 2 tabulates the test results of three neural networks trained by BP, LM and GA-BP algorithms. All test cases are labelled with corresponding quality classes by predictive values. Table 3 tabulates the maximum, minimum and average relative errors and final classification accuracy obtained according to results in Table 2.

Average relative errors of environmental quality prediction of the BP, LM and GA-BP networks are attained 9.05%, 6.13% and 6.12% respectively as tabulated in Table 3. These three neural networks are all effective in assessing environmental quality since their predictive accuracies are all within 10%. The LM and GA-BP networks are both more accurate than the BP network in environmental quality prediction and the GA-BP network has the lowest average relative error.

Comparing quality ratings of three networks in Table 2 with actual quality ratings in Table 1, we can see that for both the LM and GA-BP networks the 18th sample with heavy pollution indicated by "D" is labelled with the wrong quality class of moderate pollution indicated by "C", while for the BP network the 14th, 18th and 19th samples are identified wrongly. The classification accuracies of the BP, LM and GA-BP models are 70%, 90% and 90% respectively. Therefore the LM and GA-BP networks are superior in final classification accuracy to the BP network.

Summarily, the LM and GA-BP networks are significantly better than the BP network with faster training speed and better generalization ability. The LM network has the fastest convergence speed, and the GA-BP network has the best performance in assessing environmental quality considering both predictive and final classification accuracies.

5 Conclusions and future works

Environmental quality assessment contributes to decision making in support of sustainable economic and social development, and it has attracted many research interests in the literature. Recent studies show that the neural network method has achieved better performance than the traditional statistical method. Due to the advantages that neural networks can learn a complex nonlinear relationship with limited prior knowledge and perform inferences for an unknown combination of input variables, this paper introduces the popular BP neural network and two other LM and GA-BP networks to environmental quality assessment in attempt to provide a model with good ability of generalization. Experimental results show that the assessment task is successfully accomplished by using neural networks.

TABLE 2 Test results for environmental quality

No.	BP algorithm			LM algorithm			GA-BP algorithm		
	Predict score	**Relative error**	**Rating**	**Predict score**	**Relative error**	**Rating**	**Predict score**	**Relative error**	**Rating**
11	3.7983	15.10%	C	3.0548	7.43%	C	3.2515	1.47%	C
12	1.6452	3.22%	B	1.7213	1.25%	B	1.7426	2.51%	B
13	7.0073	6.17%	D	6.0187	8.81%	D	5.5179	16.4%	D
14	5.1948	15.44%	D	4.3065	4.30%	C	4.2892	4.68%	C
15	2.0484	2.46%	B	2.1916	4.36%	B	2.2184	5.64%	B
16	3.5519	13.37%	C	4.4872	9.44%	C	4.0303	1.70%	C
17	2.1185	3.70%	B	2.2636	2.89%	B	2.3291	5.87%	B
18	4.9912	7.57%	C	4.9802	7.77%	C	4.7489	12.06%	C
19	2.3236	13.94%	B	2.9165	8.02%	C	2.7036	0.13%	C
20	4.9300	9.56%	C	4.8181	7.07%	C	4.0182	10.71%	C

TABLE 3 Result comparisons of three neural networks

	BP network	**LM network**	**GA-BP network**
maximum relative error	15.44%	9.44%	16.4%
minimum relative error	2.46%	1.25%	0.13%
average relative error	9.05%	6.13%	6.12%
classification accuracy	70%	90%	90%
epochs	15142	3	1571

However, a weakness of the study is that the created models are relatively hard to explain because neural networks are more concerned about the actual number of variables rather about their nature. Therefore one future direction of the research is to improve interpretability of the neural network models for the problem of environmental quality assessment.

Acknowledgements

This work is supported by the Key Laboratory of Cloud Computing & Intelligent Information Processing of Changzhou City under Grant No. CM20123004 and Applied Basic Research Program of Jiangsu University of Technology under Grant No. KYY10059. Furthermore, we are indebted to the support and encouragements received from the staff and colleagues of School of Computer Engineering.

References

[1] Environmental Terminology and Discovery Service (ETDS) 2014 *Environmental Quality* http://glossary.eea.europa.eu/EEAGlossary/E/environmental_quality-2014/16 May 2014

[2] Wang H 2002 Assessment and prediction of overall environmental quality of Zhuzhou city Hunan province China *Journal of Environmental Management* **66**(3) 329-40

[3] Zan H, Chen H 2004 Credit rating analysis with support vector machines and neural Networks: a market comparative Study *Decision Support Systems* **37**(4) 543-58

[4] Fidele B, Cheeneebash J 2009 Artificial neural network as a clinical decision-supporting tool to predict cardiovascular disease *Trends in Applied Sciences Research* **4**(1) 36-46

[5] Sekar B D, Dong M C, Shi J, Hu X Y 2012 *IEEE Sensors Journal* **12**(3) 644-50

[6] Torres F B J, Cavalcanti G D C, Ren T I 2013 *IEEE Transactions on Cybernetics* **43**(6) 2082-92

[7] Reznik L, Von Pless G, Al Karim T 2011 *IEEE Sensors Journal* **11**(3) 791-8

[8] Graves A, Liwicki M, Fernandez S, Bertolami R, Bunke H, Schmidhuber J 2009 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(5) 855-68

[9] Jin W, Li Z J, Wei L S, Zhen H 2000 The improvements of BP neural network learning algorithm *5th International Conference on Signal Processing Proceedings*, Beijing **3** 1647-9

[10] Mishra D, Yadav A, Ray S, Kalra P K 2005 Levenberg-Marquardt learning algorithm for integrate-and-fire neuron model *Neural Information Processing* **9**(2) 41-51

[11] Yuan J, He C, Gao W, Lin J, Yu P 2014 A novel hard decision decoding scheme based on genetic algorithm and neural network *Optik-International Journal for Light and Electron Optics* **125**(14) 3457-61

[12] Negnevitsky M 2004 Artificial intelligence-A guide to intelligent systems *Addison-Wesley Educational Publishers Inc*

[13] Yong C, Jian L, Wu C 2009 Environmental quality evaluation based on SVM *Computer Engineering and Applications* **45**(2) 209-11

[14] Hou Kefu 1992 Environmental system engineering *Beijing Institute of Technology Press (in Chinese)*

Authors

Yijun Liu, born in June, 1978, Changzhou, Jiangsu, China

Current position, grades: the lecturer of School of computer engineering, Jiangsu University of Technology, China.
University studies: M.Sc. from Nanjing University in China.
Scientific interest: machine learning, data mining and intelligent information system.
Publications: 10 papers.
Experience: teaching experience of 9 years, 4 scientific research projects.

Sheng He, born in November, 1971, Zongyang, Anhui, China

Current position, grades: the associate professor of school of computer engineering, Jiangsu University of Technology, China.
University studies: Ph.D. from Jiangnan University in China.
Scientific interest: data mining and bioinformatics.
Publications: 10 papers.
Experience: teaching experience of 15 years, 4 scientific research projects.

Yao Wang, born in April, 1982, Changzhou, Jiangsu, China

Current position, grades: the lecturer of school of computer engineering, Jiangsu University of Technology, China.
University studies: M.Sc. from Yangzhou University in China.
Scientific interest: computer network and intelligent information system.
Publications: 5 papers.
Experience: teaching experience of 10 years, 3 scientific research projects.

Xiumei Wang, born in October, 1971, Xuanhua, Hebei, China

Current position, grades: the lecturer of school of computer engineering, Jiangsu University of Technology, China.
University studies: B.Sc. in computer science and technology from Langfang Teachers University in China, M.Sc. from Brooklyn College, CUNY in USA.
Scientific interest: computability theory, dynamical systems.
Publications: 3 papers.
Experience: teaching experience of 10 years, 2 scientific research projects.