19

Applicability of GMDH-Based Abductive Network for Predicting Pile Bearing Capacity

Isah A. Lawal[1] and Tijjani A. Auta[2]
[1]*University of Genoa*
[2]*ATBU, Bauchi*
[1]*Italy*
[2]*Nigeria*

1. Introduction

A pile is a type of foundation commonly used in civil construction. They are made using reinforced concrete and pre tensioned concrete to provide a firmer base where the earth around a structure is not strong enough to support a conventional foundation [Pile, 2011]. Accurate prediction of the ultimate bearing capacity of a structural foundation is very important in civil and construction engineering. Conventional method of estimating the pile bearing capacity has been through pile load test and other in situ test such as standard penetration test and cone penetration test [Bustamante & Gianeselli, 1982]. Though these tests may give useful information about ground conditions, however the soil strength parameters which can be inferred are approximate.

In recent time, advances in geotechnical and soil engineering research have presented more factors that can affect pile ultimate bearing capacity. However, due to nonlinearity of these factors, the use of statistical model analysis and design has proved difficult and impractical [Lee & Lee, 1996]. So there is need to provide civil and structural engineers with intelligent assistance in the decision making process. Soft computing and intelligent data analysis techniques offer a new approach to handle these data overload. They automatically discover patterns in data to provide support for the decision-making process. Tools used for performing such functions include: Artificial Neural Networks (ANN) [Abu- keifa, 1998; Chow et al, 1995; Teh et al, 1997], Support Vector Machine (SVM) [Samui, 2011],Genetic programming [Adarsh, in-press] and Gaussian process regression [Pal & Deswal, 2008, 2010]. The results from using these tools suggest improved performances for various datasets. However, neural networks like other tool suffer from a number of limitations, e.g. long training times, difficulties in determining optimum network topology, and the black box nature with poor explanation facilities which do not appeal to structural engineers [Shahin et al, 2001]. This research work proposes the use of self-organising Group Method of Data Handling (GMDH) [Mehra, 1977] based abductive networks machine learning techniques that has proved effective in a number of similar applications for performing modelling of pile bearing capacity.

Recently, abductive networks have emerged as a powerful tool in pattern recognition, decision support [El-Sayed & Abdel-Aal, 2008], classification and forecasting in many areas [Abdel-Aal, 2005, 2004]. Inspired by promising results obtained in other fields, we explore the use of this approach for the prediction of pile bearing capacity.

2. Related work

In recent years some researchers have developed computational intelligence models for the accurate prediction of pile bearing capacity. In [Lee and Lee, 1996], the authors used back-propagation neural networks to predict the ultimate bearing capacity of piles. A maximum error of prediction not exceeding 20% was obtained with the neural network model developed by using data set generated from calibration chamber. Also, in [Pal,2008], the author investigates the potential of support vector machines based regression approach to model the static pile capacity from dynamic stress-wave data set. The experiments shows excellent correlation coefficient between the predicted and measured values of the static pile capacity investigated. Similarly, in [Samui, 2011], the author studied the potential of Support Vector Machine (SVM) in prediction of bearing capacity of pile from pile load data set. In the study the author introduces ε-insensitive loss function and the sensitivity analysis of the model developed shows that the penetration depth ratio has much effect on the bearing capacity of the pile.

However, in [Pal, 2010], the author took a different approach and investigated the potential of a Gaussian process (GP) regression techniques to predict the load-bearing capacity of piles. The results from the study indicated improved performance by GP regression in comparison to SVM and empirical relations. However, the author reported that despite the encouraging performance of the GP regression approach with the datasets used, it will be difficult to conclude if the method can be used as a sole alternative to the design methods proposed in the literature. The reason been that soft computing based modelling techniques are data-dependent. Their results may change depending on the dataset, the scale at which the experiments are conducted or the number of data available for training.

The potential for GMDH-based abductive network in pile bearing capacity prediction has not been explored before in the literature. However, compared to neural networks and other learning tools, the method offers the advantages of faster model development requiring little user intervention, faster convergence during model synthesis without the problems of getting stuck in local minima, automatic selection of relevant input variables, and automatic configuration of the model structure.

3. GMDH and AIM abductive networks

Abductory Inductive Mechanism (AIM) is a powerful supervised inductive learning tool for automatically synthesizing network models from a database of input and output values [AbTech, 1990]. The model emerging from the AIM synthesis process is a robust and compact transformation implemented as a layered abductive network of feed-forward functional elements as shown in Figure 1. An abductive network model numerical input output relationships through abductive reasoning. As a result, the abductive network can be used effectively as a predictor for estimating the outputs of complex systems [Lee, 1999], as a classifier for handling difficult pattern recognition problems [Lawal et al, 2010] or as a system identifier for determining which inputs are important to the modelling system[Agarwal. 1999]. With the model represented as a hierarchy of polynomial expressions, resulting analytical model relationships can provide insight into the modelled phenomena, highlight contributions of various inputs, and allow comparison with previously used empirical or statistical models.

3.1 Abductive machine learning

The abductive machine learning approach is based on the self-organizing group method of data handling (GMDH) [Fallow, 1984]. The GMDH approach is a proven concept for iterated polynomial regression that can generate polynomial models in effective predictors. The iterative process involves using initially simple regression relationships to derive more accurate representations in the next iteration in an evolutionary manner.

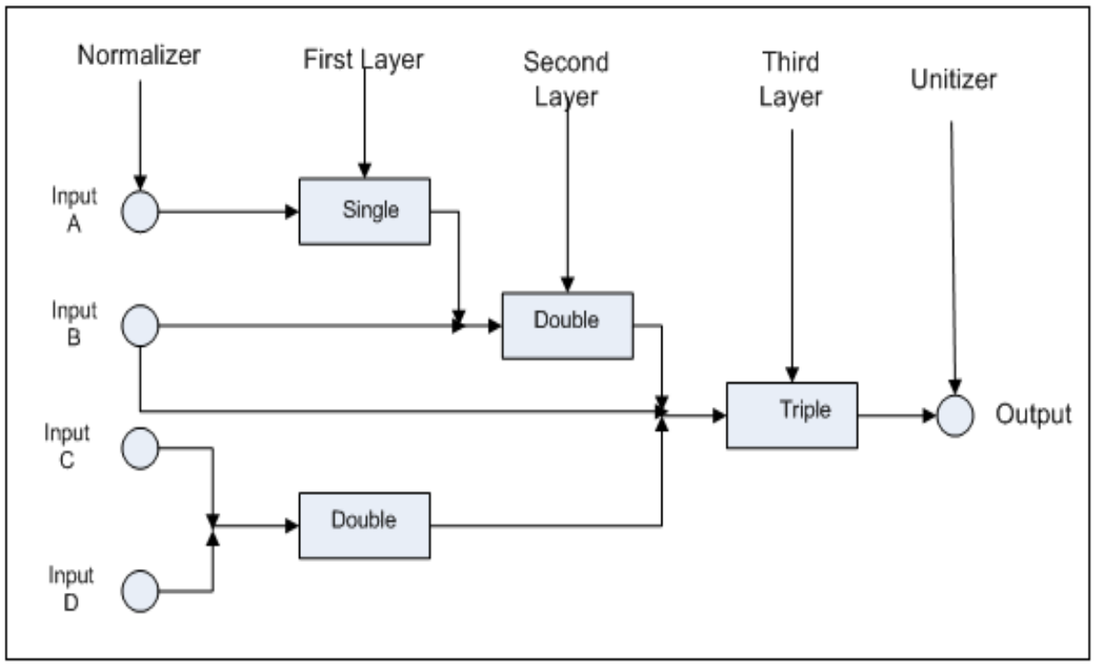


Fig. 1. Abductive network showing various types of functional elements

The algorithm selects the polynomial relationships and the input combinations that minimize the prediction error in each phase. This prevents exponential growth in the polynomial model generated. Iteration is stopped automatically at a point in time that strikes a balance between model complexity for accurate fitting of the training data and model simplicity that enables it to generalise well with new data. In the classical GMDH-based approach abductive network models are constructed by the following 6 steps [Fallow, 1984]:

i. Separating the original data into training data and testing data.

The available dataset are split into training dataset and testing dataset. The training dataset is used for estimating the optimum network model and the testing dataset is used for evaluating the network model obtained on the new data. Usually a 70-30 splitting rule is employed on the original data, but in this work a pre-determined split used by earlier published work was adopted to allow direct comparison of results.

ii. Generating the combinations of the input variables in each layer.

Many combinations of r input variables are generated in each layer. The number of combinations is $p!/\big((p-r)!\,r!\big)$. Here, p is the number of input variables and the value of r is usually taking as 2.

iii. Calculating the partial descriptors

For each input combination, a partial descriptor which describes the partial characteristics of the model is calculated by applying regression analysis on the training data. The following second order polynomial regression relationship is usually used

$$y_k = b_0 + b_1 x_i + b_2 x_j + b_3 x_j + b_4 x_i^2 \quad (1)$$

The output variables y_k in Eq. (l) are called intermediate variables.

iv. Selecting optimum descriptors.

The classical GMDH algorithm employs an additional and independent selection data for selection purposes. To prevent exponential growth and limit model complexity, the algorithm selects only relationships having good predicting powers within each phase. The selection criterion is based on root mean squared (RMS) error over the selection data. The intermediate variables which give the smallest root mean squared errors among the generated intermediate variables (y_k) are selected.

v. Iteration

Steps III and IV are iterated where optimum predictors from a model layer are used as inputs to the next layer. At every iteration, the root mean squared error obtained is compared with that of the previous value and the process is continued until the error starts to increase or a prescribed complexity is achieved. An increasing root mean squared error is an indication of the model becoming overly complex, thus over-fitting the training data and will more likely perform poorly in predicting the selection data.

vi. Stopping the multi-layered iterative computation

Iteration is stopped when the new generation regression equations start to have poorer prediction performance than those of the previous generation, at which point the model starts to become overspecialized and, therefore, unlikely to perform well with new data.

Computationally, the resulting GMDH model can be seen as a layered network of partial descriptor polynomials, each layer representing the results of iteration. Therefore, the algorithm has three main elements: representation, selection, and stopping. Figure 2 shows the flow chart of the classical GMDH-based training.

Abductory Inductive Mechanism (AIM) is a later development of the classical GMDH that uses a better stopping criterion that discourages model complexity without requiring a separate subset of selection data. AIM adopts a well-defined automatic stopping criterion that minimizes the predicted square error (PSE) and penalises model complexity to keep the model as simple as possible for best generalization. Thus, the most accurate model that does not overfit the training data is selected and hence a balance is reached between accuracy of the model in representing the training data and its generality which allows it to fit yet unseen new evaluation data.

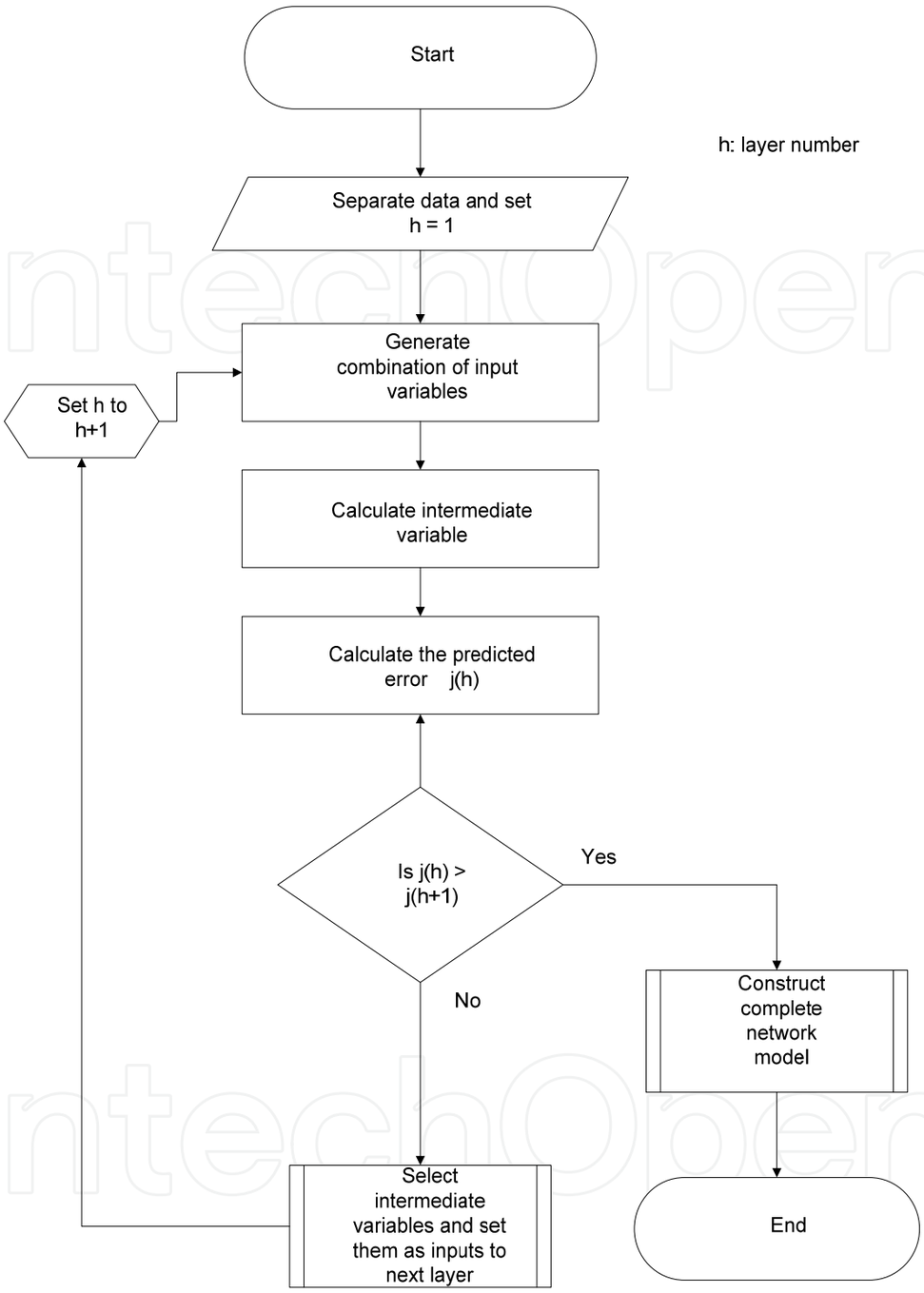


Fig. 2. The flow chart of the classical GMDH-based training.

The PSE consists of two terms [AbTech, 1990]:

$$PSE = FSE + KP \tag{2}$$

Where FSE is the average fitting squared error of the network for fitting the training data and KP is the complexity penalty for the network, expressed as [AbTech, 1990].

$$KP = CPM * (2\sigma^2/N) * K \quad (3)$$

Where CPM is the complexity penalty multiplier, K is the number of coefficients in the network, and σ² is a prior estimate of the model error variance. Usually, a complex network has a high fitting accuracy but may not generalize well on new evaluation data unseen previously during training. Training is automatically stopped to ensure a minimum value of the PSE for the CPM parameter used, which has a default value of 1. The user can also control the trade-off between accuracy and generality using the CPM parameter. CPM values greater than 1 will result in less complex models that are more likely to generalise well with unseen data while values less than the default value will result in a more complex models that are likely to over fit training data and produce poor prediction performance. Figure 3 shows the relationship between PSE, FSE and KP, [AbTech, 1990].

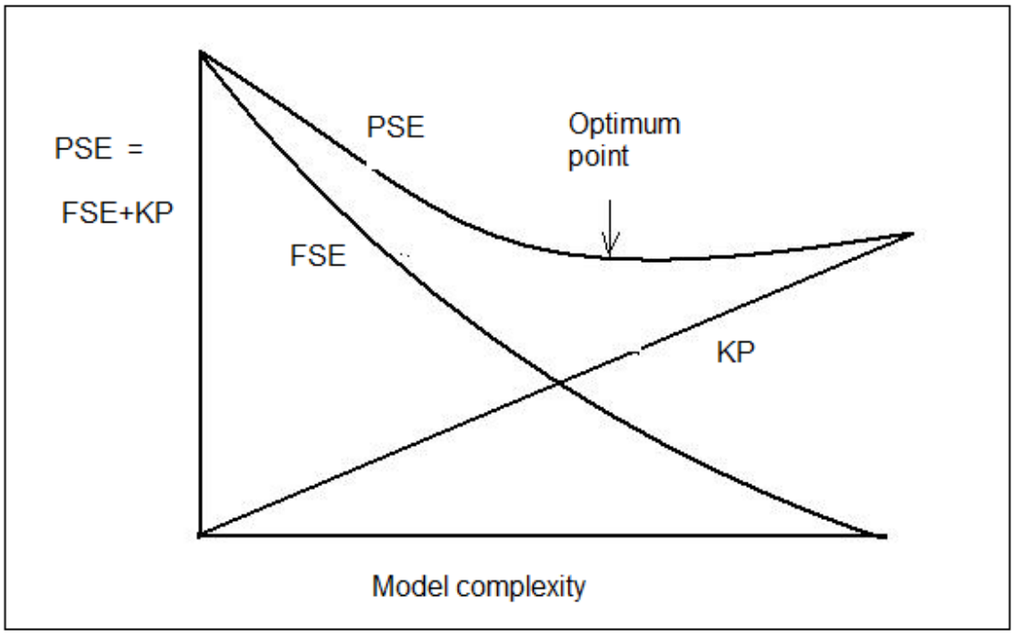


Fig. 3. The predicted square error

3.2 AIM functional elements

The used version of AIM supports several functional elements [AbTech, 1990], see Figure 1, including:

Normaliser: Transforms the original input into a normalized variable having a mean of zero and a variance of unity.

$$y = z_0 + z_1 x \quad (4)$$

Where x is the original input, y is the normalized input, z_0 and z_1 are the coefficients of the normaliser

Unitizer: Converts the range of the network outputs to a range with the mean and variance of the output values used to train the network.

Single Node: The single node only has one input and the polynomial equation is limited to the third degree, i.e.

$$y = z_0 + z_1 + z_2x^2 + z_3x^3 \quad (5)$$

Where x is the input to the node, y is the output of the node and z_0, z_1, z_2 and z_3 are the node coefficients.

Double Node: The double node takes two inputs and the third-degree polynomial equation includes cross term so as to consider the interaction between the two inputs, i.e.

$$y = z_0 + z_1x_i + z_2x_j + z_3x_i^2 + z_4x_j^2 + z_5x_ix_j + z_6x_i^3 + z_7x_j^3 \quad (6)$$

Where x_i, x_j are the inputs to the node, y is the output of the node and z_0, z_1, z_2 …and z_7 are the node coefficients

Triple Node: Similar to the single and double nodes, the triple node with three inputs has a more complicated polynomial equation allowing the interaction among these inputs.

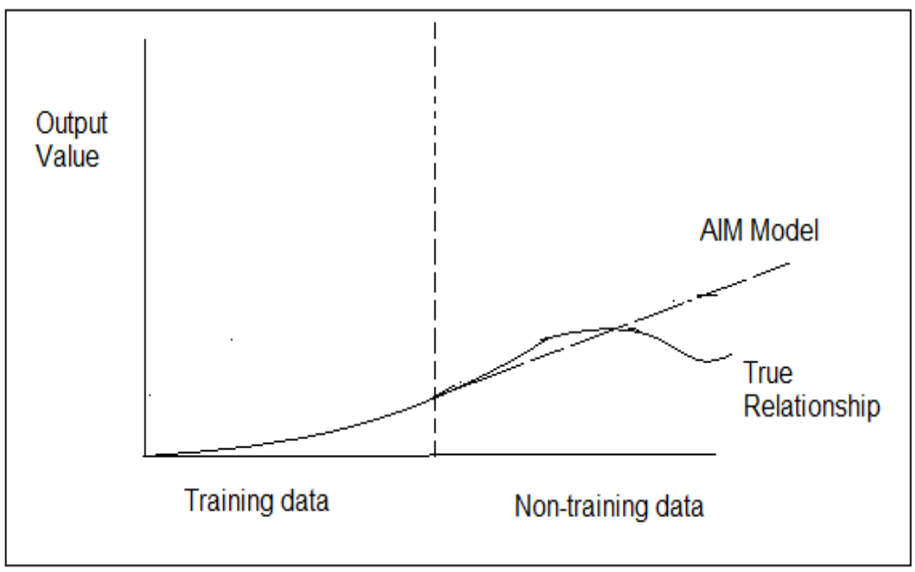


Fig. 4. Non-representative training data

However, not all terms of an element's equation will necessarily appear in a node since AIM will throw away or carve terms that do not contribute significantly to the solution. The eligible inputs for each layer and the network synthesis strategy are defined as a set of rules and heuristics that form an integral part of the model synthesis algorithm as described earlier.

On a final note, any abductive network model is only as good as the training data used to construct it. To build a good model it is important that the training database be representative of the problem space. Figure 4 [AbTech, 1990], illustrates a scenario where the training database used to create the AIM model does not cover an important portion of the problem. Training AIM using only the data to the left of the dotted line will result in a model that generalizes well within the training data range but will be inaccurate in the other region.

4. The dataset and feature discussion

To evaluate and compare the performance of the proposed approach we used the experimental dataset developed in [Lee & Lee, 1996]. The dataset were generated from a calibration chamber, in which field stress conditions were simulated with poorly graded, clean, fine and uniformly layered sand that was dried in air below 2% of water content. The Sand was deposited in the calibration chamber using a method that allows a uniform sand deposit of known relative density to be obtained. The setup was allowed to settle for 24 hours before the model pile was driven into it using a guided steel rod (hammer). The ultimate bearing capacities of the model pile were assumed to be affected by the penetration depth ratio of the model pile, the mean normal stress and the number of blows. So the dataset consist of the following features: penetration depth ratio (i.e. penetration depth of pile/pile diameter), the mean normal stress of the calibration chamber and the number of blows as input and the ultimate capacity (kN) as the output. More detailed description of the dataset generation experiment can be found in [Lee & Lee, 1996].

5. Experiments and results

This section describes the development of abductive networks model for predicting the ultimate bearing capacity using the experimental dataset described in section 3. To allow direct comparison of results, the same splitting used by earlier published work using the dataset was adopted. Two experiments were conducted. In the first experiment, the 28 instances in the dataset were split into a training set of 21 instances and an evaluation of 7 instances. While in the second experiment, 14 instances were selected for training purposes and 14 instances for evaluation. The full training set was used to synthesize an abductive network model with all the 3 features present in the dataset enabled as network inputs. The best model was obtained by adjusting the CPM value. The effect of the CPM values on the models' performance is shown in Table 1. The numbers (e.g. Var_4) indicated at the model input represent the feature selected as input to the model during training, while Var_6 represent the network output corresponding to ultimate pile bearing capacity. It is noted that the model uses 2 inputs out of the 3 inputs which indicates that almost all the features are relevant, with only little redundancy in the feature set

To measure the performance of the trained model after evaluation, two statistical measures namely Root Mean Squared Error (RMSE) and Correlation Coefficient (R^2) were used. A brief description and mathematical formulae are shown below:

5.1 Root Mean-Squared Error

The root mean square error (RMSE)) is a measure of the differences between values predicted by a model and the values actually observed from the phenomenon being modeled or estimated. Since the RMSE is a good measure of accuracy, it is ideal if it is small. This value is computed by taking the square root of the average of the squared differences between each predicted value and its corresponding actual value.

The formula is:

$$RMSE = \sqrt{\frac{[(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_i - y_i)^2]}{n}} \quad (7)$$

Where x_i and y_i are the predicted and actual values respectively while n is the size of the data used.

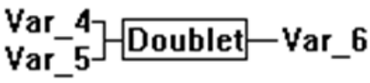
Model	CPM Values	Number of input features selected during model synthesis	Model Performance with training dataset in experiment 1
	0.5	2 out of the 3 features	RMSE = 34.11 kN Correlation = 0.98
	1	2 out of the 3 features	RMSE = 70.22 kN Correlation = 0.91
	2.5	2 out of the 3 features	RMSE = 70.22 kN Correlation = 0.91

Table 1. The effect of the CPM values on the synthesized model's performance

5.2 Correlation coefficient

A correlation coefficient is measure that determines the degree to which two variable's movements are associated. It gives statistical correlation between predicted and actual values. This coefficient is unique in model evaluations. A higher number means a better model, with a value of one (1) indicating a perfect statistical correlation and a value of zero (0) indicating there is no correlation.

$$Correlation\ Coefficient = \frac{\Sigma(y_a - y_a^1)(y_p - y_p^1)}{\sqrt{[\Sigma(y_a - y_a^1)^2]\,\Sigma(y_p - y_p^1)^2}} \tag{8}$$

Where y_a and y_p are the actual and predicted values while y_a^1 and y_p^1 are the mean of the actual and predicted values.

In experiment 1, the best model with CPM = 0.5 was evaluated using the evaluation sets. A satisfactory agreement between the predicted and measured values of the ultimate pile bearing capacity was obtained, which is shown by the cross plots in Figure 5. The maximum error of prediction was 18%. A Root Mean Squared Error (RMSE) and a correlation coefficient (R^2) of 59.22kN and 0.82 were obtained respectively in the first experiment as shown in Table 2.

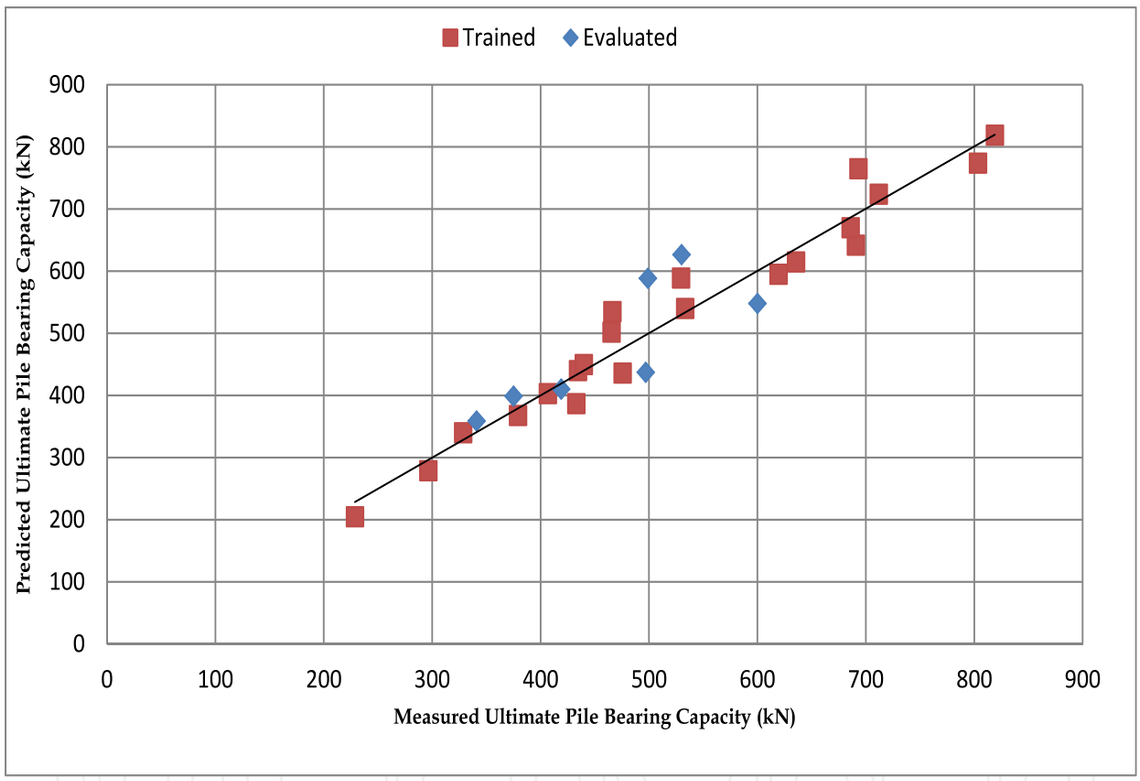


Fig. 5. The cross plot of predicted and measured values of the pile bearing capacity for experiment 1

In the second experiment, after training the abductive network with just 14 instances, the best model was evaluated with the evaluation set of 14 instances. The cross plots in Figure 6 shows the trained and the predicted values of the ultimate bearing capacity. The result showed widely scattered plots, with a RMSE of 92.5kN and Correlation Coefficient of 0.83. The reason for the poor result in this case can be attributed to the small number of the training data which was not enough for the model to learn the entire pattern in the data set. Therefore, it could be concluded that a certain number of training data sets was needed to obtain reasonable predictions.

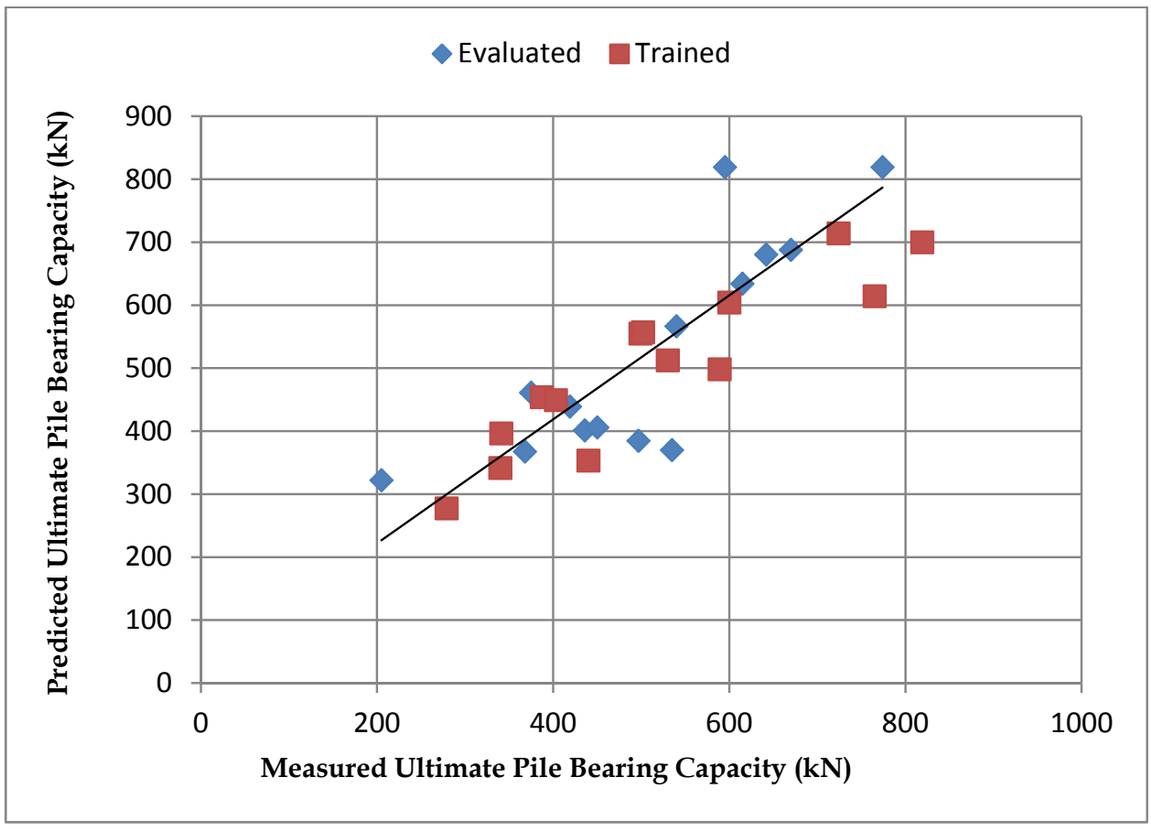


Fig. 6. The cross plot of predicted and measured values of the pile bearing capacity for experiment 2

Finally, the maximum prediction error of other modeling tool used previously on this dataset was compared with that of abductive model synthesized in this work as shown in Figure 5. The comparison indicates that the abductive network model performed much better in terms of prediction error, with almost 9% improvement compared to that obtained with Neural Network reported in [Lee & Lee 1996].

	Experiment 1		Experiment 2	
	Training set (21 instances)	Evaluation set (7 instances)	Training set (14 instances)	Evaluation set (14 instances)
RMSE (kN)	34.11	59.22	70.2	92.5
Correlation(R^2)	0.98	0.82	0.91	0.83

Table 2. Performance measure of the two experiments on the training and test data set

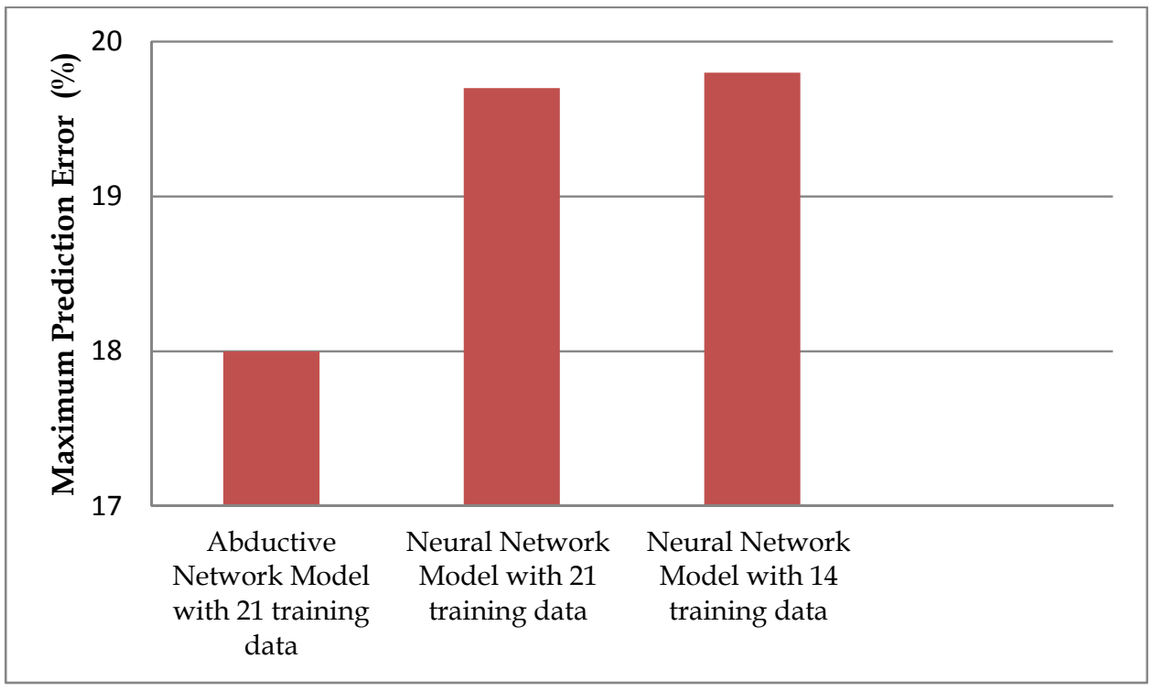


Fig. 7. Comparison of GMDH-based abductive network model with other previous model used on this dataset in term of Maximum Prediction Error (%)

6. Conclusion

This work demonstrates the use of abductive machine learning techniques for the prediction of pile bearing capacity. A RMSE value of 59.22kN and 92.5kN and a correlation coefficient of 0.82 and 0.83 were obtained with respect to the pile bearing capacity values predicted in two separate experiment conducted respectively. An improvement of almost 9% in terms of prediction error was recorded. This result indicated that the proposed abductive network approach yields a better performance compared to the other already implemented technique using the same dataset mentioned in the introduction section. However, the experiments conducted revealed that for a good prediction, a large number of training set is required to train the model before evaluation. So, to validate the performance of the abductive network approach it is recommended that data set obtained from the fields are use in further studies. This will help realize the full potential of abductive network approach in pile bearing capacity prediction.

Meanwhile, the work has outlined the advantages of abductive networks and has placed it in the perspective of geotechnical engineering problem computing point of view. Thus, researchers are encouraged to consider them as valuable alternative modeling tool. Hopefully, future work will consider the possibility of extending the approach to modelling of soil behaviour and site characterization.

7. Acknowledgement

The authors will like to acknowledge Dr. R. E. Abdel-Aal of the Department of Computer Engineering, King Fahd University of Petroleum & Minerals for providing the tool used in this study.

8. References

Abdel-Aal, R.E. (2004). Short-Term Hourly Load Forecasting Using Abductive Networks, *IEEE Transaction on Power Systems*, vol. 19, No.1, pp. 164 -173

Abdel-Aal, R.E. (2005). Improved classification of medical data using Abductive network committees trained on different feature subsets, *Computer Methods and Programs in Biomedicine*, vol. 80, no.2, pp. 141-153

AbTech, Corporation (1990). *AIM user's manual*, Charlottesville, VA

Abu-Kiefa, M. A. (1998). General regression neural networks for driven piles in cohesionless soils, *Journal of Geotechnical & Geoenviromental Engineering*, Vol. 124, No.12,

Adarsh, S.; Dhanya,R. Krishna,G. Merlin,R. & Tina,J. (in press). Prediction of Ultimate Bearing Capacity of Cohesionless soils using Soft Computing Techniques, *Journal of ISRN Artificial Intelligence*.

Agarwal, A. (1999). Abductive networks for two-group classification: A comparison with neural networks. *The Journal of Applied Business Research*, Vol. 15, No. 2, pp. 1–12.

Bustamante, M. & Gianeselli, L. (1982). Pile bearing capacity prediction by means of static penetrometer CPT. *Proceeding of the second European symposium on penetration testing*, Amsterdam, May

Chow, Y.K; Chan, W.T. Liu, I.F. & Lee, S.L. (1995). Predication of pile capacity from stress wave measurements: a neural network approach, *International Journal of Numerical & Analytical Methods in Geomechanics*, Vol. 19, pp. 107–126

El-Sayed, E.M., Abdel-Aal, R.E. (2008). Spam Filtering with Abductive Networks, *International Joint Conference on Neural Networks*, Hong Kong

Farlow, S. J. (1984). The GMDH algorithm. In *Self-organizing methods in modeling: GMDH type algorithms*, S. J. Farlow (Ed.), pp. 1–24, New York: Marcel-Dekker

Lawal, I. A.; Abdel-Aal, R. E. & Mahmoud, S. A. (2010). Recognition of Handwritten Arabic (Indian) Numerals Using Freeman's Chain Codes and Abductive Network Classifiers, *IEEE 20th International Conference on Pattern Recognition*, Istanbul, pp.1884-1887

Lee, B. Y; Liu, H. S. & Tarng, Y. S. (1999). An Abductive Network for Predicting Tool Life in Drilling, *IEEE Transactions on Industry Applications*, Vol. 35, No. 1, January

Lee, I. & Lee, J. (1996). Prediction of Pile Bearing Capacity Using Artificial Neural Networks, *Journal of Computers and Geotechnics*, Vol. 18, No. 3, pp. 189-200,

Mehra, R.K. (1977). Group Method for Data handling (GMDH): Review and Experience, *Proceedings of the IEEE Conference on Decision and Control*, New Orleans, pp. 29-34

Pal, M. & Deswal, S. (2008). Modeling Pile Capacity Using Support Vector Machines and Generalized Regression Neural Network, *Journal of Geotechnical and Geoenviromental Engineering*, Vol.134, No. 1021 pp. 134-7

Pal, M. & Deswal, S. (2010). Modeling pile capacity using Gaussian process regression, *Journal of Computers and Geotechnics*, Vol. 37, pp. 942–947,

Pile, (2011).In *Encyclopeadia Britannica*, Retrieved from http://www.britannica.com/EBchecked/topic/460377/pile pp. 1177-1185

Samui, P. (2011). Prediction of pile bearing capacity using support vector machine, *International Journal of Geotechnical Engineering,* pp. 95-102,

Shahin, M. A.; Jaksa, M. B. & Maier, H. R. (2001) Artificial Neural Network Applications In Geotechnical Engineering, *Australian Geomechanics*, Vol. 36, No.1, pp. 49-62

Teh, C.I; Wong, K.S. Goh, A.T.C. & Jaritngam, S. (1997). Prediction of pile capacity using neural networks, *Journal of Computing Civil Engineering* Vol. 11, No.2, pp. 129–38.

IntechOpen

IntechOpen