Journal of Advanced Engineering Trends
ISSN : 2682 - 2091

Vol.43, No.1. January 2024

http://jaet.journals.ekb.eg

# ESTIMATING THE CONCRETE COMPRESSIVE STRENGTH USING ARTIFICIAL INTELLIGENCE: PROOF OF CONCEPT

**Alaa Eldin Y. Abouelezz, [(1)] and Ashraf Amin [(2)]**

Associate Professor , Minia University[(1)] Instructor Civil engineering department, faculty of engineering, Minia, Egypt[(2)]

***Abstract*** *- Most structure members used nowadays and, in the future contain concrete, so it is crucial to raise the accuracy of forecasting the concrete strength all time to make the best use of it. Concrete is a non-homogeneous material that consists of various materials, each one has its unique physical properties. Based on the contribution percentage for each component of the concrete mixture, we get a different concrete compressive strength for the produced concrete. Experimental mixtures in the lab are a traditional method to design the percentages of each concrete components based on approximate relationships to achieve the required compressive strength of the concrete. This method consumes much time and wasted materials used in the experiments. Using different artificial intelligence techniques to predict the concrete strength is considered fast, inexpensive, and more accurate, so this advanced technique is in continues development which may help us to solve many engineering problems in the future. In this paper, one of the artificial intelligence techniques is applied to predict the concrete behavior under the effect of changing the contribution of the components in the mixture. In addition, some recommendations for the range of inputs make the program optimize the results. New inputs can be taken into consideration like specifying new components' properties and their effect on the resulted strength of the concrete, demonstrating that artificial intelligence can be used to predict the concrete strength.*

**Keywords**: Compressive strength, Predictive techniques, Artificial intelligence algorithms, Neural network

## INTRODUCTION

Concrete consists of some commonly used materials like aggregates, water, and cement as binder material are mixed together with adequate percentages to attain the required hardened concrete strength [1], and it is the most used material in the history of construction industry [2]. In concrete mixture design and quality control, concrete compressive strength value is the most important indicator for the quality of mixture [3]. However, many factors influence the concrete strength including material properties, ingredients percentages, curing time, environmental conditions, and concrete age [4]. High nonlinearity relationship between components and the resultant concrete characteristics complicates the mathematical modeling of compressive strength when using available data [5, 6]. A prediction model can reduce time and cost of a project, because it provides designers and the structural engineers with essential data like concrete properties which can help for better planning to the project [7, 8]. In recent years, several Artificial Intelligence (AI) methods that have been implemented to predict the properties of concretes include multiple linear regression (MLR) [9], artificial neural network (ANN) [10, 11], support vector Regression (SVR) [12, 13], adaptive neuro-fuzzy inference systems (ANFIS) [14, 15], model tree (MT) [16], genetic expression programming (GEP) [17] and multivariate adaptive regression splines (MARS) [18]. The researchers observed that the artificial intelligence (AI) methods could predict more than single properties with high precision and accuracy [19]. The software can execute the assigned task and provide the intended output using the actual dataset. During the computing, the central logic programmed in the software takes the decision, processes the data, and transfers the raw data into meaningful information [20]. The AI models are being extensively utilized in almost any field, including medical, accounting, education, manufacturing, designing, engineering, and construction and management. However, the AI methods include the conventional methods directly or indirectly. AI has become a versatile and efficient tool for research and innovation in the past few years [21]. The general approaches for forecasting concrete mechanical characteristics include traditionally nonlinear and linear static analyses and various statistical methods [22]. Statistical approaches of test design are crucial for assessing the characteristics and optimizing the concrete combinations with the desired strength.

Revised:15 June, 2022, Accepted:7 December , 2022

These approaches are usually utilized in companies for optimizing production and improving concrete performance in certain research projects [23]. For example, regression analysis looks easy and unconscious and certain weaknesses contribute to reduced inaccuracy when dealing with multiple independent variables [24]. There are high connections between the empirical data and the predicting of ANN for the characteristics examined [25]. González-Taboada et al. used genetic expression programming (GEP) to predict the mechanical properties of structural recycled concrete with high accuracy [26]. In AI techniques, physical interactions between the inputs and outputs are not considered. This is a drawback of these AI models, in which the input variables of predictive data must fall within the training data's range. Even though AI techniques seem to have a lot of potential, their performance is limited by the fact that they depend on the initial tuning parameters they are used on [27]. The detection of optimal values for parameters is a challenge procedure in AI techniques, which may seem like an optimization problem [28].

## II. Data collection

Based on previous trial mixes, 42 ordinary concrete mixes consisting of gravel table (1), sand table (2), cement, and water without any additives were tested under compression, and the concrete strength of each mix was recorded and tabulated with its gradients. Table (3) shows these mixes.

**Table 2 Properties of the coarse aggregates**

| Property type | Result |
|---|---|
| Maximum nominal size mm | 20 |
| Specific gravity | 2.71 |
| Void ratio | 36.5 % |
| Volume weight gm./cm$^3$ | 1.62 |
| Fineness modulus | 6.07 |

**Table 1 Properties of the fine aggregates**

| Property type | Result |
|---|---|
| Specific gravity | 2.55 |
| Void ratio | 41.5 % |
| Volume weight gm./cm$^3$ | 1.52 |
| Fineness modulus | 2.93 |
| Dust and fine materials | 2.50 % |

**Table 3 Concrete mixes used in modeling**

| Mixture No. | Gravel m$^3$ | Sand m$^3$ | Cement Kg | W/C | Concrete grade kg/cm$^2$ |
|---|---|---|---|---|---|
| 1 | 0.878 | 0.468 | 213 | 0.6 | 186 |
| 2 | 0.89 | .456 | 213 | 0.6 | 200 |
| 3 | 0.7 | 0.37 | 389 | 0.6 | 217 |
| 4 | 0.854 | 0.455 | 266 | 0.5 | 244 |
| 5 | 0.866 | 0.444 | 266 | 0.5 | 258 |
| 6 | 0.877 | 0.432 | 266 | 0.5 | 268 |
| 7 | 0.751 | 0.4 | 381 | 0.5 | 320 |
| 8 | 0.763 | 0.389 | 381 | 0.5 | 333 |

| | | | | | |
|---|---|---|---|---|---|
| 9 | 0.741 | 0.395 | 409 | 0.46 | 362 |
| 10 | 0.754 | 0.366 | 404 | 0.5 | 342 |
| 11 | 0.765 | 0.354 | 404 | 0.5 | 348 |
| 12 | 0.733 | 0.39 | 458 | 0.4 | 372 |
| 13 | 0.741 | 0.38 | 460 | 0.4 | 379 |
| 14 | 0.752 | 0.37 | 460 | 0.4 | 387 |
| 15 | 0.737 | 0.39 | 526 | 0.3 | 402 |
| 16 | 0.76 | 0.37 | 530 | 0.3 | 405 |
| 17 | 0.77 | 0.36 | 527 | 0.3 | 407 |
| 18 | 0.845 | .485 | 192 | 0.75 | 135 |
| 19 | 0.87 | 0.46 | 193 | 0.75 | 143 |
| 20 | 0.878 | 0.45 | 193 | 0.75 | 152 |
| 21 | 0.828 | 0.475 | 240 | 0.6 | 185 |
| 22 | 0.84 | 0.464 | 241 | 0.6 | 192 |
| 23 | 0.85 | 0.45 | 241 | 0.6 | 205 |
| 24 | 0.808 | 0.463 | 276 | 0.55 | 238 |
| 25 | 0.824 | 0.445 | 278 | 0.55 | 243 |
| 26 | 0.834 | 0.434 | 278 | 0.55 | 254 |
| 27 | 0.79 | 0.438 | 324 | 0.5 | 292 |
| 28 | 0.79 | 0.438 | 325 | 0.5 | 301 |
| 29 | 0.802 | 0.427 | 325 | 0.5 | 307 |
| 30 | 0.755 | 0.433 | 470 | 0.35 | 360 |
| 31 | 0.886 | 0.477 | 472 | 0.35 | 368 |
| 32 | 0.774 | 0.413 | 470 | 0.35 | 375 |
| 33 | 0.811 | 0.32 | 411 | 0.42 | 400 |
| 34 | 0.82 | 0.437 | 324 | 0.45 | 250 |
| 35 | 0.796 | 0.424 | 331 | 0.5 | 202 |
| 36 | 0.789 | 0.42 | 356 | 0.46 | 227 |
| 37 | 0.782 | 0.417 | 380 | 0.43 | 256 |
| 38 | 0.769 | 0.409 | 427 | 0.375 | 303 |
| 39 | 0.761 | 0.405 | 422 | 0.39 | 280 |
| 40 | 0.745 | 0.397 | 414 | 0.45 | 267 |
| 41 | 0.73 | 0.389 | 405 | 0.49 | 247 |
| 42 | 0.701 | 0.374 | 389 | 0.6 | 217 |

## III. Data augmentation

Techniques used to increase the amount of data in data analysis include adding slightly modified copies of already existing data or crating new synthetic datasets from existing datasets. It acts as a stabiliser and helps to reduce overfitting when training a machine learning model, because a lot of data is required to make an accurate regression model to get a good relationship between ingredient variables and the corresponding resultant of concrete compressive strength. Because the data was increased by multiplying component weights for all mixes by factors ranging from 2 to 60 and then fixing the concrete grade for each trial mix, we were able to get the model to be more sensitive and produce the best results. The following two conditions were used to insert data into some trial mixes designed by the British code: the first one, w/c ratio, started from a value and then increased slightly to a specified ratio where the other inputs were fixed, the second condition was to increase cement content for the previous mix to a specified weight where the other inputs were fixed. This method improved the results of the model by letting it know how important these factors are and how they might affect the strength of the concrete. The dataset is divided into two categories: training dataset (which accounts for approximately 70% of all data) and test dataset (which accounts for approximately 30% of all data), 2288 samples were used after drop duplicates of concrete grade which were 38 samples.

## IV. Regression models

### *A. Linear regression*

Using linear regression algorithms in scikit learn library in python which try to find the best fit line that can go through these data and once I have it that means which called the simple linear regression model Fig. (1). In that model that can be relied upon it in the future to predict the concrete strength.

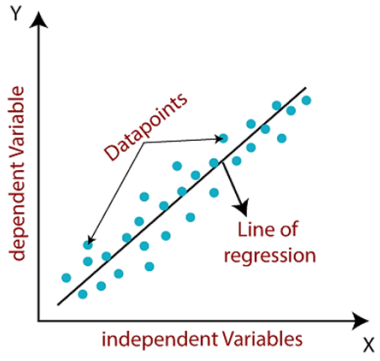


**Figure 1 linear regression model**

### *B. Ridge* regression

Ridge Regression is a linear regression extension that penalises the loss function during training by adding a regularisation penalty. A linear relationship is assumed between the input variables and the target variable in a regression model known as a linear regression. This relationship can be visualised as a hyperplane that connects the input variables to the target variable when there is only one input variable. Through an optimization process that aims to minimise the sum squared error between predictions and expected target values, the model's coefficients are discovered.

$$loss = \sum_{i=0}^{i=n}((expected\ target\ value)^i - (predictions)^i)^2 \quad (1)$$

Model aim is to obtain basically a regression algorithm that could generalize patterns which means works better on training data and testing data as well. Overfitting occurs when the model performs well on the training data and poorly on the testing dataset, so the authors are going to reduce the variability among two datasets by trying to increase the bias a little bit between the training dataset and the equation of the regression model to improve variance and this is how ridge regression works.

### *C. Lasso regression*

Regularized linear regression with an L1 penalty is known as Lasso Regression. Those input variables that don't have much bearing on the prediction process see their coefficients shrink as a result. Some coefficient values can be penalised, allowing input variables to be removed from the model and providing a form of automatic feature selection. Change the loss function to include additional costs for models with large coefficients as a way to address regression model stability. The term "penalised linear regression" refers to linear regression models that use these modified loss functions during training. The slope of the equation line can be reduced to zero using lasso regression rather than ridge regression if we have several independent variables that are useless.

$$least\ square\ regression: Min.(sum.of\ the\ squared\ residuals) \quad (2)$$

$$Ridge\ regression: Min.(sum.of\ squared\ residuals + \alpha(slope)^2) \quad (3)$$

$$lasso\ regression: Min.(sum.of\ squared\ residuals + \alpha[slope]) \quad (4)$$

Where α is the value between zero point of y axis and the point of intersection for equation line.

### *D. Decision tree regression*

Machine learning technique decision trees divide data into groups based on a predetermined condition. For example, in a tree algorithm, there are nodes, which represent decision points, and leaf nodes, which represent final outcomes. The goal is to determine the best way to split the dataset. Decision trees are used to make predictions by traversing the tree with the data provided in a specific row. Nonlinear functions can be used to implement this, where the same prediction routine is called with the left or right infant nodes, depending on how the split affects provided data. Nodes in the tree must be checked to see whether they are the final prediction or a dictionary node that contains another level of the tree.

### *E. Random forest regression*

Among machine learning algorithms, Random Forest is one of the most popular and powerful. Bootstrap Aggregation, or bagging, is a type of ensemble machine learning algorithm. Classification and regression problems can be solved using random forest algorithms. It is common practise to create a number of decision trees based on a bootstrap sample of the training dataset for each tree. If an example appears more than once in the training dataset, the sample is said to be a bootstrap sample. We call this "sampling with replacement" in our industry. Using a variety of decision trees, each with a different set of training data, makes Bagging an efficient ensemble algorithm. Because the trees used in the ensemble are unpruned, they are slightly overfit to the training dataset, unlike normal decision tree models such as classification and regression trees. This is a good thing because it makes each tree unique and less likely to have correlated predictions or errors. An average of predictions from all decision trees yields a better model performance than any single tree in it. When solving a regression problem, an ensemble's prediction is its average across all of the trees within it.

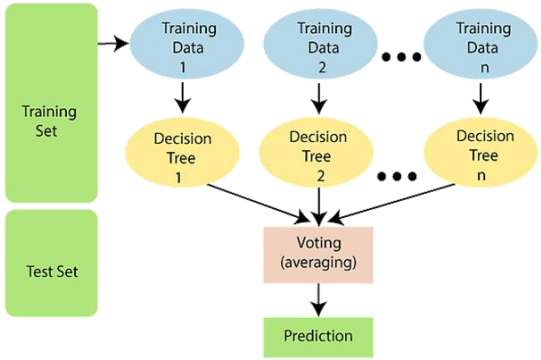


**Figure 2 Methodology of random forest algorithm**

### *F. Artificial neural network (ANN)*

The stochastic gradient descent optimization algorithm is used to fit deep learning neural network models to training data. Backpropagation of error is used to update the weights of the model. Optimization and weight update algorithms were carefully selected to fit neural networks, making this combination the most efficient method available. Even so, fitting a neural network model to a training dataset can be accomplished using different optimization algorithms. This is a good way to learn more about how neural networks work and the importance of optimization in applied machine learning by conducting this experiment. It may also be needed for neural networks with transfer functions that can't be changed and models with unusual architectures. Deep learning, or neural networks, is a flexible type of machine learning. They are models composed of nodes and layers inspired by the structure and function of the brain. A neural network model works by propagating a given input vector through one or more layers to produce a numeric output that can be interpreted for classification or regression predictive modelling. Models are trained by repeatedly exposing the model to examples of input and output and adjusting the weights to minimize the error of the model's output compared to the expected output. This is called the stochastic gradient descent optimization algorithm. The weights of the model are adjusted using a specific rule from calculus that assigns error proportionally to each weight in the network. This is called the backpropagation algorithm; the tensor flow library is used in Python for deep learning. The input layer dimension was 4, and 3 hidden layers were used. The first one contained 64 neurons, the second layer 32 neurons, and the third layer was 8 neurons. The output layer is one neuron, which represents the concrete grade. The activation function used was relu.

Compiling with: loss='mean_squared_error'

Optimizer: Adam

Learning rate = 1e-4

Train data = 70%

Test data = 30%

Dataset = 2288 samples

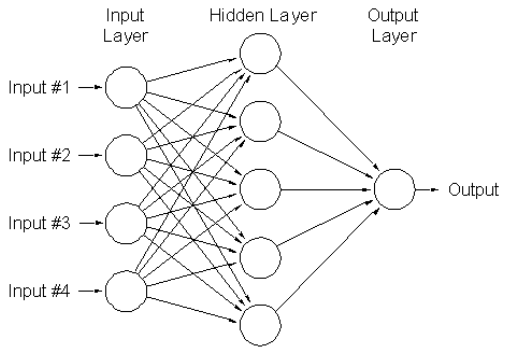


Figure 3 Neural network structure

## V. Results and Discussions

Table (4) and Fig. (4) show the accuracy of each regression model based on the validation between the training dataset and the testing dataset after the processing.

Table 4 Accuracy for the predictive methods.

| Regression type | Mean Absolute Error | Mean Squared Error | Median Absolute Error | R-Square |
|---|---|---|---|---|
| Linear regression | 59.44 | 2256.38 | 31.89 | 52.11% |
| Lasso regression | 55.29 | 2689.81 | 33.09 | 50.24% |
| Ridge regression | 53.6 | 2407.43 | 33.61 | 51.27% |
| Decision tree | 85.18 | 777.43 | 9.0 | 82.32% |
| Random forest | 90.18 | 427.03 | 8.92 | 91.19% |
| ANN | 99 | 63.28 | 3.92 | 98.58% |

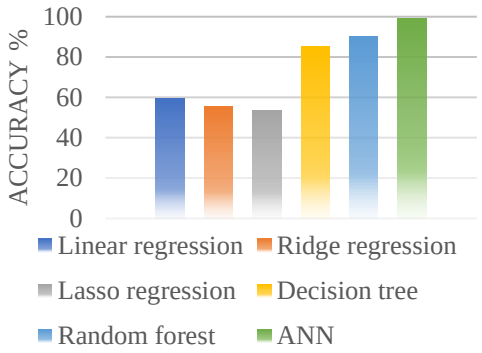


Figure 4 Accuracy for the predictive models

The correlation between actual concrete grades and the predicted ones for linear, ridge, and lasso regression models is shown in figures (5, 6, and 7), respectively. It can be observed there is a huge difference between the actual and the predicted data in these methodologies and no fit line can approximately pass through the resulted data because of the high nonlinearity between ingredient percentages and the concrete grade, and these prediction methods depend mainly on linear relationships, so it is not suitable for our investigation.

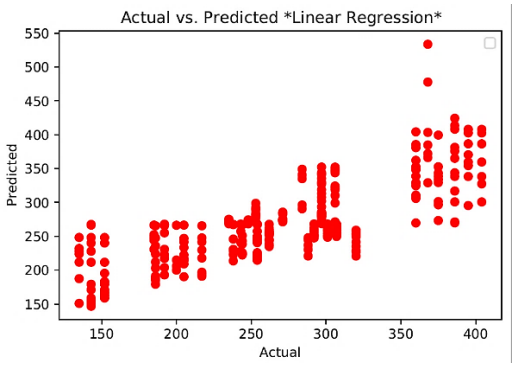


Figure 5 Training and testing data for linear regression

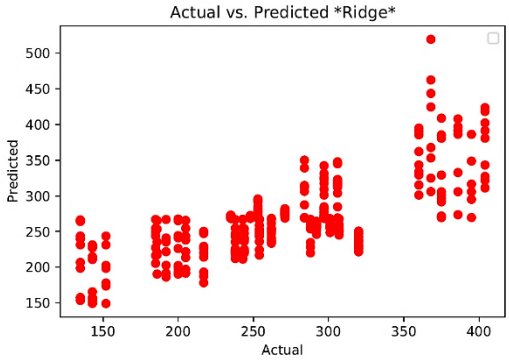


Figure 6 Training and testing data for ridge regression

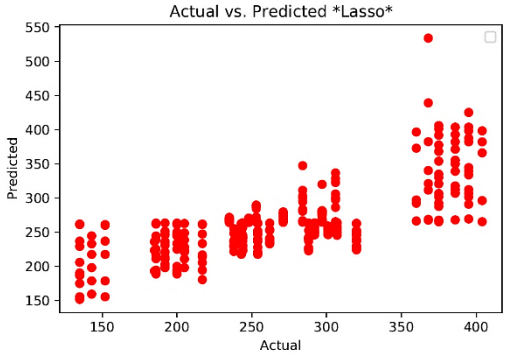


**Figure 7 Training and testing data for lasso regression**

Decision tree regression system is shown in Fig. (8), and it seems that most of the actual and predictive points can be gathered by one line, but there are some results that are not few but also not too far away from the equation line, so it shows better prediction than the linear regression models.

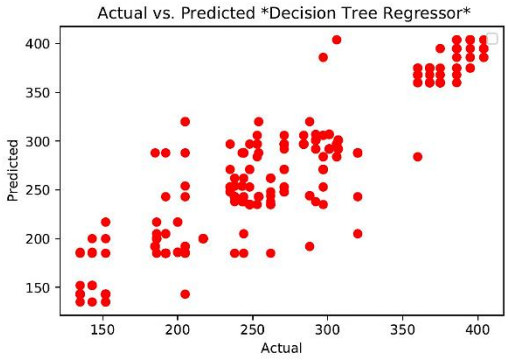


**Figure 8 Training and testing data for decision tree regression**

Random forest regression model presents more enhanced prediction results than the decision tree method because it consists of many decision trees, which make optimization to the prediction result from each tree, so most of the predictive and actual values lie near to the equation line as shown in Fig. (9).

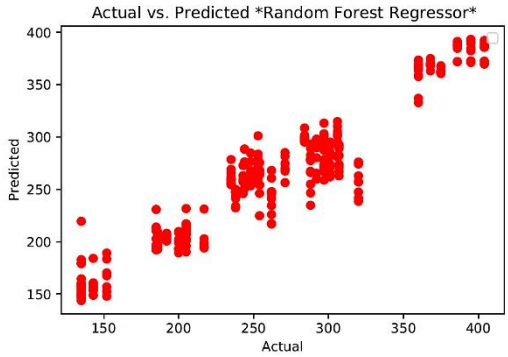


**Figure 9 Training and testing data for random forest regression**

For the ANN regression model, we just moved from machine learning for the previous methods to deep learning by building an expert system that can think and feel the sensitivity of the different variables to give an accurate result with minimum errors, as shown in Fig. (10).

The data can be gathered on one line, and this is a great indicator for the success of modelling. Fig. (11) shows the loss function of the training dataset and the testing dataset, which equals the difference between them during the processing period, because we have tried to get peak data, which helped the system to optimize itself and minimize its errors, and it can be enhanced more than now by training it more and more.

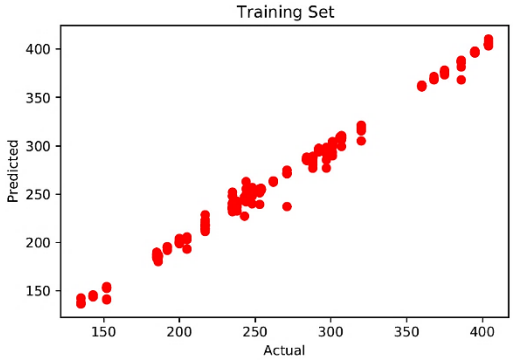


**Figure 10 Training and testing data for ANN regression**

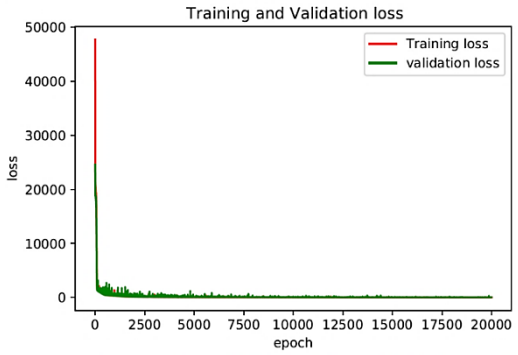


**Figure 11 Training and validation loss**

## VI. Conclusion and Recommendations

- Linear predictive models can't show how strong concrete will be exactly, and they aren't very accurate.
- Decision trees and random forests are better for classifications and not the best solutions for prediction.
- Deep learning is considered the best approach to predicting concrete strength.
- To get an expert system, more input specifications are needed to generalize the prediction to all types of concrete, whatever the type of its ingredients.
- Additives and cement replacement materials are commonly used nowadays, so it is better to involve them in the prediction models.

## VII. References

1. Bhuva, P. and A. Bhogayata, *A review on the application of artificial intelligence in the mix design optimization and development of self-compacting concrete.* Materials Today: Proceedings, 2022.
2. AzariJafari, H., et al., *Ternary blended cement: An eco-friendly alternative to improve resistivity of high-performance self-consolidating concrete against elevated temperature.* Journal of Cleaner Production, 2019. **223**: p. 575-586.
3. Chou, J.-S. and A.-D. Pham, *Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength.* Construction and Building Materials, 2013. **49**: p. 554-563.
4. Ni, H.-G. and J.-Z. Wang, *Prediction of compressive strength of concrete by neural networks.* Cement and Concrete Research, 2000. **30**(8): p. 1245-1250.
5. Yeh, I.-C., *Modeling of strength of high-performance concrete using artificial neural networks.* Cement and Concrete research, 1998. **28**(12): p. 1797-1808.
6. Kasperkiewicz, J., J. Racz, and A. Dubrawski, *HPC strength prediction using artificial neural network.* Journal of Computing in Civil Engineering, 1995. **9**(4): p. 279-284.
7. Haefner, N., et al., *Artificial intelligence and innovation management: A review, framework, and research agenda☆.* Technological Forecasting and Social Change, 2021. **162**: p. 120392.
8. Azim, I., et al. *Prediction model for compressive arch action capacity of RC frame structures under column removal scenario using gene expression programming*. in *Structures*. 2020. Elsevier.
9. Chen, L., *A multiple linear regression prediction of concrete compressive strength based on physical properties of electric arc furnace oxidizing slag.* International Journal of Applied Science and Engineering, 2010. **7**(2): p. 153-158.
10. Getahun, M.A., S.M. Shitote, and Z.C.A. Gariy, *Artificial neural network based modelling approach for strength prediction of concrete incorporating agricultural and construction wastes.* Construction and Building Materials, 2018. **190**: p. 517-525.
11. Mohammadhassani, M., et al., *Identification of a suitable ANN architecture in predicting strain in tie section of concrete deep beams.* Structural engineering and mechanics: An international journal, 2013. **46**(6): p. 853-868.
12. Sonebi, M., et al., *Modelling the fresh properties of self-compacting concrete using support vector machine approach.* Construction and Building materials, 2016. **106**: p. 55-64.
13. Yaseen, Z.M., et al., *Predicting reinforcing bar development length using polynomial chaos expansions.* Engineering Structures, 2019. **195**: p. 524-535.
14. Khotbehsara, M.M., et al., *Effect of SnO2, ZrO2, and CaCO3 nanoparticles on water transport and durability properties of self-compacting mortar containing fly ash: Experimental observations and ANFIS predictions.* Construction and Building Materials, 2018. **158**: p. 823-834.
15. Safa, M., et al., *Potential of adaptive neuro fuzzy inference system for evaluating the factors affecting steel-concrete composite beam's shear strength.* Steel and Composite Structures, An International Journal, 2016. **21**(3): p. 679-688.
16. Behnood, A., et al., *Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm.* Construction and Building Materials, 2017. **142**: p. 199-207.
17. Gholampour, A., A.H. Gandomi, and T. Ozbakkaloglu, *New formulations for mechanical properties of recycled aggregate concrete using gene expression programming.* Construction and Building Materials, 2017. **130**: p. 122-145.
18. Ashrafian, A., et al., *Prediction of compressive strength and ultrasonic pulse velocity of fiber reinforced concrete incorporating nano silica using heuristic regression methods.* Construction and Building Materials, 2018. **190**: p. 479-494.
19. Fan, D., et al., *Precise design and characteristics prediction of Ultra-High Performance Concrete (UHPC) based on artificial intelligence techniques.* Cement and Concrete Composites, 2021. **122**: p. 104171.
20. Galaz, V., et al., *Artificial intelligence, systemic risks, and sustainability*. Technology in Society, 2021. **67**: p. 101741.
21. Haque, M.A., et al., *Evaluating the mechanical strength prediction performances of fly ash-based MPC mortar with artificial intelligence approaches.* Journal of Cleaner Production, 2022. **355**: p. 131815.
22. Zhao, R., et al., *Unveiling substitution preference of chromium ions in sulphoaluminate cement clinker phases.* Composites Part B: Engineering, 2021. **222**: p. 109092.
23. Mohammadhassani, M., et al., *An evolutionary fuzzy modelling approach and comparison of different methods for shear strength prediction of high-strength concrete beams without stirrups.* Smart Struct. Syst, 2014. **14**(5): p. 785-809.
24. Shariati, M., et al., *Estimation of moment and rotation of steel rack connections using extreme learning machine.* Steel and Composite Structures, 2019. **31**(5): p. 427-435.
25. Ling, Y., et al., *Prediction of engineering properties of fly ash-based geopolymer using artificial neural networks.* Neural Computing and Applications, 2021. **33**(1): p. 85-105.
26. González-Taboada, I., et al., *Prediction of the mechanical properties of structural recycled concrete using multivariable regression and genetic programming.* Construction and Building Materials, 2016. **106**: p. 480-499.
27. Lee, T.-S., et al., *Mining the customer credit using classification and regression tree and multivariate adaptive regression splines.* Computational Statistics & Data Analysis, 2006. **50**(4): p. 1113-1130.

28. Ashrafian, A., et al., *Compressive strength of Foamed Cellular Lightweight Concrete simulation: New development of hybrid artificial intelligence model.* Construction and Building Materials, 2020. **230**: p. 117048.