JSTT JOURNAL OF SCIENCE AND TRANSPORT TECHNOLOGY

Journal of Science and Transport Technology
Journal homepage: https://jstt.vn/index.php/en

UNIVERSITY OF TRANSPORT TECHNOLOGY

Using GA-ANFIS machine learning model for forecasting the load bearing capacity of driven piles

Dam Duc Nguyen[1,*], Hai Phu Nguyen[2], Dung Quang Vu[1], Indra Prakash[3], Binh Thai Pham[1]
[1]University of Transport Technology, Hanoi 100000, Vietnam
[2]People's Committee of Phu Xuyen District, Hanoi 100000, Vietnam
[3]DDG (R) Geological Survey of India, Gandhinagar, 382010, India

Article info
Type of article:
Original research paper

DOI:
https://doi.org/10.58845/jstt.utt.2023.en.3.2.26-33

***Corresponding author:**
E-mail address:
damnd@utt.edu.vn

Received: 09/06/2023
Revised: 26/06/2023
Accepted: 28/06/2023

Abstract: This paper is aimed to apply hybrid machine learning model namely GA-ANFIS, which is a combination of Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithm (GA), for the prediction of total bearing capability of driven piles. A database of 95 Pile Driving Analyzer (PDA) tests carried out at the win power project in Hoa Binh province, Vietnam was used to develop hybrid model. The database was split into 70:30 ratio for training (70%) and validating (30%) model. Accuracy of the model was evaluated using statistical standard indicators: Coefficient of determination (R^2), Mean Absolute Error (MAE), and Root mean squared error (RMSE). Results indicated that the GA-ANFIS model has a good performance in correct prediction of the total bearding capability of driven piles on both training (R^2 = 0.976) and testing (R^2 =0.925) datasets. Therefore, the GA-ANFIS hybrid model is a promising tool for quick and accurate prediction of the total bearing capability of driven piles for the consideration in design and construction of the structures.

Keywords: adaptive inference system, anfis, driven pile, load capacity, artificial intelligence, machine learning.

1. Introduction

In the designing of high civil engineering structures such as bridges, buildings, tower and other offshore structures on soil, correct estimation of load bearing capacity of piles is required for long term stability of structures [1]. Pile foundation is designed where top soil is of low bearing capacity. Piles are designed to transfer the load to deeper high bearing capacity soil or rock. The piles transfer the loads into the soil on sides by “skin resistance” and at the end of the piles through the tip under the compression (“end bearing”). Sub-surface data of different soil layers and rocks is required for the designing of the pile foundations.

The actual load capacity of the pile at the site can be determined large deformation Pile Driving Analyzer test, static compression experiments along the shaft and Osterberg experiment [2, 3]. In particular, the static compression experiments along the axis have a higher accuracy than the large PDA deformation test (due to the actual simulation of the work of the pile). It is the most direct and reliable method [1, 4]. However, PDA experiments are faster and cheaper in comparison to static compression experiments. Moreover, they are very effective with closed piles, and piles between the river and in the middle of the sea. Experiments are conducted on piles at the site.

Field testing of piles bearing capacity at site is not easy and involve movement of heavy equipment at the test locations, therefore, attempts have been made to forecast the loading capacity of the pile using alternative methods.

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) methods have been successfully applied in the field of civil engineering including for the estimation of load-bearing capacity of the piles [5, 6] and subsidence of the piles [7, 8]. In this study, we have used hybrid AI model GA-ANFIS, which is a combination of Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithm (GA), for the prediction of total bearing capability of driven piles. It is noticed that it is the first time ANFIS and GA was combined and developed a hybrid model for prediction of load-bearing capacity of the piles. In order to develop model, in total 95 PDA pile loading test data including 5 Wind Power Plant project, Hoa Binh, Vietnam were used along with physical parameters of piles and geo-mechanical parameters of in-situ soils. Performance of the model was evaluated using Standard statistical indicators namely Coefficient of determination (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). Matlab software was used for the data processing and development of model.

2. Data used

The data used in this study was collected from 95 fields PDA tests conducted at the 5 Wind Power Plant Project, Hoa Binh, Vietnam.

Table 1. Table showing statistical values of the parameters used in the development of model for forecasting the load capacity of centrifugal driven concrete piles

Parameters	Abbreviation	Unit	Minimum	Maximum	Average	StD
Inputs						
Pile diameter (D)	X1	mm	500.00	800.00	694.737	143.192
Embedded length (L)	X2	m	24.00	59.00	39.001	11.619
Settlement of piles (S)	X3	mm	11.838	51.182	24.306	6.435
Modulus of dynamic elasticity (Ed)	X4	Ton/cm^2	497.300	539.900	525.041	14.920
Cross sectional area at the pile top (At)	X5	cm^2	1,055.575	2,902.800	2,099.371	775.553
Cross sectional area at the pile tip (Ab)	X6	cm^2	1,963.495	5,026.548	3,950.139	1,462.042
Average standard penetration number along the pile shaft (SPT-NS)	X7	-	5.000	14.939	11.587	2.375
Average standard penetration number along the pile tip (SPT-NB)	X8	-	9.000	39.000	19.971	5.925
Cohesion along the pile shaft (Cs)	X9	kPa	7.510	44.931	20.400	11.177
Friction angle along the pile shaft (ψs)	X10	degree	0.000	12.000	7.211	3.692
Cohesion along the pile tip (Cb)	X11	kPa	0.000	203.750	48.795	49.295
Friction angle along the pile tip (ψb)	X12	degree	0.000	27.500	8.974	9.214
Outputs						
Total bearing capacity (RU)	Y1	Tons	162.000	896.000	446.029	191.535

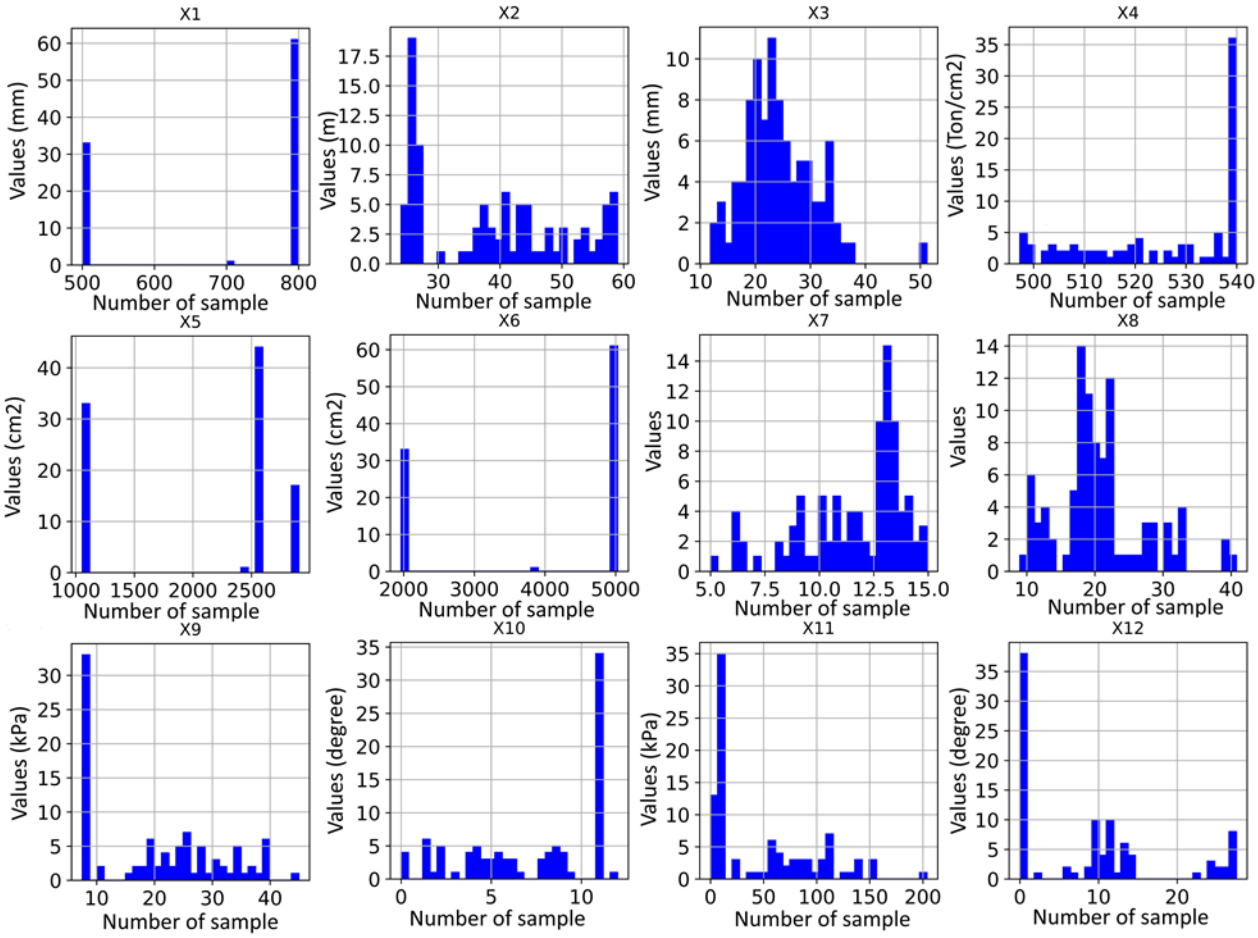

Fig 1. Data distribution of input variables

To forecast the load capacity of centrifugal driven concrete piles, a total of 12 input parameters namely: Pile diameter (D), Embedded length (L), Settlement of piles (S), Modulus of dynamic elasticity (Ed), Cross sectional area at the pile top (At), Cross sectional area at the pile tip (Ab), Average standard penetration number along the pile shaft (SPT-NS), Average standard penetration number along the pile tip (SPT-NB), Cohesion along the pile shaft (Cs), Friction angle along the pile shaft (ys), Cohesion along the pile tip (Cb), Friction angle along the pile tip (yb), Total bearing capacity (RU) were used. PDA experiment data of determined Data of Load capacity of driven piles obtained from PDA experiment was used as output parameter. The maximum, smallest, medium and data deviation of these parameters are shown in Table 1. Figure 1 shows the data distribution of the input variables. The database is divided into two parts in the ratio of 70:30 [9]. 70% data was used for training the model and 30% for testing/ validation the model.

3. Research method

3.1. ANFIS model

The ANFIS is an Adaptive Neuro Fuzy Inference System, developed by Jang in the year 1993 [10, 11]. It combines the advantages of artificial neural networks and fuzzy inference methods, allowing complex and unclear problems to solve problems by combining qualitative and quantitative knowledge.

The ANFIS model can be used to model the relationship between the input and the output of an unclear system by using the if-then laws to describe the reasoning rules of the system [12].

One of the outstanding features of ANFIS is the ability to learn automatically from input data. ANFIS uses an adaptive learning algorithm to adjust the model's parameters so that its output is best suited to training data.

ANFIS model is widely used in areas such as automatic control, data prediction and analysis, and can be applied to solve problems in areas such as economics [13], and medicine [14]. The ANFIS model has been shown to be very effective in solving complex problems, especially in unclear prediction and control of systems. In this study, ANFIS was selected to forecast the load capacity of centrifugal driven concrete piles.

3.2. GA technique

GA (Genetic algorithm) is a genetic optimization technique developed based on genetic theory in biology [15, 16]. This technique helps find optimal solutions for non-oriented optimization problems, which is no specific constraints on the structure of the solution.

In order to implement the GA algorithm, it is necessary to have a collection of individuals (also known as genes) representing solutions. These individuals will be represented in the form of bits or numbers, and are evaluated based on a target function to determine the good level of the solution. Ga algorithm includes the following basic steps [17]: (i) Initialize the original population with a set of random individuals, (ii) assess the good level of each individual using the target function, (iii) select a set of individuals It is best to continue handling (also known as a new population), (iv) to perform crossover and mutations (MUTATION) to create new individuals for new populations, (V) repeat Again the above steps until a stop condition (usually when an optimal solution is achieved or when it reaches a specific number of repetitions).

With the ability to find optimal solutions for difficult and complex optimization problems, GA techniques have become a useful tool in solving practical problems. In this study, the throttle technique is used to optimize the super parameters of the ANFIS model to create a GA-ANFIS hybrid model for the forecast of the load capacity of the centrifugal driven concrete piles.

3.3. Evaluation methods

Statistical indicators R^2, RMSE and MAE are generally used for the evaluation of accuracy of ML prediction or forecast models [18, 19]. R2 measures the relationship between independent and dependent variables [20]. It indicates how well model predicts dependent variable. The range of value of R2 is from 0 to 1, where value of nearly 1, indicate that relationship between independent and dependent variable is higher on the contrary value "0" indicate no relationship. RMSE is an assessment indicator used to measure the accuracy of predictable values compared to the actual value [21]. RMSE is calculated by taking the square root of the square average of the deviation between the predicted value and the actual value. The smaller the value of RMSE, the better the model. MAE is an assessment indicator used to measure the accuracy of predictable values compared to the actual value [22]. MAE is calculated by taking the average absolute value of the deviation between the predicted value and the actual value. The smaller the value of MAE, the better the model. There are other statistical methods such as and statistical methods such as Receiver Operating Characteristic-Area Under Curve (AUC), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Specificity (SPF), Sensitivity (SST) to evaluate predictive performance of the model. However, in the present study we have used simple statistical methods (R^2, RMSE and MAE) which also provide accurate results.

4. Results and Discussion

R^2, RMSE and MAE indicators values of NFIS model results are shown in Figure 3 and Table 2. Figure 4. Value of forecast trimming capacity and actual measurement load capacity Figure 4 shows the difference between the forecasted value and the actual measurement value of the load bearing strength. Figure 5 shows the chart of distribution of the error value of the loading capacity forecasting model of the pile. Specifically, the R^2 value of the training model using the training datasets is 0.976 and using validation datasets is 0.925; The RMSE value of the training model using training datasets is 0.582 and the using of validation datasets is 0.920; The MAE value of the model using training data is 0.356 and on the validation data is 0.544. These evaluation results show that the estimated and actual load capacity of piles are not much different (Figure 4). Therefore, the GA-ANFIS model is a very good for estimating load capacity of piles.

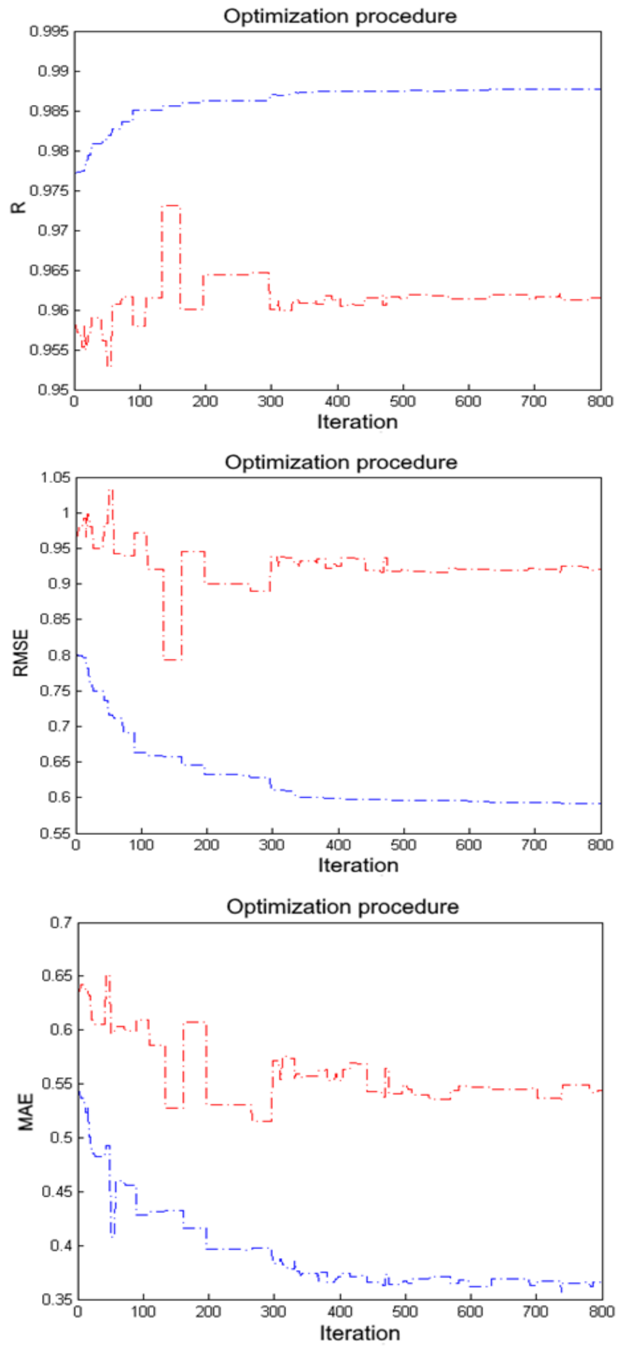

Fig 2. Optimization procedure

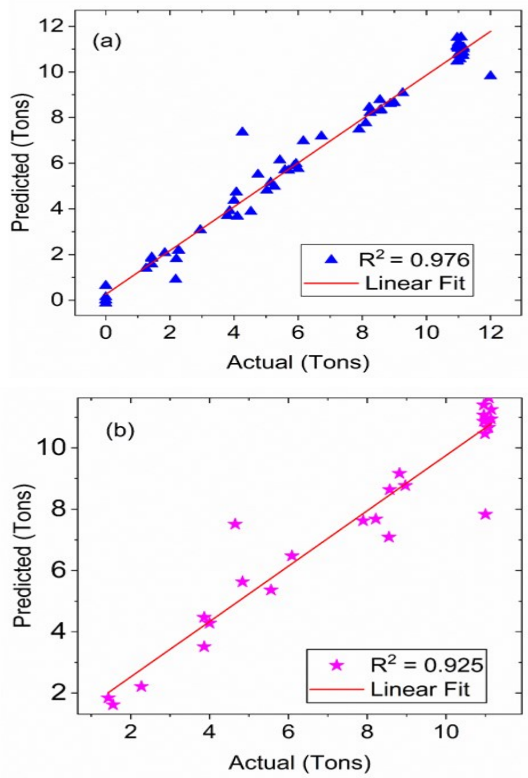

Fig 3. R^2 value of GA-ANFIS model: (a) Training datasets; (b) Validation datasets

Table 2. Summary results of GA-ANFIS models use indicators R, R^2, RMSE, MAE

Parameters	R	R^2	RMSE	MAE
Training data	0.987	0.976	0.582	0.356
Test data	0.962	0.925	0.920	0.544

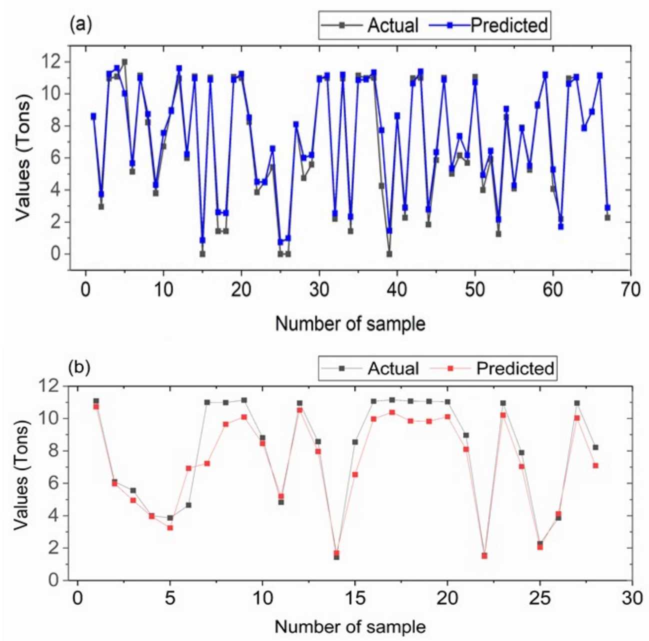

Fig 4. Value of forecast trimming capacity and actual measurement load capacity

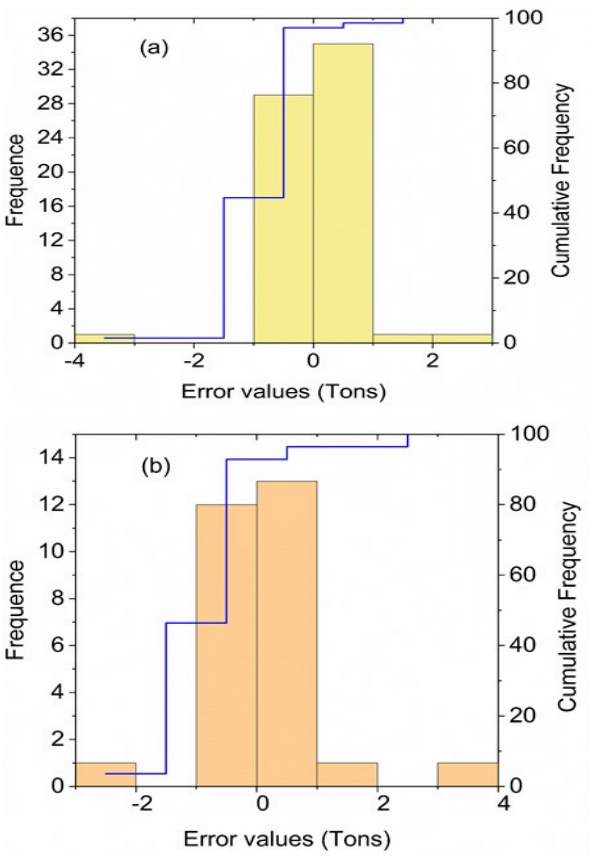

Fig 5. Chart of error distribution of GA-ANFIS model

Prediction capability of GA-ANFIS model is good because it takes advantages of two algorithms that is GA and ANFIS. Specifically, the advantage of ANFIS includes. (i) ANFIS model has the ability to learn and adapt to new data, allowing the model to optimize model parameters to make accurate predictions, (ii) ANFIS model capable of processing Fuzzy data, which is data that cannot be completely classified, (iii) ANFIS model has the ability to learn automatically, meaning that the model can automatically update and improve when it is trained on new data [23, 24]. GA model has the advantage of using techniques which include (i) Ga is capable of creating a set of diverse solutions, helping to discover a lot of solution space and increase opportunities to find good solutions, (II) GA has the ability to find many minor values. Different at the same time, increasing the ability to find the optimal solution, and (iii) GA is able to find global solutions, which means finding the best solution in the solution space [25, 26].

5. Conclusions

Results of the study indicate that the GA-ANFIS hybrid AI model has high excellent accuracy in forecasting the load capacity of the centrifugal driven concrete piles on both training (R^2 = 0.976) and testing/validation (R^2 =0.925) datasets based on the limited input parameters of piles (physical parameters) and foundation groundmass (geo-mechanical properties). Therefore, the GA-ANFIS hybrid AI model is a good potential tool that can be used for forecasting the load capacity of centrifugal driven concrete piles for proper designing and construction of the structures. In future study we will use this model in other areas also for assessing its wider applicability.

References

[1] B.T. Son, P.C. Huyen. (2011). Khả năng chịu tải của cọc từ kết quả thử động biến dạng lớn (PDA) và nén tĩnh. *Tạp chí Xây dựng*, 6, 78-81.

[2] H. Moayedi, M. Mosallanezhad, R. Nazir. (2017). Evaluation of maintained load test (MLT) and pile driving analyzer (PDA) in measuring bearing capacity of driven reinforced concrete piles. *Soil Mechanics and Foundation Engineering*, 54, 150-154.

[3] N.N. Phuc. (2022). Đánh giá hệ số cường độ sức kháng bên đơn vị trong tính toán sức chịu tải dọc trục cọc theo đất nền dựa trên kết quả thí nghiệm PDA. *Tạp chí Khoa học và Công nghệ*, 57, 118-126.

[4] V.C. Nguyen. (2011). Nghiên cứu một số phương pháp xác định sức chịu tải của cọc ngoài hiện trường.

[5] M.A. Benbouras, A.-I. Petrişor, H. Zedira, L. Ghelani, L. Lefilef. (2021). Forecasting the bearing capacity of the driven piles using advanced machine-learning techniques. *Applied Sciences*, 11(22), 10908.

[6] N. Kardani, A. Zhou, M.Nazem, S.-L.Shen. (2020). Estimation of bearing capacity of piles in cohesionless soil using optimised machine learning approaches. *Geotechnical and Geological Engineering*, 38, 2271-2291.

[7] S. Kumarand, S. Robinson. (2023). Estimating the Pile Settlement Using a Machine Learning Technique Optimized by Henry's Gas Solubility Optimization and Particle Swarm Optimization. *Advances in Engineering and Intelligence Systems*, 1(4), 50-61.

[8] M. Kumar, P. Samui. (2020). Reliability analysis of settlement of pile group in clay using LSSVM, GMDH, GPR. *Geotechnical and Geological Engineering*, 38, 6717-6730.

[9] Q.H. Nguyen, H.-B. Ly, L.S. Ho, N. Al-Ansari, H.V. Le, V.Q. Tran, I. Prakash, B.T. Pham. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering*, 6, 1-15.

[10] Đ.M. Ha. (2020). Ứng dụng hệ suy luận nơron mờ thích nghi anfis để điều khiển hệ thống điều tốc nhà máy thủy điện Sê San 4. *Trường Đại học Bách khoa - Đại học Đà Nẵng*.

[11] J.-S. Jang. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.

[12] D. Karaboga, E. Kaya. (2019). Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. *Artificial Intelligence Review*, 52, 2263-2293.

[13] A. Đokić, S. Jović. (2017). Evaluation of agriculture and industry effect on economic health by ANFIS approach. *Physica A: Statistical Mechanics and its Applications*, 479, 396-399.

[14] A. Yadollahpour, J. Nourozi, S.A. Mirbagheri, E. Simancas-Acevedo, F.R. Trejo-Macotela. (2018). Designing and implementing an ANFIS based medical decision support system to predict chronic kidney disease progression. *Frontiers in Physiology*, 9, 1753.

[15] T.T. Hoang. (2012). Kỹ thuật dự báo dựa theo hồi quy vectơ hỗ trợ và thử nghiệm áp dụng dự báo thành tích vận động viên. *Đại học Quốc gia Hà Nội*.

[16] S. Mirjalili, S. Mirjalili. (2019). Genetic algorithm. *Evolutionary Algorithms and Neural Networks: Theory and Applications*, 43-55.

[17] S. Katoch, S.S. Chauhan, V. Kumar. (2021). A review on genetic algorithm: past, present, and future. *Multimedia Tools and Applications*, 80, 8091-8126.

[18] T.A. Nguyen, H.B. Ly. (2021). Ứng dụng các mô hình học máy dựa trên thuật toán cây để giải bài toán dự báo sức kháng cắt của dầm BTCT không cốt đai. *Tạp chí điện tử Khoa học và Công nghệ Giao thông*, 1(1), 1-12.

[19] S. Dadhich, J.K. Sharma, M. Madhira. (2021). Prediction of ultimate bearing capacity of aggregate pier reinforced clay using machine learning. *International Journal of Geosynthetics and Ground Engineering*, 7, 1-16.

[20] L.J. Saunders, R.A. Russell, D.P. Crabb. (2012). The coefficient of determination: what determines a useful R2 statistic? *Investigative ophthalmology & visual science*, 53(11), 6830-6832.

[21] T. Chai, R.R. Draxler. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific model development*, 7(3), 1247-1250.

[22] C.J. Willmott, K. Matsuura. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, 30(1), 79-82.

[23] N. Walia, H. Singh, A. Sharma. (2015). ANFIS: Adaptive neuro-fuzzy inference system-a survey. *International Journal of Computer Applications*, 123(13), 32-38.

[24] M. Al-Mahasneh, M. Aljarrah, T. Rababah, M. Alu'datt. (2016). Application of hybrid neural fuzzy system (ANFIS) in food processing and technology. *Food Engineering Reviews*, 8, 351-366.

[25] S. Sivanandam, S. Deepa, S. Sivanandam, S. Deepa. (2008). Genetic algorithms. *Springer*.

[26] K. Gallagher, M. Sambridge. (1994). Genetic algorithms: a powerful tool for large-scale nonlinear optimization problems. *Computers & Geosciences*, 20(7-8), 1229-1236.