International Journal of Scientific Research and Engineering Development-– Volume 2 Issue 4, July – Aug 2019

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Case Analysis of Police Offices using of Back Propagation Neural Network with Genetic Algorithm

Aung Kyaw Oo*, Aye Zarchi Minn**

*(Department of Information Technology, Technological University (Taunggyi), and Taunggyi, Myanmar)

** (Department of Information Technology, Technological University (Taunggyi), and Taunggyi, Myanmar)

----------**********----------

Abstract:

Data mining is a procedure of extraction of information and knowledge that are hided in data, unknown by people and potentially useful from a large quantity of data with multiple characteristics. The data-mining model based on genetic neural network has been widely applied to the procedure of data mining. This paper presents a method that combines the learning algorithm of Back Propagation (BP) neural network with genetic algorithm to train BP network and optimize the weight values of the network. The system uses a single hidden layer. Input for the system is training data set of police office data. Each attribute values are all numeric-valued. At the beginning, each training data set is trained using Neural Network architecture with sigmoid function. Then, by using GA, neural network weights are chosen as parents for selection, crossover and mutation. The weights are optimized, using genetic algorithm. The fitness is calculated and the selection is performed with method of roulette wheel selection. This system can be solved cases of police office rapidly and make decisions efficiently using the historical data.

Keywords **—Data mining, Neural Network, Back Propagation (BP), Genetic Algorithm**

----------**********----------

I. INTRODUCTION

Neural networks (NNs) and genetic algorithm (GAs) are computational abstractions of biological information processing systems, and both have taken the imaginations or researcher all over the world. In general, NNs are used as learning systems and GAs as optimization system, but as many researchers have exposed, they may be joined in a number of system.

A neural network is a computational model containing of a number of connected elements, known as neurons. A neuron is a processing unit that accepts input from neurons, applies a local transformation to that input, and delivers a single output signal which approved on to other neurons. Each of the input is changed by a value associated with the connection strength, or weight represents how much importance the neuron assigns to that input source. The local transformation is referred to as the activation function and usually sigmoid in nature. Alternatively, in a collaborative approach, the GA and NN are combined into a single system in which a population of neural networks is evolved. The objective of the system is to find the optimal neural network solution. A genetic algorithm can be applied to optimizing a neural network in a variety of way. The populations of the GA are all NNs with the same basic topology with different weight values. Using a GAs instead of gradient descent algorithm to train the weights can result in faster and better convergence [1].

The chiefpurposes of this paper areto achieve an effect for assisting people to solve cases and make good decisions, to study a method that combines the learning algorithm of backpropagation (BP) neural network with genetic algorithm, to enhance the neural network weight values by combining BP

algorithm with genetic algorithm and to resolve cases of police office rapidly and make decisions efficiently using the historical data.

II. BACK PROPAGATION NEURAL NETWORK

Back Propagation (BP) neural network is a kind of feed-forward network that is now in most common use. Generally it has a multi-layer structure that contains at least three layers covering one input layer, one output layer and one or more hidden layers. There are full connections between neurons in adjacent layers and no connection between neurons in the same layer. Based on a set of training samples and a set of testing data, BP neural network trains its neurons and complete the procedure of learning. The application of BP algorithm is appropriate for data mining environment in which it is impossible to solve problems using ordinary methods. Therefore, complex function of several variables is used to complete non-linear calculation to accomplish the semi-structural and non-structural decision-making supporting procedure [2]. So, in the procedure of data mining in the command centre of police office, the BP neural network model is selected.

The basic structure of BP neural network is as follow. The learning procedure of neural network can be separated into three phases. The first one is a forward propagation phase in which a specified input pattern has been passed through the network from input layer through hidden layers to the output layer and becomes an output pattern. The second one is an error back propagation phase. In this phase, BP algorithm compares the real output and the expected output to calculate the error values. After that, it propagates the error values from output layer through hidden layer to input layer in the opposite direction. The connection weights will be altered during this phase. These three phases proceed repeatedly and alternately to complete the memory training of network until it tends to convergence and the global error tends to minimum.

The weights of the network to be trained are typically initialized at small random values. The initialization strongly affects the ultimate solution. If all weights start out with equal weight values, and if the solution requires that unequal weights be developed, the network may not train properly. Unless the network is distributed by random factors or the random character of input patterns during training, the internal representation may continuously result in symmetrical weights. Also, the network may fail to learn the set of training examples with the error stabilizing or even increasing as the learning continues. In fact, many empirical studies of the algorithm point out that continuing training beyond a certain low-error plateau results in the undesirable drift of weights. This causes the error to increase and the quality of mapping implemented by the network decreases. To counteract the drift problem, network learning should be restarted with other random weights [4]. Generally, this system generate random nonzero floating numbers in [-1, 1] as the initial weight values. The choice of initial weights is, however, only one of several factors affecting the training of the network toward an acceptable error minimum.

III. PRINCIPLE OF GENETIC ALGORITHM

Genetic algorithm is a kind of search and optimization model built by simulating the lengthy evolution period of heredity selection and natural elimination of biological colony. It is an algorithm of global probability search. It doesn't depend on gradient data and needn't the differentiability of the function that will be solved and only need the function can be solved under the condition of constraint [3]. So by using genetic algorithm to optimize the weights of BP neural network, this system can eliminate the problems of BP network and enhance the generalization performance of the network. The elementary operation of genetic algorithm consists of three operands: selection, crossover and mutation. Selection is also called copy or reproduction. By calculating the fitness of individuals, this system selects high quality

individuals with high fitness, copy them to the new population and eliminate the individual with low fitness to generate the new population. Generally used strategies of selection include roulette wheel selection, expectation value selection, paired competition selection and retaining high quality individual selection. Crossover puts individuals in population after selection into match pool and randomly makes individuals in pairs to form parent generation. Then according to crossover probability and the specified method of crossover, it exchanges part of the genes of individuals that is in pairs to form new pairs of child generation and finally to generate new individuals. Generally used methods of crossover are one point crossover, multi point crossover and average crossover. This system uses one point crossover. According to specified mutation rate, mutation substitutes genes with their opposite genes in some loci to generate new individuals.

A. Fitness Function

The fitness function defines the criterion for ranking potential solution and for probabilistically selection them for inclusion in the next generation. The fitness function or evaluation function plays an important role in GA. The fitness function is used to decide how 'good' a chromosome is. The evaluation function usually comes straight from the problem.

B. Selection

Selection is the procedure for choosing individuals as parents on which to perform crossover in order to create new solutions. The process of selection specifies which chromosomes from one generation will be sources for chromosomes in the next generation. The idea is that the 'fitter' individuals are more prominent in the selection process [4].

C. Stopping Criteria

The stopping criterion is used in the evaluation process to determine whether or not the current generation and the best solution found so far are close to the global optimum. Various stopping criteria can be used, and usually more than is employed. The standard stopping criteria is use to stop the procedure after the given number of iterations. Another stopping criteria is use to stop after the 'best' solution has not changed over a specified number of iterations. Another stopping criteria is used when the average fitness of the generation is the same or closes to the fitness of the 'best' solution. In this system, the stopping criterion is used to stop the procedure after a given number of iterations or also is controlled to stop after the best solution has not changed depending on a pre-specified error value by user[5].

IV. SYSTEM DESIGN OF THE STUDY

For this system as shown in Fig.1, the multi-layer network is used. Weights are first initialized among input nodes at input layer, one hidden layer and one output layer. Attributes numbers are set input nodes at input layer are makeup of. The system analyses people's gender, age, education degree, history of crime, salary level, alcohol, opium, family status, gambling. And, to some extent these factors affect people's social actions that may lead to commit a crime. Using these factors as input variables, a genetic neural network can be utilized to predict the present crime possibility of these people.

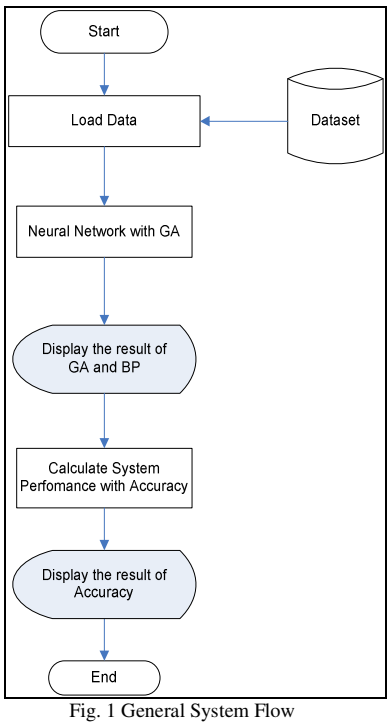

Fig. 1 General System Flow

At the beginning, the input training dataset is loaded from database. After that the training dataset is trained with backpropagation algorithm combined with Genetic Algorithm. The system displays the value of global error, fitness value, weights values, and crossover and mutation values during training time for each iteration. Then, it finally displays the output result of the given features.

A. Data Mining in Command Centre of Police Office

Every day in the command centre of police office, people receive a large number of information about cases received with various approaches. The information has been input into database to form a large amount of case information. These case information has been archived annually and periodically to form a plenty of historical case resources. By inducing and analysing these historical cases, people can get some experiences and learn some lessons that can help them solve cases and make decisions in the future. Therefore, in order to assist police departments to solve cases rapidly and make decisions efficiently, we should synthesize and organize these historical data, use proper data mining models to discover the potential and useful knowledge behind the data, and then predict and analyse the important factors in the data including the rate of crime, the constitution of crime population, the crime age structure, the area distribution of crime, the developing tendency of crime, the means and approaches of crime, the hidden areas of criminals and so on. At present all of these have become urgent tasks that need our police office to accomplish in the procedure of data processing.

B. Neural Network Train with Genetic Algorithm

In this system, the multilayer networks are used, first initializing weights made up of input nodes, hidden nodes and output nodes from the training dataset. Second, initialization is processed on weights with small random values between (+1) and (-1). They are the population of the Genetic Algorithm. The population size is user defined. Then, calculation is performed on the output of the hidden layer neurons and output layer neuron used sigmoid function. Using genetic algorithm and backpropagation algorithm, the errors in the output must be found out to calculate the fitness.

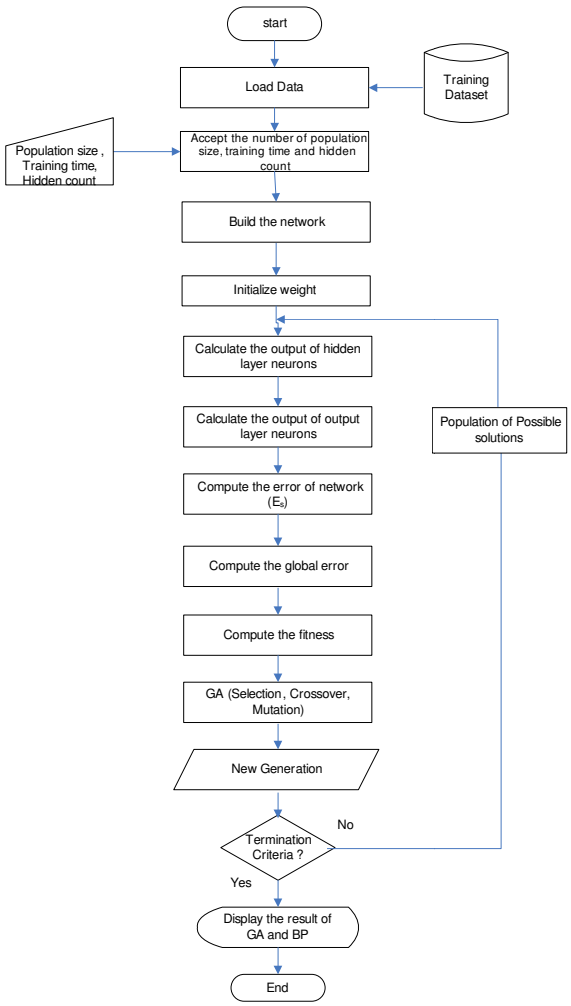

Fig. 2 System Flow of Neural Network Weights by using Genetic Algorithm

Then, the weights are evaluated by using Genetic Algorithms and backpropagation algorithm. The maximum fitness population is selected as parent populations. These populations must be calculated with crossover and mutation. The new populations are obtained and these populations are also called new generation. These generations are trained with the same training sample. The stopping criterion is used to stop the procedure after a given number of iterations or also is controlled to stop after the best solution has not changed depending on a pre-specified error value by user. It seems that Genetic Algorithms have proved suitable for solving the Neural Network weights. Both Neural Network and Genetic Algorithm are trained with police office training dataset. User is easy to understand this system.

V. IMPLEMENTATION OF THE SYSTEM

There are four main parts in this system. These are

- Data inputs
- Normalize and train network
- Testing
- Accuracy

In data inputs, user can add new data, update and delete existing data. This is the training dataset of the system.

The system shows the normalized training dataset and it can be trained.

Before training, user must define the population size and training time. Also user can select the data item which he/she want to train. Data can be selected randomly (eg : 2, 6, 9, … , etc), by group or all data. User can construct the network architecture with variable of number of neurons in hidden layer.

During the training time, the system displays the output results of the global error and fitness value, weight values, after crossover and after mutation value for each training time. Finally, the system also displays the final output of the system.

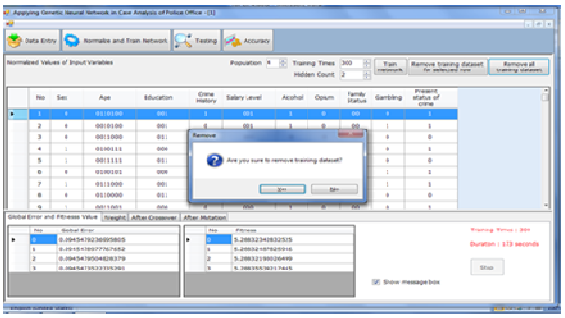

Fig. 3 Selected Training Dataset to Confirm

Training the network can be stopped by two conditions: user's defined loop count and predefined error values. When training is completed, the system displays the message and output result of

weight values, crossover values and mutation values.

In training dataset, user can remove the data as shown in Fig. 3. After neural network has been training with GA and BP, the user can test various values with BP and GA neural network. Before testing the data, there is no output results. The result is shown in Fig. 4.

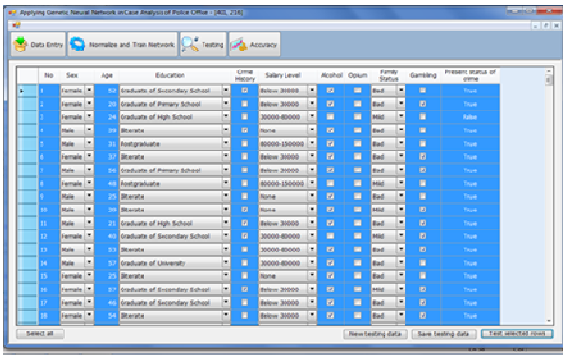

Fig. 4After Testing with BP and GA

In accuracy analysis, user can test the performance of the BP and GA algorithm by using accuracy analysis. When user clicks "Accuracy" button, the system displays the accuracy of the testing data.

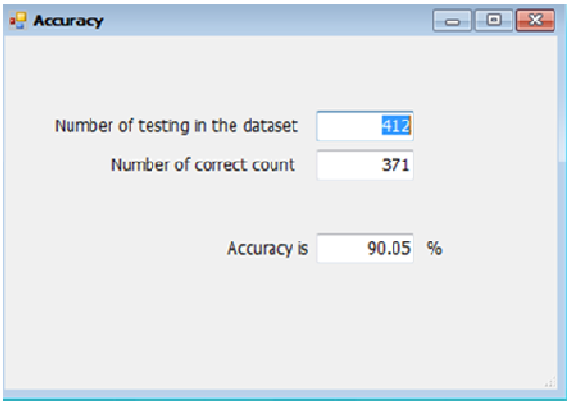

Fig. 5 Output of the Accuracy Result

VI. CONCLUSIONS

The knowledge learned by neural networks is hard to be understood by users because it is concealed in a large amount of connections. Neural networks achieve high accuracy in classification, prediction and many other applications. Neural network has been proven of their capabilities in many domains such as medical application. Genetic Algorithms are Heuristic Search Algorithms. They are basically intended to solve problems without Algorithm. Genetic Algorithms doesn't guarantee that this is the best solution but it guarantees that this is a good solution. Neural Networks and Genetic Algorithms are two highly popular areas of research, and integrating both techniques can often lead to highly successful learning systems.

Using genetic algorithm to optimize the BP network can effectively avoid the problem of local minimum. Therefore, the genetic neural network based on data-mining model has many advantages over other data mining models. By inducing and analysing these historical cases, people can get some experiences and learn some lessons that can help them solve cases and make decisions in the future. Therefore, the system can assist police departments to solve cases rapidly and make decisions efficiently.

VII. LIMITATION AND FURTHER EXTENSION

The system can only recognize numerical numbers. The best training algorithm still cannot be singled out for general neural network. Although BP algorithm has been widely used, it does not guarantee the global optimal solution. Although the genetic algorithm does not require normalization of the data, the backpropagation programs all normalized the data. Besides, the selection of some parameters in BP training also lacks of systematic guideline.

Genetic Neural Network has been also widely applied to medical treatment and genetic engineering and many other fields. In insurance companies, stock companies, banks and credit card companies, people apply genetic neural network to detect the deceptive actions of customers to reduce the commercial deceptions. Other applications can test the genetic algorithm, such as biology, engineering research, image processing and pattern

recognition, and physical science. Genetic algorithms provide a potential method to learn neural networks for such tasks and hence to simplify the development of such networks and to make them more robust.

REFERENCES

[1] TalibS.Hussain, "Methods of Combining Neural Network and Genetic Algorithms", Queen's University.
[2] Aye AyeNwe, "Printed Circuit Board Design by Using Genetic Algorithm", M.C.Tech. (Thesis), June, 2001.
[3] LIU Han-li, LI Lin, ZHU Hai-hong, "Genetic Neural Network Based Data Mining and Application in Case Analysis of Police Office", School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan, P.R.China.
[4] HtweHtwe Linn, "Travelling Sales Man Problem by Using Genetic Algorithm", M.C.Sc. (Thesis), September, 2007.
[5] JacekM.Zurada, "Introduction to Artificial Neural Systems", Info Access Distribution Pte Ltd, 1992.

