Comparison of Genetic and Hill Climbing Algorithms to Improve an Artificial Neural Networks Model for Water Consumption Prediction

Sajjad Rezaei - Department of Industrial Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.
Farbod Zorriassatine1 - Department of Industrial Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran.

Abstract
No unique method has been so far specified for determining the number of neurons in hidden layers of Multi-Layer Perceptron (MLP) neural networks used for prediction. The present research is intended to optimize the number of neurons using two meta-heuristic procedures namely genetic and hill climbing algorithms. The data used in the present research for prediction are consumption data of water subscribers in Fasa City of Fars Province (Iran) between the years 2010 to 2013. Ultimately, using the respective data set, the data of the subsequent year 2014 can be predicted. In the present research it was observed that the mean square errors of per data (MSEPd) for the abovementioned algorithms are less than 0.2, indicating a high performance in the neural networks’ prediction. Correlation coefficients using genetic and hill climbing algorithms were respectively equal to 0.891 and 0.759. Thus, GA was able to leave a better effect on optimization of neural network.

Key words: Water Consumption Prediction, Genetic Algorithm, Hill Climbing Algorithm, Artificial Neural Network, Multi-Layer Perceptron, Correlation Coefficient
Introduction

Prediction has always mattered to people playing a crucial role in their life. Accordingly, human beings have permanently searched for ways to predict the events and have deployed different methods to reach their goals.

Water is among the resources of the world which can be properly managed for consumption using prediction techniques because water and the competition to dominate its limited resources will be one of the most challenging issues of the third millennium. Population growth, industrial and social development, and climate change each have limited the available clean water resources. Water crisis has inflicted Iran sooner than most of the world’s other countries and will undoubtedly inflict irreparable damages on the country’s economy in case of lack of appropriate and coherent planning. Therefore, the heavy burden of timely, high-quality, and cost-effective supply of water requires accurate management and intelligent decision-making. One of the strategies in this respect is prediction of future levels of the subscribers’ water consumption. Based on this strategy, it will be possible to supervise and control not only water consumption but also water subscribers’ behaviors to take appropriate action accordingly where necessary.

There are different methods to forecast future values of subscribers’ consumption. The examples are regression techniques, Artificial Neural Networks (ANN), and so on. In the present research, Back-Propagation Multi-Layer Perceptron (MLP) ANNs were used. This kind of neural network was primarily proposed by Rosen Blatt in 1958. Some researchers resumed their research to make machines with the ability to solve problems such as pattern recognition based on this sort of neural network. Verbas (1974) invented back-propagation training and learning technique which was an MLP with more powerful training rules (Martin et al., 2014).

Najafi & Givi (2006) compared both ANN methods and pedotransfer functions to predict soil bulk density and realized finally that ANN was superior to PTF.

In another paper, it was shown that neural network yields an excellent result with correlation coefficient of over 80% for prediction.


Unal et al. (2012) by Back- propagation predicted effect of water pressure on lettuce product with a correlation coefficient of 80%. Also, properties of cereals were predicted by means of artificial neural network taking into account quality of cereals, moisture, cereal yield, protein and nitrogen contents of cereals, and so on. The prediction results were satisfactory (Goyal, 2013). In another research work, energy consumption and greenhouse gas emission in apple production of Iran was predicted using ANNs. The results indicated that the prediction had a correlation coefficient of 90% (Taghavifar, H., Mardani, A., 2015).

Furthermore, applications of ANNs include:


One of the important issues in neural network techniques is to determine the number of neurons in hidden layers, for which of course, no unique method has been reported. One of the methods is to use optimization methods such as genetic algorithm, HC, and so on. Sayyed Ali Moasher and Seyyed Mahmood Tábata-
baie (2013) predicted nitrate amount in underground water by means of optimized neural network based on GA. The respective prediction yielded a correlation coefficient of 83% based on optimized neural network.

In another paper entitled “utilization of genetic algorithm in optimization of ANN for earthquake prediction”, use of GA led to efficiency improvement and high accuracy of predictions (Qiwen Zhang and Cheng Wang, 2008).

In a research by, Ashkan Nabavi-Pelesarei, Reza Abdi, and Shahin Rafiee (2014) using multi-target GA and ANN managed to model and optimize energy consumption and emission of greenhouse gases resulting from almond production. The results showed a correlation coefficient of over 99%.

Another research which took advantage of GA and ANNs for optimization of soybean hydration resulted in a correlation coefficient of 93% with mean square error (MSE) of 5.9299 (Tushar Gulati, Mainak Chakrabarti, Muralidhar Duvuuri and Rintu Banerjee, 2010).

As implied before, in addition to GA, HC algorithm has been also used in the present research to optimize ANN.

In a paper entitled “convergence of genetic – hill climbing algorithm for image correlation”, Andrew D.J. Cross, Richard Myers, and Edwin R. Hancock (2000) specified that deployment of HC algorithm accelerated convergence to a large extent and was able to obtain good results with respect to correlation of images.

In another work (Brototi Mondala, Kousik Dasgupta, Paramartha Duttab, 2012) the optimal spot was determined to allocate servers or virtual (computer) servers for responding to large volume of computations.

A hybrid safe HC optimization algorithm was utilized for solving optimization problems in designing and production of industries (Yildiz, 2009). It was shown that the respective algorithm was particularly superior regarding quality.

In another research (Tuping Jiang, Gang Ren, and Xing Zhao, 2013) it was demonstrated not only forbidden search and hill climbing algorithms can acquire the best result but also have a high computational efficiency and faster convergence rate.

Arid countries like Iran need to be rescued from any probable threat concerning water resources during crisis because phenomena such as sinkholes have been formed in different parts of Iran including Fars Province due to reasons like water deficiency. Also, performing this research with the techniques depicted above, the future values of water consumption will be predicted aimed at helping Fars Province Water Organization to improve controlling of urban water resources in order to optimize water distribution and reform consumption behaviour. Other objectives might also be sought for such as: enhancement of the ability to utilize available information resources for better planning intended to correct consumption patterns, and, improvement of organizational efficiency in implementation of the tasks related to control and supervision over consumption. The assumptions of the research include:

1. The input training data are certain. Therefore, there is no probable, random, and fuzzy state or any other state that might violate certainty assumption.
2. The data are independent.

**Materials and Methods**

The data set used in the present research consists of the data of subscribers’ consumption in the period from 2010 to 2013 belonging to Fasa County of Fars Province – Iran. Since Fars Province in Iran suffers greatly from drought, the data of the aforementioned city were used in this research. The number of subscribers was 28,938 households whose consumption data were distributed during 4 years in 24 columns. The data set used benefited from data certainty, being free from noises, and data independence. Fasa County is located in Fars Province 145 kilometers south of Shiraz.
ANN programming in MATLAB software, was used as a suitable tool for analyzing functions and recognizing patterns and processing any operational functions in order to predict future values of subscribers’ water consumption during the year 2014 given the data of the three previous years.

The intent of the neural networks is to construct patterns that act like human's brain. One of the function of neural networks is to create an output pattern based on the input pattern given to the network. Neural networks acquire and process inside the data using their processing elements (or artificial neurons), and ultimately, generate an output from it.

One of the significant points in this respect is to determine the number of neurons within the hidden layer. Among the methods to solve this problem is utilization of meta-heuristic algorithms. In the present research, genetic and HC meta-heuristic algorithms were used.

GA was founded based on Darwin’s theory of evolution. The principal idea is based on the concept that any datum can be acquired from combination of two parent data such that it would acquire some features from the mother and the others from the father. The strength of the algorithm is the fact that any parent who is superior to others has a greater chance to select and reproduce offspring, and accordingly, the next generations tend to improve with resumption of reproduction process. This algorithm does not need error function gradient.

HC is one of the extremely simple local search methods which can provide the researchers with proper results taking into account small size of problem space. Also, this algorithm is one of those which have been applied for solving optimization problems during the recent years thanks to its high speed, simplicity, and effectiveness (Brototi Mondala et al., 2012).

Feed-forward (Back-propagation) neural networks are the networks in which the relationship is one-way and data is transferred from one neuron only to the next neuron whereas the relationship is two-way in feed-backward networks. In the present research, the feed-forward MLP architecture was used, which divides network processors into several input, hidden, and output layers.

BP algorithm, which is based on gradient descent method, is used for solving optimization problems. In MLP, the objective is to reduce cumulative error (Ec) of the network. Ec is the sum of E(k) square errors, such that:

$$E(k) = \sum_{i=1}^{q} (t_i(k) - O_i(k))^2$$

Where, t(k) is the kth ideal output, and O(k)

![Figure 1. Structure of genetic and HC algorithms with neural network](image-url)
is the kth actual output of the network. The algorithm has been designed such that it is supposed to update network weights aimed at reducing gradient of network’s cumulative error. In general, this algorithm works as follows. First, the training pattern is introduced to the network. The error between ideal output and actual network output is then computed and propagated as a signal in reverse direction (back propagation). And by definition, the connection weights of the network are updated. The process resumes until the weights are obtained and the network is trained. Figure (2) illustrates this error correction mechanism.  

There are different types of stimulation functions for computational neurons. Here, sigmoid function was used as computational function (transfer function) of all neurons. 

\[ E_c = \sum_{k=1}^{n} E(k) = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{q} [t_i(k) - O_i(k)]^2 \]

Where; n is number of training patterns, I is index of output layer neuron with total number equal to “q”. The weights are updated such that \( E_c \) and/or \( E(k) \) is minimized.  

Levenberg_Marguardt training function was used for prediction because this function can bring about favorable results in computations and mathematics during fitting process in order to minimize error squares (Levenberg, 1944).  

**Evaluation Criteria**  
One of the critical issues here is the evaluation of neural network performance. It must be checked whether neural network has been able to behave accurately in prediction or not. Two factors i.e. MSEPDP and R were used respectively representing mean square error for each data and correlation coefficient (A.K.Suykens, Johan. Vandewalle, Joos, P. L., 1998). For this purpose, mean square error index shall be expressed for per data. The respective index is simplified as below:  

\[ \text{MSE} = \sum_{i=1}^{n} \left( \frac{\hat{y}_i - y_i}{y_i} \right)^2 \quad \text{MSEPDP} = \sum_{i=1}^{n} \left( \frac{\hat{y}_i - y_i}{y_i} \right)^2 - n \]

Where \( y_i \) and \( \hat{y}_i \) respectively are actual and estimated values and, n represents number of observations.  

Now, if the derived equation is divided by number of data MSE value will be obtained in terms of data:  

\[ \text{MSEPDP} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\hat{y}_i}{y_i} \right)^2 - 1 \]

And, validation of neural network based on MSEPDP is as follows.(table1)  

**Correlation Coefficient**  
Correlation coefficient is another method for  

\[ r = \frac{n \sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{n \sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2}} \]

**Figure 2. Overall design of MLP consisting of error correction and weight reversal**
assessing performance of neural network. This coefficient reflects correlation level between model prediction results and actual data which is calculated as below. Obviously, the closer value to unity indicates further proximity of predicted values to actual values.

\[ R = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)(y_i - \hat{y}_i)}{\sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \sum (y_i - \hat{y}_i)^2}} \]

In equation above, \( y_i \) is the actual value and \( \hat{y}_i \) is the estimated value. The closer values of \( R \) to unity indeed signify more acceptable result.

**Results**

As mentioned in the previous sections, artificial neural network technique of MLP type was used to predict consumption values of subscribers in Fasa County – Fars Province. Also, genetic and hill climbing algorithms encoded in MATLAB programming were utilized in order to determine the number of neurons in the hidden layer.

**Genetic Algorithm and Neural Network**

The first step in this algorithm is to convert phenotype to genotype. It means that genetic average can be extracted from the actual data of the problem. Since the objective of the problem is finding the optimal value of number of neurons, the data can be simply regarded as a number. Due to prevention from overfitting, there is no need to have a high number of neurons. Therefore, number of neurons in the hidden layer is considered to be more than twice the number of neurons in the input layer, i.e. 32, which can be easily illustrated with 5 bits. Thus, every chromosome is a 5-bit data representing an integer between 0 and 31.

**Selection**

In each stage, some of data in the previous stage were selected as parent based on their fitness. Roulette wheel technique was applied in order to have a fair selection. As such, the algorithm started working with 20 chromosomes which had been generated randomly. At each stage, 18 of chromosomes were chosen as parents based on fitness and using Roulette wheel. The remaining two were the best of the previous generation which represented the most suitable in terms of eligibility. They have qualified for this stage directly without involvement in the competition.

**Crossover**

Crossover operator was used for generat-
ing new children or chromosomes. Uniform crossover was used i.e. each gene was generated from one child chromosome with uniform probability from one of the parents, and, the same gene was given to the second child in another parent. In this way, two new children were reproduced from each of the two parents.

**Mutation**

Genetic mutation is one of the ways to avoid local optimum in genetic algorithm. During this stage, one gene was chosen and toggled from a chromosome with low probability i.e. converted to zero if it was zero and vice versa. At the end of this stage, the new generation was reproduced and made ready to be sent to the subsequent iteration. Probability values of 95% and 5% were respectively used for performing crossover and mutation operations. Therefore, taking into account all abovementioned points and developing the codes related to GA, the number of neurons obtained based on GA due to optimization of neural network was equal to N = 16.

The diagrams acquired from GA and neural network training are shown below:

A: Training data; B: Test data; C: All data

As observed from Figure 4, the correlation coefficient of the data selected for training was 0.92 and the correlation coefficient of the test data was around 0.89.

**Hill Climbing Algorithm and Artificial Network**

This method is among the local search techniques where a number of individuals are considered similar to GA technique. However, unlike GA where one child is generated from two parents, one child is assumed to be generated from one child with slightest variations. In other words, it only considers the children around and moves toward better points. Thus, a point is randomly chosen in the problem space, and then, using increment and decrement functions, the algorithm moves in the direction with the greatest progress in terms of the fitness value. Hence, it approaches the

![Figure 4. Correlation coefficient of the values predicted based on the available consumption values taking into account training data; test based on GA/ANN](image)

![Figure 5. Correlation coefficient of the values predicted based on the available consumption values taking into account training data; test based on Hill Climbing/ANN](image)
best solutions by moving to both sides in the problem space. Also, random restart was utilized to prevent it from being stuck at local optimal points.

As observed from Figure 5, the correlation coefficient of the data selected for training was 0.99 and the correlation coefficient of the test data was around 0.75. Thus, in the table 2 the results acquired from data input to optimized neural network based on genetic and HC algorithms are shown.

Also, the following figures illustrate the general diagram and summation of consumptions during the previous and predicted years.

At the end, some instances of the trainings performed in MATLAB software using GA/ANN are illustrated based on the available data of a subscriber for better prediction of data (aimed to assess performance of neural network).

**Conclusions**

Undoubtedly, gaining knowledge about future consumption levels in future periods is one of the measures that can be taken for prevention of crisis in Iran’s water resources. Since, this crucial information can serve as the basis of planning for establishing the required facilities and equipment including the required water volumes during specific periods depending on the regional climates, the required installations for transmission of water resources to consumers, prevention of water loss, and even, water management during times of crisis such as drought that threaten Iran. Ultimately, this approach will be effective in control of costs and gaining national revenues via mitigating the water consumption. As stated earlier, this prediction was done using an optimized neural network tool through genetic and HC algorithms. According to Table (2), it is observed that the number of optimal neurons for the hidden layer of network was equal to 13 and 16 using genetic and HC algorithms, respectively. Their correlation coefficients respectively equaled 0.891 and 0.759 after applying them in the neural network architecture. As values of correlation coefficient closer to unity signify further matching of predicted data with real data, GA therefore yielded more favorable results compared to HC algorithm. It is also noteworthy that MSEPD values of both algorithms were less than 0.2 according to Tables

<table>
<thead>
<tr>
<th>Name of algorithm</th>
<th>No. of iteration</th>
<th>Training Function</th>
<th>No of neuron</th>
<th>No. of hidden layer</th>
<th>R</th>
<th>MSEPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA/ANN</td>
<td>1000</td>
<td>Levenberg-marquardt</td>
<td>16</td>
<td>1</td>
<td>0.891</td>
<td>0.0031</td>
</tr>
<tr>
<td>HC/ANN</td>
<td>1000</td>
<td>Levenberg-marquardt</td>
<td>13</td>
<td>1</td>
<td>0.759</td>
<td>0.0049</td>
</tr>
</tbody>
</table>

Table 2: The results acquired from data input to optimized neural network based on genetic and HC algorithms

![Figure 6: Sum of consumptions during 2010 until 2013 and the predicted year (2014)](image-url)
(1) and (2) meaning that the performance of neural network was quite successful in prediction. Figure 6 illustrates the total consumptions in three previous years (2010 to 2013) and the prediction year 2014. This diagram indicates consumption trend in the city under study was ascending whereas precipitation amount has decreased year by year and underground water level has declined as well. It is an alarm for Fars Province’s Water Company for more appropriate consumption control and management. Another result and the analysis based on Figure 6 is the fact that no increase was observed in water production despite the increasing consumption trend. Perhaps, the only way for responding to the demands of future years is to change the subscribers’ consumption model based on hours of each day. Water Organization can also monitor the subscriptions regarding water theft and leakage etc. based on
consumption pattern of each subscriber as a result of the present research. Furthermore, instances of neural network trainings can be observed in Figures 7 through 10 based on the available data and estimation of the same data using neural network.

The restrictions of carrying out this research were the absence of water temperature and pressure information in the database system of the respective system. More accurate prediction can be delivered based on the respective data, if available. Also for future research works, the dataset can be analyzed taking into account fuzzy approach and utilization of ANFIS neural network for prediction of future consumption values and comparison with the results of the present research, and likewise, deployment of data search tools such as clustering and classification for prediction.

**Acknowledgements**

The author should appreciate Water Organization (ABFA) of Fars Province – Iran for their collaboration in execution of this research.
References