Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env.Pharmacol. Life Sci., Vol 4 [Spl issue 1] 2015: 212-218

Online ISSN 2277-1808
Journal's URL:http://www.bepls.com
CODEN: BEPLAD
Global Impact Factor 0.533
Universal Impact Factor 0.9804

FULL LENGTH ARTICLE **OPEN ACCESS**

Comparison and Evaluation of Artificial Neural Network (ANN) Training Algorithms in Predicting Soil Type Classification

Hamidreza Hassannejad Mohammad Sirus Pakbaz, Rasool Mehdizadeh
Department of Civil Engineering, Islamic Azad University of Najafabad Branch, Iran

ABSTRACT

The present study uses different ANN training algorithms to predict soil type classification and evaluates the output of their training. Studies were done on the soil of Shahrekord (Iran), using a database consists of 120 soil samples. The used data includes the results of standard impact and penetration tests, classification and Atterberg limits. Because of diverse training algorithms in neural networks, the selection of the optimal training function can increase the accuracy of network predictions. For this purpose, several training functions and algorithms were used in the ANN modeling. To evaluate the performance of proposed models, relations of coefficient of residual mass (CRM), coefficient of determination (COD) and root mean square error (RMSE) were used. The comparison and evaluation of training with various algorithms show that the Levenberg-Marquardt training function with very high accuracy in network training is the optimal training function used in the prediction of geotechnical parameters of soil, including soil type.

Keywords: *Artificial Neural Network, Predicting Soil Type Classification, Levenberg-Marquardt algorithm, Artificial Neural Network Training Algorithms*

INTRODUCTION

The initial steps in the design and implementation of civil projects include determining geotechnical characteristics and parameters of soil in the project site. Geotechnical engineering includes studying the properties of land layers for construction purposes. In this context, underground identification or local search is an essential prerequisite by which geological position, geotechnical parameters and other information involved in the construction and operation of geotechnical engineering are obtained.

To determine the geological structure of an area, local explorations and geologic boreholes are needed. Given the importance and magnitude of a project, some boreholes with different depths are dug. Then various geotechnical tests including aggregation, hydrometer, plate loading, standard impact and penetration tests are done and then soil classification, Atterberg limits, type of land, cohesion, internal friction angle of the soil and the reaction coefficient of the bed are determined. But the situation of geotechnical parameters in the area between speculations is not specified clearly. On the other hand, some geological parameters like the shaping type and degree of weathering are created in a complex process and are described qualitatively. Therefore, applying them with other quantitative parameters in numerical modeling to model their spatial distribution needs many simplifications and assumptions. Moreover, multiple speculations are not economically justified. Thus, to predict geotechnical parameters, one can use the information of adjacent boreholes and train them by the ANN. The ANN is focused only on the patterns in data which result in a high speed of success in estimating soil properties.

2. Database

In the present study, a database consists of 120 soil samples was used. The used data is related to a part of Shahrekord, Chahar Mahal and Bakhtiari province, obtained by the Technical and Soil Mechanics Laboratory of the Roads and Urban Development on the province. Samples were named by the Unified Soil Classification System. Various geotechnical tests were performed on samples including aggregation, hydrometer, plate loading, and standard impact and penetration tests (SPT), and then project soil classification, Atterberg limits, type of land, cohesion, internal friction angle and the reaction coefficient of the bed were determined. Seven input parameters were used in modeling: sample moisture, liquid limit, plasticity range, SPT, sample coordinates (longitude, latitude and altitude). In order to use the sample coordinates, a hypothetical center pint was considered in the project. The middle part of each sample was determined as its representative, and its coordinates was given in modeling. The target parameter in this artificial neural network is soil type classification. To use this classification in modeling, a number is assigned to each class, according to Table 1.

Table 1. Soil Type Classification's Codes

Soil Type Classification	SM	SC	GC	CL	GC-GM	GP	GP-GC	GM
Code	1	2	3	4	5	6	7	8

Table 2 shows the scope of changes in some parameters of modeling.

Table2.

Parameters	MC (%)	LL (%)	PL (%)	SPT
Minimum	6.8	10	0	0
Maximum	25	69	33	100
Mean	12.09	31.23	12.53	65.65
Standard Deviation	4.99	9.83	5.95	19.69

3. Artificial neural network

The concept of artificial neural networks was first described by Segal (1911). The ANN is one of those dynamic systems that transmit knowledge or rule within data to the network by processing the experimental data. The development of artificial neural networks started 50 years ago. Their driver was the understanding of the structure and function of the brain and simulating it to use the high power of the brain in different applications. Moreover, in cases where obtaining numerical relationships to relate independent variables to dependent parameters is difficult, the ANN performs significantly. In fact, neural networks are trained by a limited series of actual data. If effective parameters on the phenomenon under study are properly selected and given to the network, it can be expected to receive logical solutions from the network. Thus, a neural network does not need the regression analysis of dependent variables. Of course, this analysis requires basic data which is difficult to achieve in many cases. Thus a neural network can offer a better model according to available data sources.

Today the use of this network has increased in geotechnical engineering in areas such as estimating of ultimate capacity of shallow foundations, forecasting cement injection parameters in the barrier waterproof membrane, evaluating soil liquefaction, and anticipating pile subsidence.

In the present study, we used propagation networks. These networks are multi-layer with a nonlinear transfer function and Widrow-Hoff learning rule. The input vector and target are used for training this type of network to approximate a function, find a relationship between input and output and classify inputs. Having a bias, a sigmoid layer and a linear output layer, a propagation network can approximate any function with a limited number of discontinuity points. The term propagation refers to the performance in gradient calculation in multi-layer nonlinear network. A multi-layer neural network that is trained with the error propagation method is also called multi-layer perceptron network. In this method, there are an input layer, an output layers and a hidden layer. The weight of the relationship between the hidden layer and the input layer determines the hidden layer performance. Moreover, the activities between hidden layers and the weight of the relationship between them and the output layer determine the output layer performance. Each layer consists of some neurons (nodes). These neurons are connected to each other by their weight which indicates their power. A simple ANN architecture is shown in Figure 1.

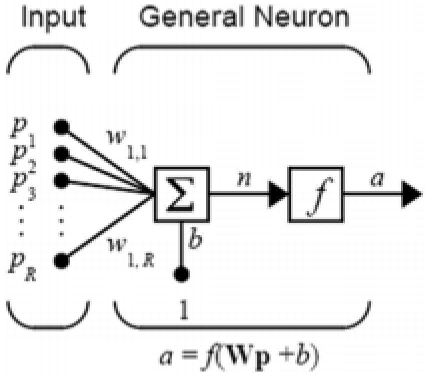

Figure 1 Simple ANN architecture

The neural network used in this study has an input layer, an output layer and a hidden layer. The input layer includes 7 variables (MC, LL PL, SPT, X, Y, Z) and the output layer has only one parameter (soil type classification). The hidden layer has 15 neurons. The number of nodes in the hidden layer was determined by trial and error.

Various training functions are used in the ANN. In order to predict with minimum error, an optimal training function must be used in the ANN. The study used these algorithms: batch gradient descent training, batch gradient descent training with momentum, variable learning speed, back-propagation, conjugated gradient, quasi-Newton and Levenberg-Marquardt.

The RMSE, CRM and COD were used to evaluate the performance of the proposed model. The CRM represents the difference between actual and estimated values. The RMSE is very useful to calculate the performance of prediction models. It has been used by various researchers.

Mathematical parameters are listed below.

$$CRM = 1 - \frac{\sum_{i=1}^{n}(P_i)}{\sum_{i=1}^{n}(M_i)}$$ Relations of Coefficient of Residual Mass

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - M_i)^2}{n}} \times 100$$ Root Mean Square Error (RMSE)

$$COD = 1 - \frac{\sum_{i=1}^{n}((M_i - P_i)}{\sum_{i=1}^{n}(M_i - \bar{M})}$$ Coefficient Of Determination

where Mi, Pi, $\bar{M}$ and n are actual values, estimated values, the mean of actual values and the sample size, respectively. The lower the RMSE, the model accuracy will be higher. The positive values of CRM indicate that the model estimates the soil type less than its actual value and vice versa. If COD equals one (100%), i.e., when we use independent variables, no error occurs which is the best possible case.

4. Introduction of training algorithm

The algorithms used in the network training are described below.

4.1 Batch gradient descent training

In the batch method, weights and biases are updated after applying all members of the training set. The gradients calculated for each input are summed up until the weights and biases are updated. In this method, weights and biases are updated in the reverse direction of the gradient of performance function. Its training function is Traingd.

4.2 Batch gradient descent training with momentum

Momentum allows the network to react to error level changes in addition to gradient changes. It also makes the algorithm trajectory softer by ignoring minor errors. Its training function is Traingdm.

4.3 Variable learning rate

With the increasing pace of learning in flat surfaces and its reduction in high gradient points, one can improve the convergence speed. There are various types of learning algorithms with variable learning speed. Jacobs proposed a learning rule called delta-bar-delta, where each network parameter (weights and biases) has its own learning speed. The SuperSAB algorithm uses a rule similar to delta-bar-delta but has more complex rules for adjusting the learning speed. The used training functions are Traingda and Traingdx. The Traingdx function combines two learning speed methods of adaptive learning speed and momentum. This method is similar to Traingda except that it has an additional parameter called momentum factor.

4.4 Resilient Backpropagation training function

This algorithm removes the harmful effects of the size of partial derivatives. Only the derivative sign is used for weight updates, and the derivative size has no effect on the weight updates. This algorithm has a much higher performance than the standard gradient descent algorithm. In addition, this method requires less memory. The back-propagation training takes place using the training function Trainrp.

4.5 Conjugated gradient algorithm

This is a balanced algorithm. It does not require the calculation of second derivatives and ensures convergence to the minimum quadratic function. In most conjugated algorithms, learning speed is used to determine the step size in updates. In most conjugated algorithms, the size of each step is set for each iteration. For this purpose, a search operation is done among all conjugated gradients to select the most appropriate one to minimize the performance function. The conjugated gradient comes in different types, including:

The Fletcher-Reeves conjugated gradient algorithm with the Traincgf training function, the Pollack-Ribier conjugated gradient algorithm with the Traincgp training function, and the Paul-Bill conjugated gradient algorithm with the training Traincgb function.

Scaled conjugated gradient algorithm

This algorithm was designed to stay away from the time-consuming linear search. It is very complex and is based on the combination of the two methods of conjugated gradient and Levenberg-Marquardt. Its training algorithm needs more iterations to converge compared to the rest of conjugated gradient algorithms, but the amount of computation per iteration is markedly reduced because linear search is not performed in this method. Its training function is Trainscg.

Quasi-Newton algorithms

Newton methods usually have a better and more rapid convergence than conjugated gradient algorithms. But they are very complex and computationally expensive. Two quasi-Newton algorithm are:

The BFGS algorithm which is located in the Trainbfg training function and the one-step secant algorithm which is located in the Trainoss training function. The latter requires less space and computation than the former.

Levenberg-Marquardt algorithms

This method usually have more rapid convergence than other algorithms. But this algorithms need to maintain large matrixes in memory. This requires a lot of space.

Its training function is Trainlm.

Predicting soil type

In all the methods used in network training, training data is divided into three categories: 80% of data is used for training, 10% for validation and the remaining 10% for network testing or verification. Since data is randomly selected in the network training, the training process is continued until achieving the optimal neural network. In this study, among 120 data, 96 data (80%) was used for training, 12 data (10%) for validation and 12 data (10%) for verification. MATLAB 2014 was used for network training. After training with the proposed algorithms, all the input data was trained to evaluate the prediction accuracy of neural network. Table (3) shows the evaluation of the prediction accuracy of training algorithms with the RMSE, COD and CRM indicators.

Table 3. Evaluation of the prediction accuracy of training algorithms

Algorithms	RMSE	CRM	COD
Batch Gradient Descent	0.5007	0.0051	87.88%
Batch Gradient Descent with Momentum	0.8433	0.0028	68.65%
Variable Learning Rate	0.6449	-0.0156	81.74%
Resilient Backpropagation	0.3141	0.0055	95.69%
Fletcher-Reeves Conjugated gradient algorithm	0.8689	0.0137	66.68%
Pollack-Ribier Conjugated gradient algorithm	0.9011	-0.0109	64.59%
Paul-Bill conjugated gradient algorithm	0.8734	-0.0066	66.21%
Scaled conjugated gradient algorithm	0.6242	0.0048	82.7%
BFGS algorithm	0.4785	-0.0095	89.87%
one-step secant algorithm	0.7260	0.0048	76.6%
Levenberg-Marquardt algorithms	0.0521	-0.0013	99.88%

Table (3) shows that the Levenberg-Marquardt training algorithm with an excellent performance could best predict soil type classification. Figures (2), (3), (4) and (5) show the high accuracy of Levenberg-Marquardt algorithm.

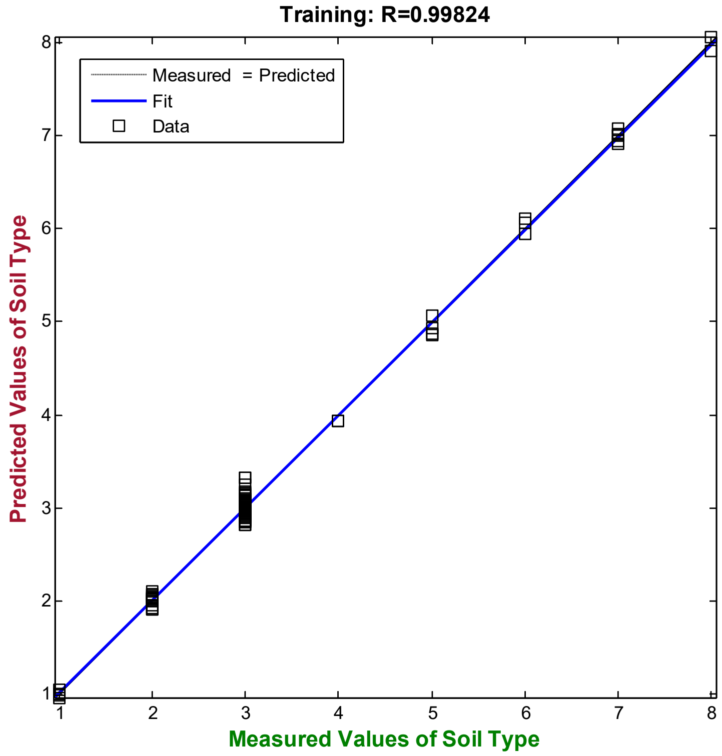

Figure 2 Comparison between the predicted values of soil type and the Measured values of Training data

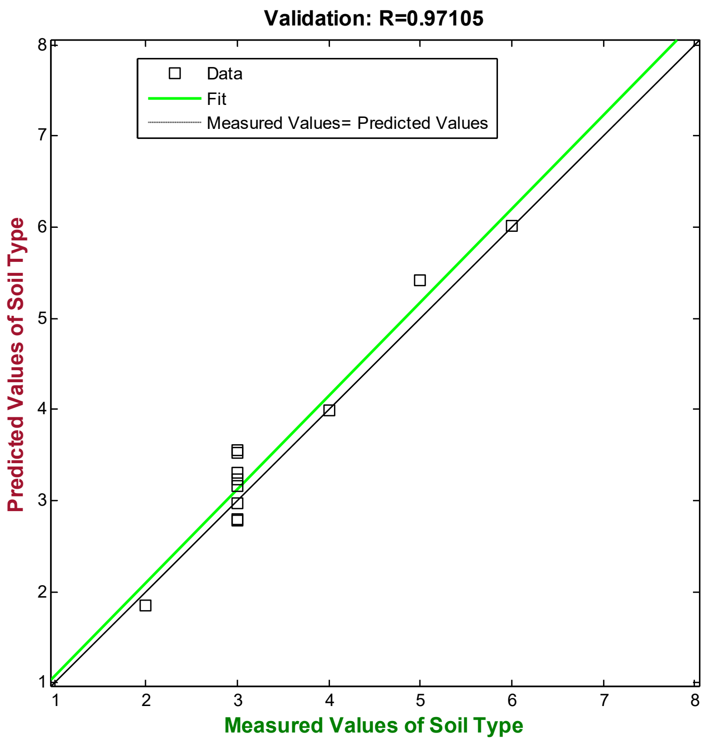

Figure 3 Comparison between the predicted values of soil type and the Measured values of Validation data

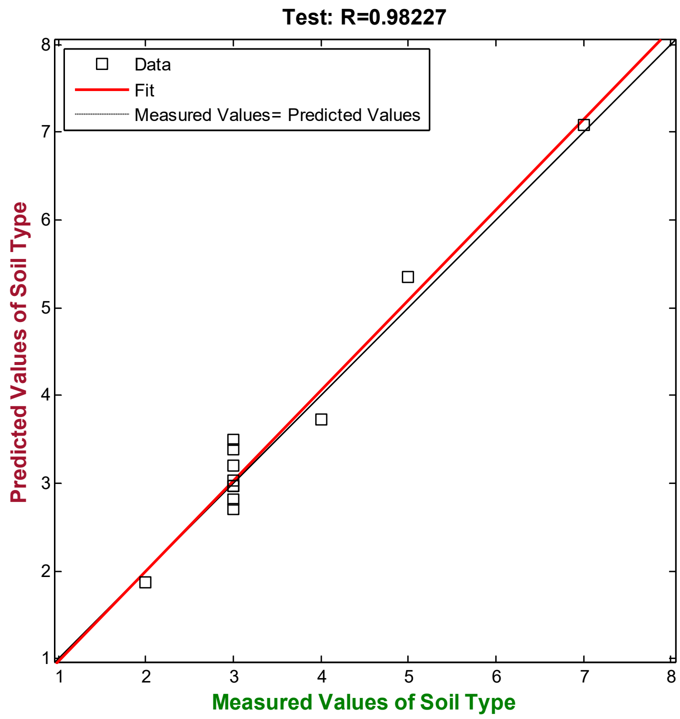

Figure 4 Comparison between the predicted values of soil type and the Measured values of Testing data

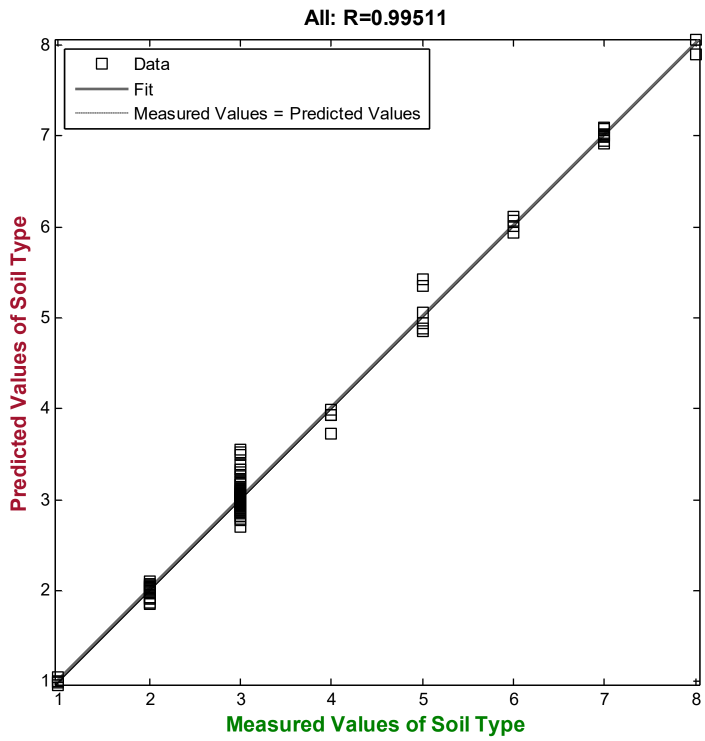

Figure 5 Comparison between the predicted values of soil type and the measured values of All data

CONCLUSIONS

The initial steps in the design and implementation of civil projects include determining geotechnical characteristics and parameters of soil in the project site. To characterize the geology of a region, we need local explorations and several boreholes. Given the importance and magnitude of the project, some boreholes must be dug with various depths. Then various geotechnical tests are done on them. Multiple boreholes are not economically justified. Thus to predict geotechnical parameters, one can use information of adjacent boreholes and train them by the ANN.

In this study, a model was proposed to predict soil type classification, and the best training neural network algorithm was known. A database consists of 120 soil samples was used. The database is related to the boreholes of the Parsian Hospital construction project in Shahrekord (Iran). Classification, sample moisture, liquid limit, plasticity range, number of impacts of standard penetration test, and sample coordinates are 8 parameters used in modeling. For training the network, the back-propagation or multi-layer perceptron was used. The neural network was trained by these algorithms: batch gradient descent training, batch gradient descent training with momentum, variable learning speed, back-propagation, Fletcher-Reeves conjugated gradient, Pollack-Ribier conjugated gradient, Paul-Bill conjugated gradient, scaled conjugated gradient, quasi-Newton BFGS, quasi-Newton one-step secant and Levenberg-Marquardt. To evaluate the performance of the proposed model, the RMSE, CRM and COD were used. By comparing measured values and predicted values, it can be concluded that the Levenberg-Marquardt algorithm with COD= 99.88% and RMSE =0.0521 is the best training algorithm to predict soil type classification. Then the back-propagation algorithm and quasi-Newton BFGS can train the network with a good accuracy.

It should be noted that the data used in this study was limited. The accuracy of neural network training can much increase in case of increasing the number of data.

REFERENCES

1. Zhou, Y.X., WU, X.P., 1994, Use of Neural Network in the Interpretation of Site Investigation Data, Computers and Geotechnics, Vol. 16, No.2, pp. 105-122
2. Goh, A.T.C.,1995, Modeling soil correlations using neural networks. Technical note, Computing in Civil Engineering, ASCE 9 (4), 275–278
3. Juang, C.H., Jiang, T., Christopher, R.A., 2001, Three-dimensional site characterisation: neural network approach, Geotechnique 51 (9), 799–809.
4. Kurup, P.U., Griffin, E.P., 2006, Prediction of soil composition from CPT data usinggeneral regression neural network. Journal of Computing in Civil Engineering, ASCE 20 (4), 281–289
5. Doe, M. C. and C.Thirumalayah. 2000, Real time forecasting using neural networks. Artificial Neural Networks in Hydrology, edited by R.S. Govindarajue and A. Ramachandra Rao, Chapter 3
6. Ghafooripur, A., Asgari, M., Using neural networks to predict the soil profile with a nearby borehole data, The first international conference on earthquake, Technical and Engineering Faculty of Qom, Iran, 1384, 400-415
7. Menhaj, M., Principles of Neural Networks (Computational Intelligence), Professor Hesabi publication Center, Iran, 1377, 17-49
8. Tizpa, P., Jamshidi Chenari, R., Karimpour Fard, M., Machado, S., ANN prediction of some geotechnical properties of soil from their index parameters, 2014, Saudi Society for Geosciences 2014
9. Alvarez Grima, M.& Babuska, R., 1999, Fuzzy model for the prediction of unconfined compressive strength of rock samples. Int. J. Rock Mech. Min, Sci 36, pp. 339-349
10. Gokceoglu, C., 2002, A fuzzy triangular chart to predict the uniaxial compressive strength of Ankara agglomerates from their petrographic Composition, Engineering Geology 66, pp. 39–51
11. Finol, J., Guo, Y. K.& Jing, X. D., 2001, Arule based fuzzy model for the prediction of petrophysical rock parameters, Journal of Petroleum Science and Engineering 29, pp. 97–113.

