Recent Trends in Civil Engineering and Built Environment Vol. 2 No. 1 (2021) 700-709

RTCEBE

Homepage: http:// publisher.uthm.edu.my/periodicals/index.php/rtcebe
e-ISSN :2773-5184

Concrete Strength Prediction Using Artificial Neural Network Machine Learning Algorithm

Kueh Lee Gek[1], Nickholas Anting Anak Guntor[1, *]

[1]Faculty of Civil Engineering and Built Environment,
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/rtcebe.2021.02.01.076
Received 30 January 2021; Accepted 28 April 2021; Available online 30 June 2021

Abstract: Compressive strength of concrete is an important parameter in the design of reinforced concrete structure and it is also used by industry to show the performance of the concrete. However, it is difficult to predict compressive strength of concrete since it is affected by many factors such as age, ratio of water to cement conditions of curing and compaction. This study is to predict the concrete strength using artificial neural network machine learning algorithms. The dataset of this study will be collected from the published literature that consist 191 rows of and focused on the parameter that will affect the compressive strength of the concrete. The measurable parameter for input variable consists of water, cement, fine aggregate, coarse aggregate, various types of admixtures and et cetera. Besides that, from this parameter it will extend into a few variables such as types of aggregate and type of sand. Compressive strength categorized as an output variable. Artificial neural network is one of the algorithms used in machine learning for modeling the data. An artificial neural network model will be developed for predicting the compressive strength of concrete. The modeling framework can be carried out by using KNIME software without writing a programming code. Dataset will split into 2 partitions: 80% for training set and 20% for testing set. Some input variables are removed due to high correlation to each other to achieve the best possible performance. The result shown 5 hidden layer and 14 hidden neurons per layers performed the best neural network architecture in this study. The R-squared value for the generated model is 0.838, it is considered high and quite good for the model. The graph shown predicted strength obtained from the model is quite similar for actual strength. It means that the model has a reliable prediction capability. With the help of this model, it can be used to verify if the designed proportion of the design mix can fulfil the requirement for target strength and do not required to produce a lot of concrete sample in lab.

Keywords: Compressive Strength, Neural Network, Performance, R-Squared

1. Introduction

**Corresponding author: nickholas@uthm.edu.my*

Concrete is a composite material that mainly contain of cement, water, sand and aggregate, which are mixed together and undergo the chemical reaction known as hydration process to become harder and obtain strength. Concrete is the most popular materials that used in the construction in the world. Compressive strength is concrete's capacity to withstand the surface loads without cracking or deflection by using compression testing machine. The parameter such as age, ratio of water to cement, conditions of curing, component and compaction can affect the compressive strength of concrete and the prediction of concrete compressive strength usually using the traditional method based on linear and non-linear methods of regression equations [1].

It is difficult to acquire an accurate equation therefore it is required a lot of techniques and experience. Machine learning is one of the techniques were used to predict the concrete's compressive strength to arrive at these difficulties [2]. A large quantity of data can be analyzed by machine learning. Researchers have explored the potential of artificial neural network, a non-linear modeling technique in order to predict the compressive intensity of concrete due to its ability to efficiently learn input-output relationships for any complex problem [3]. Thus, this is study is purposely conduct to develop a machine learning algorithm using neural network to predict the compressive strength of concrete based on historical data parameters.

2. Literature Review

This chapter describes several literature reviews on the innovative approach for concrete, compressive strength of concrete abd concrete mix design. According by the types of machine learning and neural network algorithms. Lastly, KNIME platform, the software use to construct machine learning model in this study.

2.1 Innovative Approach for concrete

Normally, the conventional concrete is mixing of 4 basic components, cement, water, sand and aggregate. Portland cement is the most prevalent cement type that is commonly used to manufacture concrete. According to [4], manufacture of cement is a very energy-intensive process, resulting in the emission of extremely large quantities of greenhouse gases, while the extensive use of aggregate is correlated with a huge consumption of natural resources. In order to achieve environmental sustainability, [5] investigated that it is possible to replace Portland cement with high-volume fly ash into and replace fine aggregate by crumb rubber to increase concrete's compressive strength and mechanical properties. He found that the compressive strength's value was decreasing when the crumb rubber was increasing and the use of fly ash as a partial replacement of cement can reduce the greenhouse carbon dioxide emission.

2.2 Compressive Strength of Concrete

According to [6], concrete strength is actually controlled by the proportioning of cement, fine and coarse aggregates, water and various type admixtures. Meanwhile, the water to cement ratio is the chief determining factor for concrete strength. He probed that the lower the water cement ratio, the greater the compressive force. Usually, the compressive strength checks are done at about 7 or 28 days from the day the concrete is casted.

2.3 Concrete Mix Design

According to [7], the design of the concrete mix is a method of selecting the components of the required concrete and deciding their optimal proportion that would produce concrete in the most economical way possible that meets a certain compressive strength and desired workability. Furthermore, the design of the concrete mix is based on the concept of workability, required strength and durability of hardened concrete and on-site conditions that help determine the specifications of workability, strength and durability.

2.4 Machine Learning Approach

Machine learning approaches are being increasingly used to model concrete material's behavior and have become an important field of study. In many areas of civil engineering, the use of machine learning based applications has recently increased, ranging from engineering design to project planning. Machine learning is an artificial intelligence subfield. Artificial intelligence is a computational approach that aims to replicate human capacity for cognition through symbol manipulation and symbolically organized knowledge bases in order to solve engineering problems that defy solution using traditional methods. Machine learning can be classified into three types: supervised learning, unsupervised learning and reinforcement learning.

[8] investigated that supervised machine learning is quite popular in classification problems because the goals are frequently to get the computer to gain the classification system that we have developed. After having sufficient training for the program, it is able to provide expectations for any new data. Additionally, the learning algorithm compares its output with the expected output correctly, and detects the errors to adjust the model.

Besides that, [9] found that the aim of unsupervised learning is to split the training dataset into clusters in such a way that data in all clusters exhibits a high degree of similarity across all clusters. Unsupervised algorithms of machine learning are used where knowledge provided is neither categorized nor labelled. For example, the input cannot be directed to the algorithm without the aspect of known data, from which the unsupervised term arises. All these data are fed to the algorithm of machine learning and are used to train the model.

For reinforcement machine learning, there is no information provided about the ideal category signal or specific objectives. Reinforcement algorithms are forced by trial and error to know the optimal target. In reality, reinforcement learning allows an agent to determine the optimal behaviour within a particular context in order to maximize the output of the model [9]. According to [10], the behaviour of the algorithm is determined by a sequence of rewards and penalties based on whether its decision against a defined objective are correct or incorrect, as described by the researcher and application of this type of learning in brain disorders have been very limited so far.

2.4 Artificial Neural Network Algorithms

In general, artificial neural networks are one of the algorithms that can be used in machine learning for modelling the data. Artificial neural networks have hundreds or thousands of artificial neurons, namely processing units, linked to nodes. The processing units are made up of input and output units. Artificial neural networks are commonly used in multiplex states where the typical methods of computation are not effective enough to solve them. It can model data based on the human neuron model. Another important parameter affecting the artificial neural network is the number of neurons in artificial neural network layers. The number of neurons in the input layer must be proportional to the number of independent or input variables and the number of output neurons is equal to the number of output or dependent variables in a prediction problem based on cause and effect relationship [11].

[12] applied artificial neural network for predicting the compressive strength of concrete. In addition, [13] adopted artificial neural network to predict the recycled aggregate concrete compressive strength. The results indicate the artificial neural network is a well-organized model to be used as a method to predict compressive strength of RAC which is consists of various form and sources of recycled aggregate. [13] also used artificial neural network for predicting the 28 days compressive strength of concrete with different mix design.

[14] probed that the coefficient of determination R^2 is adopted for evaluating the accuracy of qualified networks. The coefficient is used to measure how well the independent variables which is called as input variables account for the measured dependent variable which called as output variables. The stronger the relationship between predictions if the R^2 is higher.

2.5 KNIME Analytics Platform

In the early year of 2004, Nycomed Chair for Bioinformatics and Information Mining developed the Konstanz Information Miner namely as KNIME at the University of Konstanz. Since 2010, KNIME has become open source and free to use because it has been licensed under GPLv3. When KNIME published first edition in 2006, it was used by several pharmaceutical companies and soon afterwards, software vendors began building tools based on KNIME. KNIME can replace writing a programming code with a visual programming, to use the drag-and drop functioning nodes to build the workflow. Every node performs algorithm-based data processing and is able to interconnect with other nodes, allowing the complex data processing workflows to be generated and documented. Node is the smallest processing unit in KNIME.

3. Methodology

This chapter is about methodology that discussed the framework of this study to construct the machine learning model in this study. The steps include data collection, data preparation, data pre-processing, train and apply machine learning model, model evaluation, make prediction, data analysis and discussion, recommendation and conclusion. The data set will be collected from published literature and focused on parameters that influence concrete's compressive strength. A data collection of 191 mixes was collected from the literature containing the composition of the mixture with comparable physical and chemical properties The data set need to separate into two partitions: training set and testing set, the ratio of training set is 80% and testing set is 20%. Data set consist of 34 input variables (cement, water aggregate and et cetera) and only one output variables (28 days compressive strength). KNIME software will be used in this study to construct a machine learning algorithm without writing the programming code. To set up the workflow, just using drag and drop feature nodes just like a pipelining concept. Figure 3 shown the framework of methodology in this study.

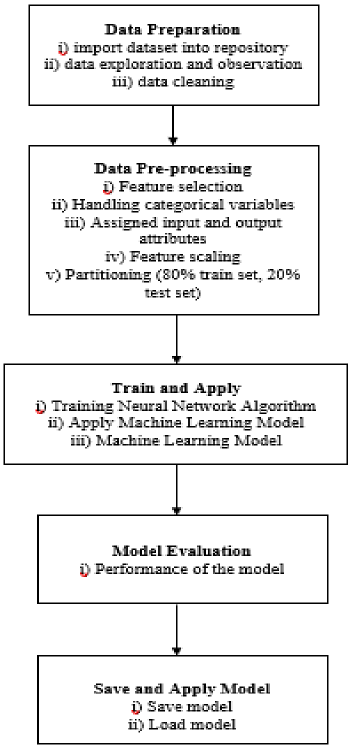

Figure 1: Framework of methodology

4. Results and Discussion

This chapter discusses the findings of artificial neural network modelling framework as mentioned previously.

4.1 Data Exploration

The following image is a statistical table in Knime, providing all the statistical values such as mean, median, standard deviation, maximum value, minimum value, variance and kurtosis for the variables along with the histogram which are skewed. From the statistical table, it shown that the minimum value of output variable is 18 MPa and its maximum value is 57.7 MPa. The mean and median value are 35.273 MPa and 34.730 MPa. It has low standard deviation which is 8.513 MPa. From the histogram above, most of the data are on the left side of the histogram but also a few larger values are on the right, this can be said the data to be skewed to the right. It means the mean is larger than the median.

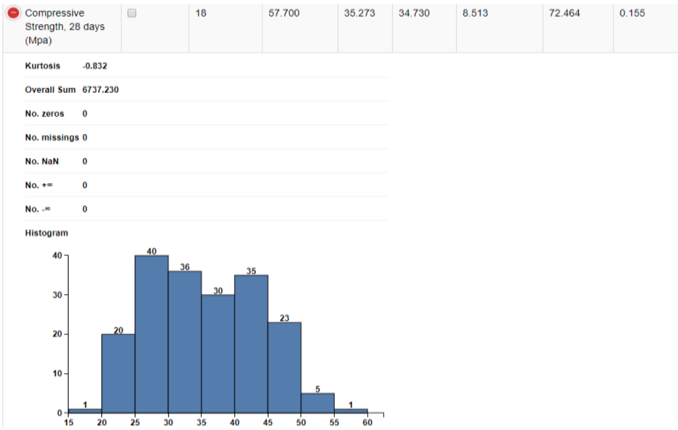

Figure 2: Statistical table

Outlier determination is must during construct a machine learning model. Below figure show the box plot view for output variables, compressive strength. There is not outlier for this variables.

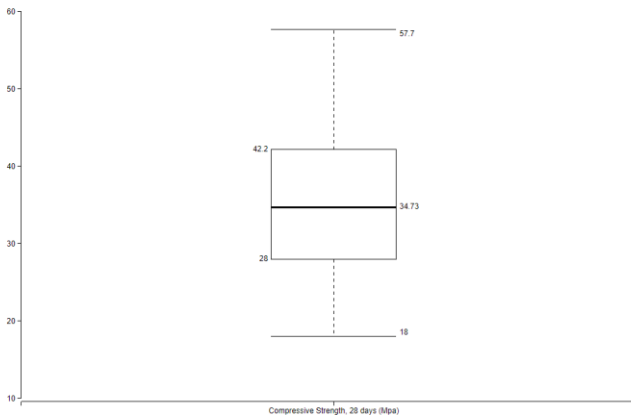

Figure 3: Box plot view for compressive strength, 28 days (MPa)

4.2 Data-preprocessing

In order to avoid multicollinearity effect, identify which features that are highly correlated to each other is very important. Figure 5 shown the Rank Correlation between all the variables. For those R value higher than 0.9, one of it will be excluded from the selection to train the algorithm. Figure 6 shown the variable that excluded from the selection to train the algorithm using the column filter node.

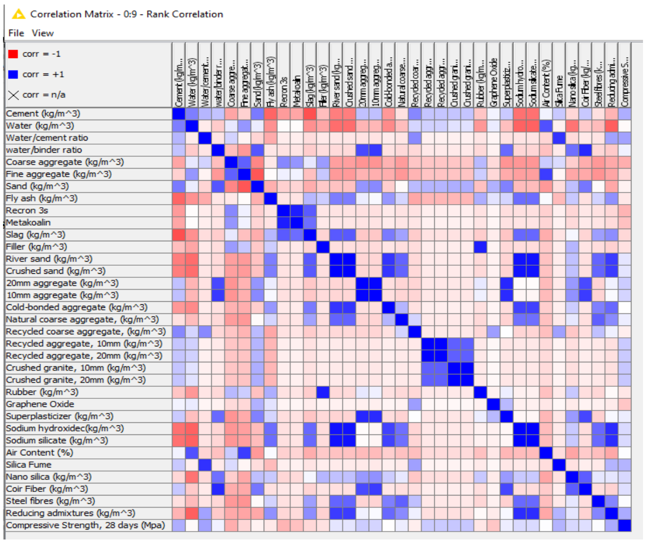

Figure 4: Rank correlation between variables

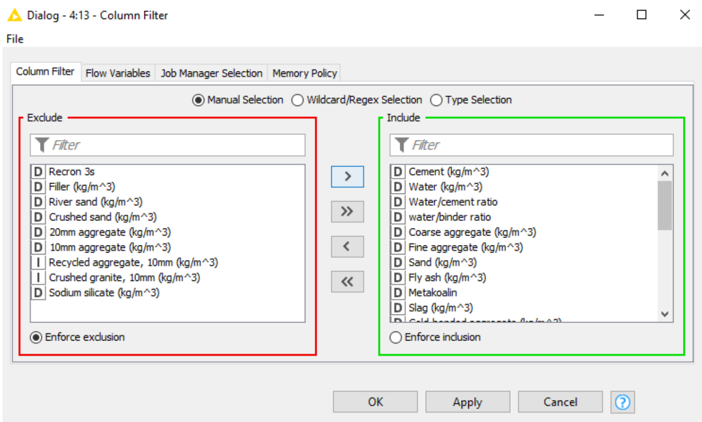

Figure 5: Variables Excluded

Normalize and partitioning the data set into 80% of train set and 20% of test set. After data are prepared and transform properly, it is ready to be used to train machine learning algorithm. Below shown the result of R-Squared with different number of layer.

Table 1: Results regarding the neural network models based on number of hidden layers

No. of hidden layer	R-Squared, R^2	Mean Squared Error, MSE	Root Mean Squared Error, RMSE
1	0.813	0.006	0.080
2	0.827	0.006	0.077
3	0.808	0.007	0.081
4	0.757	0.008	0.091
5	0.838	0.006	0.074
6	0.715	0.010	0.099
7	0.783	0.007	0.086
8	0.684	0.011	0.104
9	0.555	0.015	0.123
10	-0.059	0.036	0.190

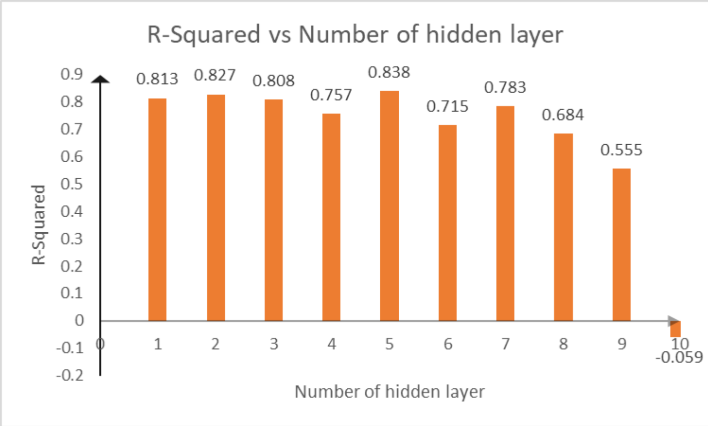

Figure 6: Bar chart for number of hidden layer and R-Squared

Table 2: Results regarding the neural network models based on number of hidden neurons per layers

No. of hidden neurons per layer	R-Squared, R^2	Mean Squared Error, MSE	Root Mean Squared Error, RMSE
2	0.097	0.031	0.176
4	0.393	0.021	0.144
6	0.574	0.015	0.121
8	0.580	0.014	0.120
10	0.789	0.007	0.085
12	0.700	0.01	0.101
14	0.838	0.006	0.074
16	0.749	0.009	0.093
18	0.738	0.009	0.095
20	0.816	0.006	0.079
22	0.783	0.007	0.086
24	0.799	0.007	0.083
26	0.655	0.012	0.109
28	0.742	0.009	0.094
30	0.732	0.009	0.096

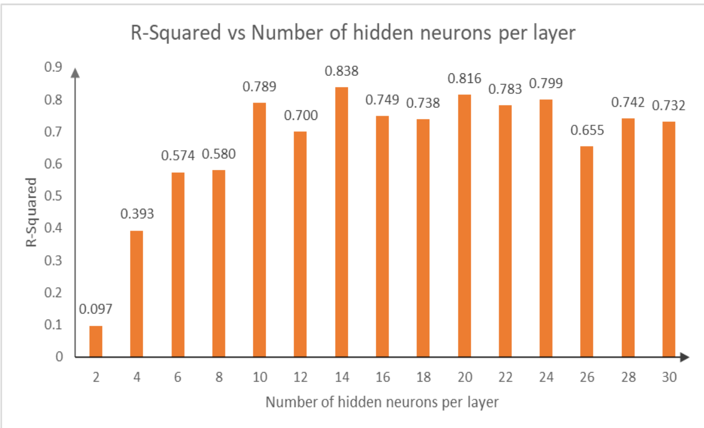

Figure 7: Bar chart for number of hidden neurons per layer and R-Squared

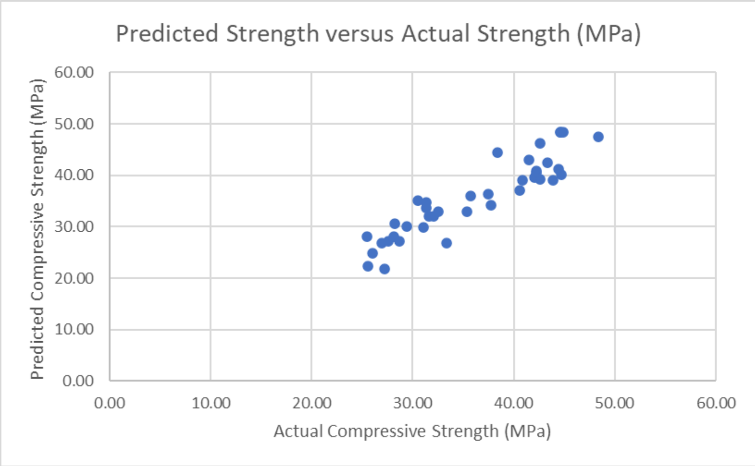

Figure 8: Predicted compressive strength versus actual compressive strength (MPa)

Table 1 and Figure 6 shown the result of the models based on number of hidden layers. As can be observed the performance of model was improved as the number of hidden layers increased until number 2. During the number of hidden layers at 3, the R-Squared value was decreased. Increasing the number of hidden layers up to 5 will increase the performance of the model which can be observed by the R-Squared. However, the performance started to drop once the number of hidden layers was increasing. This is due to smaller number of hidden layers will make the neural network less complex and more efficient as reported by [15] . R-Squared was gradually decreased from the number of hidden layers at 7 until 10. As can be found from the table below the number of hidden layers at 10 shown the negative value of R-Squared. According to [16], the negative R-Squared value indicate that the forecast appears to be less accurate than the average value of the dataset over time. Thus, the neural network architecture with 5 hidden layers was chosen as the best neural network architecture.

Based on Table 2 and Figure 7, the performance of model was improved as the number of hidden neurons per layer increased until 10 which can be observed by using R-Squared value. 14 hidden neurons per layers performed the best neural network architecture. R-Squared was decreased again when the number of hidden neurons per layer at 16 and increased at 20 hidden neuron per layer. According to [17], too many neurons in the hidden layer can cause over-fitting problems, resulting in good network learning and memorization of data, but lacking the ability to generalize. Figure 8 shown the relationship between the predicted and actual compressive strength. The predicted strength obtained from the model is quite close for actual strength. It indicated a quite strong correlation between input and output parameter of this model.

4. Conclusion

Based on the results, the scope has been concluded. Dataset of compressive strength at 28 days (MPa) was collected from 11 published journal and shown in the Chapter 3. It consists of 191 rows of data and 35 columns with various types of variables. It consists of 34 input variables such as cement, water, different types of aggregates and et cetera. Output variables only consist of 28 days compressive strength result. All the experimental data used to estimate the compressive strength of the concrete was taken from the research work that carried out by a few researchers.

For this study, the artificial neural network algorithm is used for the 28-day prediction of concrete compressive strength. 191 set of data were collected and test variables were split into 34 input and 1 output (compressive strength). The dataset was split into training set and testing set. Testing set is use to evaluate the performance of the generated model. The model is created by the Knime. The algorithm will be trained to predict output based on features in training dataset. The machine learning model was generated by following the several process: data preparation, data pre-processing, algorithm processing and model validation. Data preparation include import data into repository, data exploration and data cleaning. During data pre-processing, the step needs to do is feature selection, handling categorical variables if has, assigned input and output attributes, feature scaling and also partitioning to split the dataset into training and testing set. Once the dataset was well prepared and transformed, it is ready to be used to train the neural network algorithm. The target output is 28 days compressive strength. After the model is trained, the performance of trained model will be evaluated.

For the generated model, 5 hidden layer and 14 hidden neurons per layers performed the best neural network architecture in this study. The R-squared value for the generated model is 0.838, it is considered high for the model. The average or mean absolute error of the model is 0.061. This mean the absolute differences between predicted compressive strength and actual compressive strength will be in between ±0.061. It was concluded that the neural network model indicates the good accuracy of the prediction for the 28 days of compressive strength. With the help of this model, we can efficiently get the different mix proportion of concrete at different time and no need to produce the large number of concrete samples in lab. On the other hand, it can also be used to verify if the designed proportion of the design mix can fulfil the requirement for target strength.

Acknowledgement

The authors would like to thank the Faculty of Civil Engineering and Built Environment, University Tun Hussein Onn Malaysia for its support.

References

[1] Naderpour, H., Rafiean, A. H., & Fakharian, P. (2018). Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *Journal of Building Engineering*, *16*(January), 213–219. https://doi.org/10.1016/j.jobe.2018.01.007

[2] Al-Shamiri, A. K., Kim, J. H., Yuan, T. F., & Yoon, Y. S. (2019). Modeling the compressive strength of high-strength concrete: An extreme learning approach. *Construction and Building Materials*, *208*, 204–219. https://doi.org/10.1016/j.conbuildmat.2019.02.165

[3] Siddique, R., Aggarwal, P., & Aggarwal, Y. (2011). Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks. *Advances in Engineering Software*, *42*(10), 780–786. https://doi.org/10.1016/j.advengsoft.2011.05.016

[4] D’Alessandro, A., Fabiani, C., Pisello, A. L., Ubertini, F., Materazzi, A. L., & Cotana, F. (2016). Innovative concretes for low-carbon constructions: a review. *International Journal of Low-Carbon Technologies,* (October 2016), 289–309. https://doi.org/10.1093/ijlct/ctw013

[5] Adamu, M., Mohammed, B. S., & Shahir Liew, M. (2018). Mechanical properties and performance of high volume fly ash roller compacted concrete containing crumb rubber and nano silica. *Construction and Building Materials*, *171*, 521–538. https://doi.org/10.1016/j.conbuildmat.2018.03.138

[6] Deepa, C., SathiyaKumari, K., & Sudha, V. P. (2010). Prediction of the Compressive Strength of High Performance Concrete Mix using Tree Based Modeling. *International Journal of Computer Applications*, *6*(5), 18–24. https://doi.org/10.5120/1076-1406

[7] Alilou, V. K., & Mohammad, T. (2011). Prediction of 28-day compressive strength of concrete on the third day using artificial neural networks. *International Journal of Engineering*, *3*(6), 565–576.

[8] M., G., A., T., C., S., & P., R. (2013). Soybean Under Water Deficit: Physiological and Yield Responses. *A Comprehensive Survey of International Soybean Research - Genetics, Physiology, Agronomy and Nitrogen Relationships*. https://doi.org/10.5772/54269

[9] Salehi, H., & Burgueño, R. (2018). Emerging artificial intelligence methods in structural

engineering. *Engineering Structures*, *171*(November 2017), 170–189. https://doi.org/10.1016/j.engstruct.2018.05.084

[10] Vieira, S., Hugo, W., Pinaya, L., Mechelli, A., & Andre, S. (2020). learning. In *Machine Learning*. https://doi.org/10.1016/B978-0-12-815739-8.00001-8

[11] Tiryaki, S., & Aydin, A. (2014). An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. *Construction and Building Materials*, *62*, 102–108. https://doi.org/10.1016/j.conbuildmat.2014.03.041

[12] Feng, D. C., Liu, Z. T., Wang, X. D., Chen, Y., Chang, J. Q., Wei, D. F., & Jiang, Z. M. (2020). Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Construction and Building Materials*, *230*, 117000. https://doi.org/10.1016/j.conbuildmat.2019.117000

[13] Khademi, F., Akbari, M., Jamal, S. M., & Nikoo, M. (2017). Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete. *Frontiers of Structural and Civil Engineering*, *11*(1), 90–99. https://doi.org/10.1007/s11709-016-0363-9

[14] Yeh, I. C. (1998). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*. https://doi.org/10.1016/S0008-8846(98)00165-3

[15] Zhou, Y., Zheng, S., Huang, Z., Sui, L., & Chen, Y. (2020). Explicit neural network model for predicting FRP-concrete interfacial bond strength based on a large database. *Composite Structures*, *240*(January), 111998. https://doi.org/10.1016/j.compstruct.2020.111998

[16] Nerdy, F. (2017). What Is R Squared And Negative R Squared. Retrieved from http://www.fairlynerdy.com/what-is-r-squared/

[17] Rohani, A., Abbaspour-Fard, M. H., & Abdolahpour, S. (2011). Prediction of tractor repair and maintenance costs using Artificial Neural Network. *Expert Systems with Applications*, *38*(7), 8999–9007. https://doi.org/10.1016/j.eswa.2011.01.118