



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

A Robust Design Approach for GA-based Back Propagation Neural Networks Designed to Classify Data of Different Types

¹Chien-Yu Huang, ²I-Chiang Wang and ³Tien-Hui Chen

¹Department of Information Management, National Taitung Junior College, Taiwan

²Department of Marketing and Logistics Management, Southern Taiwan University, Taiwan

³Department of Travel Industry Management, Far-East University, Taiwan

Abstract: Genetic Algorithms (GAs) have a proven ability to improve the classification performance of Back-propagation Neural (BPN) networks by optimizing their topology and parameter settings. However, before they are used to optimize the BPN network, their parameters should be calibrated to improve the quality of the results. Accordingly, the current study develops a robust design method in which the Taguchi method is employed to establish appropriate values for the main GA parameters, namely the crossover rate, the mutation rate and the size of the population. The calibrated GA is used to optimize the parameters of BPN networks designed to classify three different types of data, continuous, ordinal and nominal, to immunize the noise of the data type. The classification performance of each GA-optimized BPN is verified using datasets downloaded from the server of the University of California's Department of Information and Computer Science. The results demonstrate that the process of calibrating the GA's parameters prior to its use in optimizing the BPN network yields a significant improvement in the network's classification performance.

Key words: Back-propagation neural network, genetic algorithm, classification

INTRODUCTION

Data classification methods provide an efficient means of resolving the "data overload" problem faced by many decision-makers. Typically, data classification methods are designed to identify the unique characteristics of the input data so that they can be assigned to appropriate classes. Common data classification methods include the nearest neighbor classifier, ID3 classification trees, fuzzy decision trees and Artificial Neural Networks (ANNs) (Lotfi *et al.*, 2006; Giordano *et al.*, 2007; Umer and Khiyal, 2007; Wu *et al.*, 2008; Gketsis *et al.*, 2009; Saravanan and Ramachandran, 2010; Wang *et al.*, 2011).

ANN schemes have been widely applied for such uses as forecasting and classification (Sun *et al.*, 2000; Zhang *et al.*, 2001; Srivaree-Ratana *et al.*, 2002; Ambrogi *et al.*, 2007; Fahimifard *et al.*, 2009; Solaimani, 2009; Gu *et al.*, 2011; Tahir and Manap, 2012). In addition, Back-propagation Neural (BPN) networks are one of the most commonly employed of all ANNs and have been successfully applied to solve many data classification problems (Salchenberger *et al.*, 1992; Nussbaum *et al.*, 1995; Liu *et al.*, 2001; Wakaf and Saii, 2009). However, BPN networks have notoriously unreliable performance when trained using the Gradient Steepest Descent Learning Algorithm (GSDLA). GSDLA-trained BPN networks have two major drawbacks, namely: (1) the

GSDLA converges slowly to the optimal solution and (2) when trapped into some local areas, the GSDLA tends to converge to local sub-optimal solutions during the iterative optimization procedure (Ikuno *et al.*, 1994).

Genetic Algorithms (GAs) have a proven ability to improve the classification performance of ANN schemes (Sexton *et al.*, 1998; Goffe *et al.*, 1994; Ramasamy and Rajasekaran, 1996; Panda *et al.*, 2007; Soltani *et al.*, 2007; Hsu *et al.*, 2009; Liu, 2010). GAs have also been successfully applied to identify the optimal BPN network topology and parameter settings Sexton compared the optimization performance between a genetic-algorithm-derived BPN network and conventional BPN network and found that in most cases genetic algorithms could improve the BPN network (Sexton *et al.*, 1998). Sexton *et al.* (1999) further compared the performance of optimizing the BPN network derived using Simulated Annealing (SA) and GAs on certain topologies of a BPN network model. The values assigned to the GA's parameters, i.e. the crossover rate, the mutation rate and the size of the population, have a direct impact on the results obtained for a particular optimization problem (Kirkpatrick *et al.*, 1983; Sexton *et al.*, 1999).

In decision-making, the data type and the distribution pattern of the data values have a direct bearing on the choice of statistical test used for their evaluation. It is therefore important to recognize or at least to immunize the attribute noise of the data points in a sample drawn to

represent the entire population. Broadly speaking, these attributes may be categorized as continuous, ordinal or nominal (Sheats and Pankratz, 2002).

Choosing a suitable classification model and assigning appropriate level settings to each factor within the model is a difficult task for even experienced decision-makers. Generally, trial-and-error experimental techniques are used to identify the significant factors and to assign appropriate factor level settings. However, such approaches are invariably time-consuming and thus expensive. Therefore, a straightforward and systematic technique for calibrating the parameters of a GA such that it can then be used to optimize the topology and parameter settings of a BPN network classification scheme is required.

Taguchi Orthogonal Arrays (OAs) (Taguchi, 1986) enable robust, noise-immune design solutions to be obtained from a minimum number of experimental trials. Many studies have demonstrated the use of OAs in calibrating heuristic algorithms. For example, Gupta (1999) considered the problem of a tabu search-based heuristic mechanism designed to solve the two-stage bow shop problem. In the proposed approach, the Taguchi method was used to analyze the effects of four factors, namely the initial solution, the type of move, the neighborhood size and the list size, on the performance of the heuristic mechanism. The results enabled the optimal factor settings to be determined and therefore improved the effectiveness of the heuristic algorithm. Chien and Tsai (2003) developed an analytical model for the prediction of tool flank wear and then used a GA-based optimization scheme, in which the parameters were calibrated using the Taguchi method, to determine the optimal cutting conditions when machining 17-4PH stainless steel. They showed that the model was capable of successfully predicting tool flank wear.

The Taguchi method has emerged as the method of choice for analyzing the main and interaction effects of the individual factors of a design problem and for screening and ranking these factors such that optimal design solutions can be obtained at a minimal experimental cost (Roy, 1990). In the present study, the Taguchi method is employed to calibrate the parameters of a GA scheme used to optimize the topology and parameter settings of a BPN network designed to classify data of different types.

The objectives of the present study can be summarized as follows:

- To provide decision-makers with the ability to immunize the noise of data types robustly such that

they can establish the data collection method and determine the construction of a highly efficient classification model

- To calibrate the GA used to design the BPN network parameters and topology to enhance its optimization performance
- To investigate the optimal combination of the GA's controllable factors and their level settings by using a robust noise-immune experimental design
- To use ANOVA and the Analysis of the Mean (ANOM) (Roy, 1990) to rank and screen the controllable factors of the GA to reduce the experimental cost

MATERIALS AND METHODS

This section describes the use of the Taguchi method to screen and rank the main and interaction effects of the controllable factors in the GA-based BPN classification model so that the network's classifying performance is rendered robust when the data types are immune to noise. The design process starts by identifying the various attributes of the data types of common interest to decision-makers. Appropriate level setting ranges are then specified for the controllable factors of the GA; i.e. the crossover rate, the mutation rate and the size of the population. These factors are then arranged in a Taguchi orthogonal array so that their main and interaction effects can be systematically examined when the GA is applied to optimize the topology and parameter settings of three BPN networks designed to classify data of different types. Figure 1 shows a schematic overview of the design methodology applied in the present study. The basic steps in this framework can be summarized as follows:

Step 1: Identify the attributes of different data types: In general, data classification problems involve the processing of a different number of attributes. In practice, this implies that the topology and parameter settings of the BPN network employed to carry out the data classification process depend on the nature of the input data. Accordingly, this step of the proposed methodology identifies the attributes of three representative datasets (containing continuous, ordinal and nominal data, respectively) downloaded from the server of the Department of Information and Computer Science (ICS) at the University of California

Step 2: Identify the BPN network topology and parameters: The BPN network parameters include

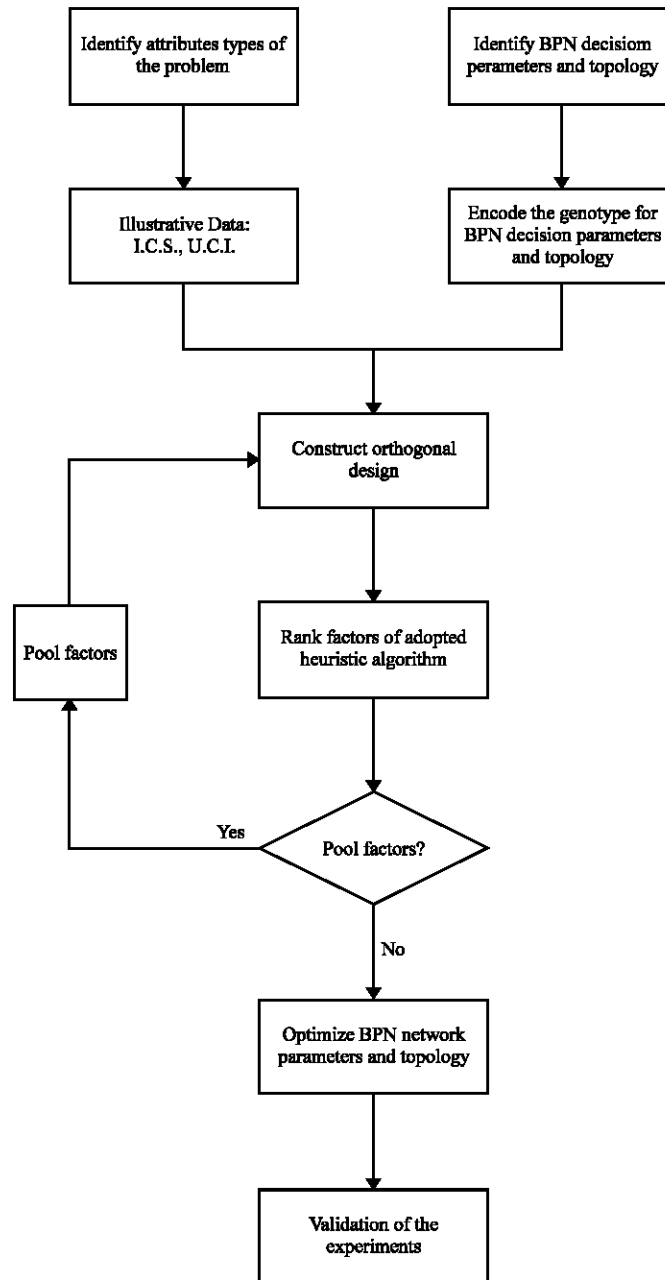


Fig. 1: Framework of proposed robust design procedure for GA-based BPN network

the initial weight value, learning rate, momentum factor and number of hidden layers, number of neurons in the first hidden layer and number of neurons in the second hidden layer. In this step, the calibrated genetic algorithm is used to specify appropriate network values for the BPN. Appropriate ranges are assigned to each of these BPN network parameters to form an initial search space for the GA optimization procedure

Step 3: Code the BPN network parameters in binary form: The chromosomes of the genetic algorithm should be in the form of a simple binary sequence comprising a series of 0 and 1 sec. This study thus uses a binary method to encode and decode the parameter and topology settings of the BPN model, which are optimized using a GA

Step 4: Identify the parameters of the GA used to optimize the BPN network model: Appropriate

level settings are assigned to the crossover rate, mutation rate and size of the population parameters of the GA to establish an initial search space for the Taguchi design procedure

Step 5: Rank and screen the GA factors: Employing the ANOVA statistical method, the effects of each of the main GA parameters in the BPN network (and the interactions between them) on the classification performance are analyzed so that the parameters with the greatest effect can be identified

Step 6: Calibrate and Pool the GA factors: Based on the ANOVA results obtained in step 5, the effects of pooling one or more of the controllable factors are considered so that an informed decision can be made regarding a potential trade-off between the quality of the robust design solution and the experimental cost

Step 7: Optimize the BPN parameters and topology: The calibrated GA is used to optimize the BPN topology and parameter settings

Definition of different data types: Broadly speaking, data can be classified as either continuous or categorical in nature (Table 1). Qualitative data are always categorical, whereas quantitative data are generally continuous but may be categorical in some cases. Categorical data may be further classified as either ordinal or nominal. Ordinal data (e.g., ranking or grading data) have an order to their levels of assignment, with each data point being more (or less) severe than the previous one. Conversely, nominal data (e.g., race or gender) have no inherent order to their levels of assignment (Sheats and Pankratz, 2002). Consequently, the data points in the sample used in this study include continuous data, ordinal data and nominal data.

Identifying the BPN network parameters and specifying appropriate ranges: Many studies have demonstrated the ability of GAs to optimize the topology and parameter settings of BPN networks so that their classification performance can be improved. However, before the GA can be applied to carry out this optimization process, it is first necessary to specify exactly which BPN network parameters are to be considered in the optimization procedure and to assign an appropriate range to each one. As discussed in step 2 above, six BPN network parameters are considered in the current GA optimization procedure, namely the initial weight value, the learning rate, the momentum factor, the number of hidden layers, the number of neurons in the first hidden layer and the number of neurons in the second hidden layer. Table 2 indicates the ranges assigned to each of these six

Table 1: Overview of basis data types

Data type	Description	Examples
Continuous	Variables that are measured and can take on any value along a continuum	Height weight
Categorical	Variables whose values fall into distinct categories or defined levels	- -
Ordinal	Variables for which an order exists in the levels assigned	Grading ranking
Nominal	Variables for which there is no hierarchical order to the category level	Race Sex

Table 2: BPN network parameters and setting ranges

Name of parameters and topology	Range
Initial weight values factor	0<real number<1.0
Learning rate	0<real number<1.0
Momentum factor	0<real number<1.0
Number of hidden layers	1<integer<2
Number of neurons in the first hidden layer	1<integer<63
Number of neurons in the second hidden layer	1<integer<63 (if hidden layers>1)

parameters. In accordance with conventional BPN network theory, the initial weight value, the learning rate and the momentum factor parameters are assigned ranges of 0~1.0. Although one hidden layer is generally sufficient in most BPN networks, two hidden layers may be required for more complex classification problems. Accordingly, in the current study, the hidden layer parameter is specified as either 1 or 2. The number of neurons in each hidden layer is related to the total number of hidden layers in the BPN structure (Zurada, 1995; Khaw *et al.*, 1995; Sexton *et al.*, 1998). As a result, in the current study, the number of neurons in the first and second hidden layers is specified in a more expanded range of 1~63 in both cases.

Code the BPN parameters in binary form: In general, GA optimization procedures commence with a population of randomly-chosen chromosomes, where each chromosome represents a potential solution to the specified problem (Khoo and Loi, 2002; Kamrani and Gonzalez, 2003; Siani and de Peretti, 2007). For computational simplicity, Goldberg (1989) proposed that these chromosomes should be expressed in the form of a simple binary sequence comprising a series of 0 and 1 bits. The search space for the current GA-based BPN optimization process is defined by the BPN network parameter settings specified in Table 2. Therefore, in accordance with the GA methodology, each parameter range is transformed into a binary sequence and the individual binary sequences are then concatenated to form a single chromosome string. As shown in Fig. 2, this process results in the construction of a string with a total of 39 bits of binary code. The initial weight range (i.e., 0~1.0) is mapped using a seven-bit sub-string. When decoding this string back into a real number, the maximum decoded value is given by $2^7-1 (= 127)$. Therefore, having identified the optimal

chromosome via the GA iterative procedure, the corresponding real value of the weight factor is given by the quotient obtained when decoding the optimal seven-bit sub-string and then dividing the result by 127. As shown in Fig. 2, the learning rate, momentum factor, number of hidden layers, number of neurons in the first hidden layer and number of neurons in the second hidden layer are coded using eight-bit, ten-bit, two-bit, six-bit and six-bit sub-strings, respectively. The real optimal values of each parameter are obtained using the same decoding and division procedure as that described above for the weight factor.

Construct orthogonal design: The OAs used in the Taguchi method enable the optimal design parameters to be established from a minimum number of experiments and therefore reduce the total experimental cost. Moreover, the use of an OA enhances the reproducibility of the experimental results and enables an objective setting of the experimental factor levels. Various OA layouts are commonly employed and the choice of an appropriate array depends on the degrees of freedom of the particular experiment. The OA notation $L_4 (2^3)$ indicates that the experiment requires a total of four trial runs, where each trial run involves a maximum of three factors, the values of which can be assigned on one of two different levels. Similarly, an $L_9 (3^4)$ OA prescribes nine experimental runs, each involving a maximum of four factors with three permissible level settings. As shown in Table 3, the Taguchi procedure used to calibrate the controllable factors of the GA is performed using an OA of the latter type. In this table, the four columns correspond to the three controllable factors of the GA (i.e., the crossover rate, mutation rate and the size of population and the

	0101010	10101010	0101010101	10	010101	101010
Sub-string	1	2	3	4	5	6

Fig. 2: Binary code representations of BPN network parameters

Table 3: $L_9 (3^4)$ orthogonal array for current classification model of GA

L_9 experiment/column	A	B	AB	C	(E_{RMEN})			
					Glass	Dermatology	Hayes-Roth	S/N_i
1	1	1	1	1	0.0244778	0.00946149	0.00848975	35.959
2	1	2	2	2	0.0282868	0.00982626	0.00389588	35.172
3	1	3	3	3	0.0221551	0.00792481	0.00352253	37.243
4	2	1	2	3	0.0202524	0.00811437	0.00520779	37.754
5	2	2	3	1	0.0279098	0.00972380	0.00521220	35.226
6	2	3	1	2	0.0274912	0.00937310	0.00363589	35.442
7	3	1	3	2	0.0267135	0.01121120	0.00369510	35.462
8	3	2	1	3	0.0225782	0.00882387	0.00411497	36.957
9	3	3	2	1	0.0212302	0.00922303	0.00519036	37.268

A: Crossover rate, B: Mutation rate, C: Size of population, AB: Intraction between crossover and mutation rate

interaction between the crossover rate and the mutation rate). The table indicates that in the first experimental run, the three controllable factors and the interaction parameter are all assigned their Level 1-settings. Similarly, in the second experimental run, the crossover rate is assigned its Level-1 setting, while the mutation rate, size of population and interaction effect are all assigned their Level-2 settings. The remaining trial runs are arranged in such a way as to achieve an orthogonal property (Roy, 1990).

The quality of the design solution achieved in each experimental run is assessed using the following signal-to-noise metric:

$$S/N = -10\text{Log}_{10}(\text{MSD})$$

where, MSD is the mean-squared deviation of the experimental result from the target value of the specified quality characteristic, i.e., $\text{MSD} = (y_1^2 + y_2^2 + \dots + y_n^2)/n$ where y_1, y_2, \dots etc. are the deviation of the experimental result y_i from the target value (where $i = 1 \dots n$) and n is the total number of repetitions of y_i .

Rank selected factors: Having performed the experimental trials prescribed in the OA, the experimental outcomes are analyzed using conventional ANOVA and ANOM statistical methods. In the analysis procedure, the ANOVA statistical technique is applied to rank the controllable factors of the GA (and the interaction effect between them) in terms of the extent to which they affect the classification performance of the BPN. The ANOM approach is then used to establish the optimal level settings for each of the controllable factors and the interaction term.

Pool selected factors: The ANOVA-based ranking and screening results may indicate that one or more of the controllable factors can be pooled, since variations in their values have a negligible effect on the BPN network classification performance. Having pooled the least

influential factors, a further optimization experiment is then conducted using the remaining controllable factors to ensure that the quality of the solution is not seriously degraded.

Optimize the BPN network topology and parameters using calibrated GA: In the discussions above, in the calibration process the GA parameters are assigned in accordance with the level settings in each row of the orthogonal array and the GA is then used to search for the optimal BPN parameter settings for each of the classification problems (Glass, Dermatology and Hayes-Roth in this study). Finally, the classification performance of each BPN network is evaluated by computing (E_{RMSN}), as shown in Table 3. Once the GA has been calibrated, it is used to optimize the BPN networks used to solve the three data classification problems.

In general, the quality of the BPN network classification results is evaluated using the root-mean-square normalized error (E_{RMSN}) metric:

$$E_{RMSN} = \frac{1}{P \cdot K} \sqrt{\sum_{i=1}^P \sum_{j=1}^K (d_{ij} - o_{ij})^2}$$

where, P is the number of input examples, K is the number of output neurons, d_{ij} is the expected output value of the jth neuron in the ith example and o_{ij} is the actual output value of the jth neuron in the ith example.

Illustrated example: In this study, the performance of GA-optimized BPN networks was evaluated by considering their effectiveness in solving data classification problems involving three different types of data, namely continuous, ordinal and nominal.

Table 4 summarizes the characteristics of the databases downloaded for evaluation purposes from the server of the Department of Information of Computer Science (ICS) at the University of California. The major features of these three databases are discussed.

Glass database: Continuous data values; 100 training samples and 258 testing samples. Glass splinters left at the scene of a crime may provide investigators with valuable clues. Therefore, a typical data classification problem involves analyzing the properties of a glass splinter such that its type can be identified. The nine attributes of the data records in this database include nine input attributes, namely the refractive index, sodium, magnesium, aluminum, silicon, potassium, calcium, barium and iron and one output attribute, the glass type (float-processed building windows, non-float processed building windows,

Table 4: Problem data types

Data type	No. of attributes	No. of training	No. of tests	Data base
Continuous	9	100	258	Glass
Ordinal	34	100	114	Dermatology
Nominal	4	100	32	Hayer-Roth

float-processed vehicle windows, non-float processed vehicle windows, containers, tableware or headlamps). Therefore, in using the GA to optimize the structure of a BPN to perform this classification process, a total of nine neurons are specified in the input layer and one neuron is specified in the output layer.

Dermatology database: Ordinal values; 100 training samples and 114 testing samples. In the dermatological field, performing differential diagnoses of erythematous diseases is a common requirement. The dataset includes 34 input attributes and a single output attribute. As a result, during the GA optimization procedure, the input and output layers are assumed to have 34 neurons and one neuron, respectively. In the dataset created for this domain, the input samples are classified in a range of 0 to 3 to indicate the presence or absence of 34 features. Every feature used for classification purposes is assigned a degree in the range of 0 to 3, where 0 indicates that the feature is not present, 1 and 2 indicate the relative intermediate values and 3 indicates the largest amount possible. The seven diseases classified in the output neuron are psoriasis, seboric dermatitis, lichen planus, pityriasis rosea, cronic dermatitis and pityriasis rubra pilaris.

Hayes-Roth database: Nominal values; 100 training samples and 32 testing samples. Each record in this dataset has four input attributes, namely hobby, age, education and marital status and one output attribute corresponding to three different classes. As a result, during the GA calibration procedure the BPN is specified as having four nodes in the input layer and a single node in the output layer.

Calibration of GA control factors using Taguchi method: The controllable factors of the GA were calibrated by performing a Taguchi design process in which the GA was used to identify the optimal topologies and parameter settings of the BPNs required to solve the classification problems described above. In performing the Taguchi design process, the controllable factors of the GA were assigned three levels, i.e., crossover rate (0.5, 0.7, 0.95) (Maniezzo, 1994; Srinivas and Patnaik, 1994), mutation rate (0.001, 0.01, 0.05) (Perez and Holzmann, 1997) and size of population (10, 20, 30) (Maniezzo, 1994; Perez and Holzmann, 1997). The treatment levels of these controllable factors are summarized in Table 5.

Table 5: Factor levels of GA

Factors	Level-1	Level-2	Level-3
Crossover rate	0.5	0.7	0.95
Mutation rate	0.001	0.01	0.05
Size of population	10	20	30

RESULTS AND DISCUSSION

As discussed in Previous Section, the controllable factors of the GA were calibrated using a Taguchi design procedure based on an $L_9 (3^4)$ OA. As shown in Table 3, this array prescribes nine experimental trials involving four factors, each with three level settings. Columns 1, 2 and 4 are assigned to the controllable factors as follows: Column 1: Factor A (crossover rate), Column 2: Factor B (mutation rate) and Column 4: Factor C (size of population). Column 3 is used to investigate the effect of the interaction between factors A and B. The numerals in columns A, B and C correspond to the factor level settings shown in Table 5. In the calibration procedure, the optimal BPN network structure for each of the three data classification problems was identified using the factor level settings specified in the nine experimental trials. The quality of the corresponding solution (i.e., the classification performance of the resulting BPN network) was evaluated using the E_{RMSN} metric. The corresponding results are presented in the Glass, Dermatology and Hayer-Roth columns of Table 3.

Ranking selected factors: Having conducted the experimental trials, the E_{RMSN} data were analyzed using the ANOVA statistical method to rank the controllable factors and the interaction factor in terms of their relative influences on the classification performance of the BPN network. The full details of this analysis process are presented in Roy (1990) and are therefore omitted here. The ANOVA analysis results for the current Taguchi experiments are presented in Table 6-8, in which the column headings are defined as follows:

- f = Degree of freedom
- S = Sum of squares
- S' = Pure sum of squares (pooled)
- V = Mean squares (variance)
- F = Variance ratio
- P = Percentage contribution

Note that in the discussions which follow, the term “interaction” (indicated by the symbol “×” between the two interacting factors) describes the condition in which the effect of one factor’s influence upon the result is dependent on the level setting of the other. The right-hand column of Table 6 indicates that the percentage contributions of factors A, B and C are 4.4, 14.2 and 69.8%, respectively, while that of the interaction term A×B

Table 6: ANOVA analysis for GA controllable factors

Column	Factors	f	S	V	P
1	A	2	0.36977	0.18489	4.4
2	B	2	1.18620	0.59309	14.2
3	Interaction A×B	2	0.96498	0.48249	11.6
4	C	2	5.82840	2.91420	69.8
	Other error	0	0	0.00000	0.
	Total	8	8.34930		100%

is 11.6%. In other words, these factors can be ranked in terms of reducing influence as follows: factor C, factor B and factor A. To evaluate the potential for reducing the experimental cost and difficulty, the following section considers the step-by-step pooling of factor A, the interaction term A×B and Factor B, respectively and evaluates the corresponding effects on the quality of the design solution.

Factor pooling analysis: Table 7 presents the ANOVA analysis results for the case where Factor A and interaction factor A×B are pooled. It can be seen that factor C (with an F value of 8.7332 in Table 7) has a significantly greater effect on the classification performance of the BPN than factor B. Conducting an F-test, the F values are found to be $F_{v_1=2, v_2=4} = 6.9443$ at a 95% confidence level (F Table $F_{0.05}(f_1, f_2)$, 95% confidence). Table 7 shows that $F_B = 1.7774$ and $F_C = 8.7332$. Since, $F_B = 1.7774$ is less than the value computed in the F-test, i.e., $F_{v_1=2, v_2=4} = 6.9443$, factor B can also be pooled with factor A and the interaction factor A×B. The corresponding ANOVA analysis results are presented in Table 7 and 8. The results indicate that the value of Factors A and B (i.e., the crossover rate and mutation rate) can be arbitrarily assigned within the defined range without significantly influencing the performance of the GA-optimized BPN network classification scheme.

Using calibrated GA to identify the BPN network parameters and topology: As described previously, the optimal levels of each GA parameter are obtained using the ANOM statistical method. Observing the entries in Column 1 of Table 3 corresponding to factor A, it can be seen that “Level 1” occurs in experimental runs 1, 2 and 3. The average effect of A_1 (i.e., the Level-1 setting of factor A) can be calculated by processing the S/N ratios corresponding to these three experiments as follows:

$$\bar{A}_1 = (SN_1 + SN_2 + SN_3) / 3 = (35.959 + 35.172 + 37.243) / 3 = 36.125$$

The average effects of the level-2 and level-3 settings for factor A can be computed in a similar manner. The procedure is then repeated for the three level settings of factors B and C and the interaction term. The corresponding results are summarized in Table 9 and it

Table 7: ANOVA analysis for GA controllable factors (factors A, A×B pooled)

Column	Factors	f	S	V	F	S'	P
1	Factor A	2	0.36977	Pooled			
2	Factor B	2	1.18620	0.59309	1.7774	0.5188	6.2
3	Interaction A×B	2	0.96498	Pooled			
4	Factor C	2	5.82840	2.9142	8.7332	5.1610	61.8
Other error		4	1.33475	0.33369		1.3348	32.0
Total		8	8.34930				100%

Table 8: ANOVA analysis for GA controllable factors (factors A, B, A×B pooled)

Column	Factors	f	S	V	F	S'	P
1	Factor A	2	0.36977	Pooled			
2	Factor B	2	1.18620	Pooled			
3	Interaction A×B	2	0.96498	Pooled			
4	Factor C	2	5.82840	2.9142	6.936	4.9881	0.59742
Other error		6	2.52090	0.42016		2.5209	0.40258
Total		8	8.34930				100%

Table 9: ANOM analysis of main effects (S/N)

Factor	Level 1	Level 2	Level 3
Crossover rate A	36.125	36.141	36.562
Mutation rate B	36.392	35.785	36.651
Interaction A×B	36.119	36.731	35.977
Size of population C	36.151	35.359	37.318

can be seen that the S/N ratio of factor A is maximized when the level-3 setting is specified. In the Taguchi experiments, a higher S/N ratio indicates improved BPN network classification performance. Hence, the results presented in Table 3 indicate that the GA parameters should be specified as follows: Factor A: Level 3 (cross-over rate of 0.95); Factor B: Level 3 (mutation rate of 0.05); and Factor C: Level 3 (size of population of 30).

Effects of calibrating the GA factors on the BPN classification performance: The experimental results were performed using MATLAB version 6.5 (Math Works) running on a Pentium(R) CPU 1.73 GHz PC.

It can be deduced that when all three of the GA's controllable factors are calibrated, the BPN network classification performance is 0.0105 (simulation test). When only Factors B and C are calibrated (i.e., as considered in Table 7), the BPN classification performance is found to be 0.0111. When only Factor C is calibrated (i.e., as considered in Table 8), the BPN network classification performance is 0.0114. Finally, when none of the controllable factors are calibrated, the BPN network classification performance is 0.0134. RMSE is reduced from 0.0134 to 0.0105; the error rate represented by RMSE is effectively reduced by about 21.64%. These results confirm that the use of the orthogonal array in calibrating the GA's controllable factors extends the solution search space, thereby increasing the probability of achieving improved performance.

CONCLUSION

This study employed the Taguchi design method to calibrate the controllable factors of a GA used to optimize

the parameters of BPN networks designed to classify data of different types, namely continuous, ordinal and nominal. Overall, the results show that the calibrated GA yields a significant improvement in the classification performance of the BPN networks.

Furthermore, the ANOM and ANOVA results not only indicate appropriate level settings for each of the controllable factors in the GA but also indicate which of the controllable factors have the greatest effect on the BPN performance, thereby enabling decision-makers to identify those factors which need to be carefully controlled.

Tracking the genetic algorithm used in this study increases the search space to raise the possibility of finding the optimal solution. In conclusion, the approach proposed in this study provides a method for developing robust noise-immune classification models for a variety of data types.

REFERENCES

- Ambrogi, F., N. Lama, P. Boracchi and E. Biganzoli, 2007. Selection of artificial neural network models for survival analysis with genetic algorithms. *Comput. Stat. Data Anal.*, 52: 30-42.
- Chien, W.T. and C.S. Tsai, 2003. The investigation on the prediction of tool wear and the determination of optimum cutting conditions in machining 17-4PH stainless steel. *J. Mater. Process. Technol.*, 140: 340-345.
- Fahimifard, S.M., M. Homayounifar, M. Sabouhi and A.R. Moghaddamnia, 2009. Comparison of ANFIS, ANN, GARCH and ARIMA techniques to exchange rate forecasting. *J. Applied Sci.*, 9: 3641-3651.
- Giordano, F., M. La Rocca and C. Perna, 2007. Forecasting nonlinear time series with neural network sieve bootstrap. *Comput. Stat. Data Anal.*, 51: 3871-3884.

- Gketsis, Z.E., M.E. Zervakis and G. Stavrakakis, 2009. Detection and classification of winding faults in windmill generators using wavelet transform and ANN. *Electr. Power Syst. Res.*, 79: 1483-1494.
- Goffe, W.L., G.D. Ferrier and J. Rogers, 1994. Global optimization of statistical functions with simulated annealing. *J. Econ.*, 60: 65-99.
- Goldberg, D.E., 1989. *Genetic Algorithms in Search Optimization and Machine Learning*. Addison-Wesley, New York, USA.
- Gu, Y., Y. Li, J. Xu and Y. Liu, 2011. Novel model based on wavelet transform and GA-fuzzy neural network applied to short time traffic flow prediction. *Inform. Technol. J.*, 10: 2105-2111.
- Gupta, J.N.D., 1999. Designing a tabu search algorithm for the two-stage flow shop problem with secondary criterion. *Prod. Plan. Control: Manage. Oper.*, 10: 251-265.
- Hsu, H.M., Y. Hsiung, Y.Z. Chen and M.C. Wu, 2009. A GA methodology for the scheduling of yarn-dyed textile production. *Expert Syst. Appl.*, 36: 12095-12103.
- Ikuno, Y., H. Hawabata, Y. Shirao, M. Hirata, T. Nagahara and Y. Inagaki, 1994. Application of an improved genetic algorithm to the learning of neural networks. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, E77: 731-735.
- Kamrani, A.K. and R. Gonzalez, 2003. A genetic algorithm-based solution methodology for modular design. *J. Intell. Manuf.*, 14: 599-616.
- Khaw, J.F.C., B.S. Lim and L.E.N. Lim, 1995. Optimal design of neural networks using the Taguchi method. *Neurocomputing*, 7: 225-245.
- Khoo, L.P. and M.Y. Loi, 2002. A tabu-enhanced genetic algorithm approach to agile manufacturing. *Int. J. Adv. Manuf. Technol.*, 20: 692-700.
- Kirkpatrick, S., C.D. Gelatt Jr. and M.P. Vecchi, 1983. Optimization by simulated annealing. *Science*, 220: 671-680.
- Liu, Y., W. Liu and Y. Zhang, 2001. Inspection of defects in optical fibers based on back-propagation neural networks. *Opt. Commun.*, 198: 369-378.
- Liu, Y.H., 2010. Different initial solution generators in genetic algorithms for solving the probabilistic traveling salesman problem. *Applied Math. Comput.*, 216: 125-137.
- Lotfi, F., F. Nadir and B. Mouldi, 2006. Arabic words recognition by fuzzy classifier. *J. Applied Sci.*, 6: 647-650.
- Maniezzo, V., 1994. Genetic evolution of the topology and weight distribution of neural networks. *IEEE Trans. Neural Networks*, 5: 39-53.
- Nussbaum, M.A., D.B. Chaffin and B.J. Martin, 1995. A back-propagation neural network model of lumbar muscle recruitment during moderate static exertions. *J. Biomech.*, 28: 1015-1024.
- Panda, B.P., M. Ali and S. Javed, 2007. Fermentation process optimization. *Res. J. Microbiol.*, 2: 201-208.
- Perez, C.A. and C.A. Holzmann, 1997. Improvements on handwritten digit recognition by genetic selection of neural network topology and by augmented training. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Computational Cybernetics and Simulation, Volume 2, October 12-15, 1997, Orlando, FL.*, pp: 1487-1491.
- Ramasamy, J.V. and S. Rajasekaran, 1996. Artificial neural network and genetic algorithm for the design optimization of industrial roofs-a comparison. *Comput. Struct.*, 58: 747-755.
- Roy, K.R., 1990. *A Primer on the Taguchi Method*. Van Nostrand Reinholds, New York.
- Salchenberger, L.M., E.M. Cinar and N.A. Lash, 1992. Neural networks: A new tool for predicting thrift failures. *Decis. Sci.*, 23: 899-916.
- Saravanan, N. and K.I. Ramachandran, 2010. Incipient gear box fault diagnosis using Discrete Wavelet Transform (DWT) for feature extraction and classification using Artificial Neural Network (ANN). *Expert Syst. Appl.*, 37: 4168-4181.
- Sexton, R.S., R.E. Dorsey and J.D. Johnson, 1998. Toward global optimization of neural networks: A comparison of the genetic algorithm and backpropagation. *Decision Support Syst.*, 22: 171-185.
- Sexton, R.S., R.E. Dorsey and J.D. Johnson, 1999. Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing. *Eur. J. Operat. Res.*, 114: 589-601.
- Sheats, R.D. and S. Pankratz, 2002. Understanding distributions and data types. *Semin. Orthodontics*, 8: 62-66.
- Siani, C. and C. de Peretti, 2007. Analysing the performance of bootstrap neural tests for conditional heteroskedasticity in ARCH-M models. *Comput. Stat. Data Anal.*, 51: 2442-2460.
- Solaimani, K., 2009. A study of rainfall forecasting models based on artificial neural network. *Asian J. Applied Sci.*, 2: 486-498.
- Soltani, S., H. Babaei, K.A. Zeynali and A. Jouyban, 2007. Modelling vasorelaxant activity of some drugs/drug candidates using artificial neural networks. *J. Pharmacol. Toxicol.*, 2: 411-426.
- Srinivas, M. and L.M. Patnaik, 1994. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Trans. Syst. Man Cybernet.*, 24: 656-667.

- Srivaree-Ratana, C., A. Konak and A.E. Smith, 2002. Estimation of all-terminal network reliability using an artificial neural network. *Comput. Oper. Res.*, 29: 849-868.
- Sun, M., A. Stam and R.E. Steuer, 2000. Interactive multiple objective programming using tchebycheff programs and artificial neural networks. *Comput. Oper. Res.*, 27: 601-620.
- Taguchi, G., 1986. *Introduction to Quality Engineering*. Asian Productivity Organization, New York.
- Tahir, N.M. and H.H. Manap, 2012. Parkinson disease gait classification based on machine learning approach. *J. Applied Sci.*, 12: 180-185.
- Umer, M.F. and M.S.H. Khiyal, 2007. Classification of textual documents using learning vector quantization. *Inform. Technol. J.*, 6: 154-159.
- Wakaf, Z. and M.M. Sarii, 2009. Frontal colored face recognition system. *Res. J. Inform. Technol.*, 1: 17-29.
- Wang, Z., D. Cui and Q. Xu, 2011. Artificial endocrine system and its application for nearest neighbor rule condensation. *Inform. Technol. J.*, 10: 2004-2013.
- Wu, T.K., S.C. Huang and Y.R. Meng, 2008. Evaluation of ANN and SVM classifiers as predictors to the diagnosis of students with learning disabilities. *Expert Syst. Appl.*, 34: 1846-1856.
- Zhang, G.P., G.E. Patuwo and M.Y. Hu, 2001. A simulation study of artificial neural networks for nonlinear time-series forecasting. *Comput. Operat. Res.*, 28: 381-396.
- Zurada, J.M., 1995. *Introduction to Artificial Neural Systems*. PWS Publishing, Boston, MA, USA.