Prediction of 28-day Compressive Strength of Concrete at the Job Site using Artificial Neural Network

Mary Joanne C. Aniñon* and Elizabeth Edan M. Albiento
College of Engineering and Technology
Mindanao State University – Iligan Institute of Technology
Iligan City, 9200 Philippines
*maryjoanne.aninon@g.msuiit.edu.ph

Abstract

Recently, there has been a great interest in applying artificial neural networks (ANNs) to predict the compressive strength of concrete in various compositions. This study aimed to predict the 28-day compressive strength of concrete delivered at the job site using ANN. The datasets used to construct, train and test the ANN model were obtained experimentally by the authors. Feature importance analysis was applied to evaluate the significance of input variables on the output variable. Feature selection was employed to eliminate the least relevant features based on the importance scores to improve the model prediction performance. The results demonstrated that the ANN model could predict the 28-day compressive strength of delivered concrete with high accuracy and robustness. It was also indicated that the ANN model with feature selection outperformed the ANN model without feature selection. The R values of the ANN model with feature selection were increased by 0.76 and 1.69% in training and testing sets, respectively, compared with the model without feature selection. Furthermore, it was found that the MSE values for training and testing sets were decreased by 0.8381 and 1.8882 MPa, respectively. This study revealed that the C/A ratio was the most influential feature of the compressive strength of delivered concrete followed by the FA/CA ratio, ER, W/C ratio, slump and temperature.

1. Introduction

Concrete is a frequently used construction building material because of its good mechanical properties, versatility and availability (Xu *et al*., 2021). The compressive strength of concrete is often considered a notable structural design element. Generally, it is attained through a compression test after 28 days of curing the concrete specimen, which is tedious, cumbersome and

expensive. Thus, to ameliorate this situation, several studies were performed to forecast the compressive strength of concrete using conventional statistical methods and machine learning techniques (Muliauwan *et al*., 2020; Silva *et al*., 2020; Xu *et al*., 2021).

The relationship between compressive strength and the components of concrete is nonlinear and complex; therefore, it can be challenging to determine the compressive strength of concrete using conventional statistical methods such as empirical formula (Pengcheng *et al*., 2020). Lately, the compressive strength of concrete is determined using machine learning approaches such as artificial neural network (ANN) (Muliauwan *et al*., 2020; Lin and Wu, 2021), support vector machine (SVM) (Silva *et al*., 2020; Ren *et al*., 2021) and random forest (Pengcheng *et al*., 2020). These studies have shown that machine learning-based models are more feasible than traditional statistical approaches.

The abovementioned studies have shown convincing results in predicting the compressive strength of concrete. However, the concrete delivered and placed on an actual construction site is exposed to highly variable conditions. The water content and water to cement (W/C) ratio may range because of various environmental situations and different uncertainties encountered all through the mixing and transportation of concrete (Derousseau *et al*., 2019). Hence, when used as input variables to any machine learning model, the model predicts the compressive strength of laboratory-sampled concrete, not the concrete delivered and placed at the job site (Derousseau *et al*., 2019). However, the abovementioned studies did not consider these possibilities when the water content and W/C ratio were used as input variables for their proposed models. Therefore, despite the high prediction accuracy of these studies, the challenge of predicting the compressive strength of delivered concrete at the job site remains relevant.

The average electrical resistivity (ER) of fresh concrete strongly correlates to the W/C ratio (Obla *et al*., 2018); hence it is used to represent the actual W/C ratio of delivered concrete. Therefore, it can be used as one of the input variables when developing a concrete compressive strength prediction model. To the best of the author's knowledge, the ER of fresh concrete has never been reported in the literature as an input variable of any machine learning method probably because of the data unavailability.

This study aimed to develop an ANN model that predicts the 28-day compressive strength of concrete delivered at the job site. The input variables

included the ER of the fresh concrete (Ω-m), slump (mm), maximum aggregate size (mm), FM of fine aggregates, temperature (°C) of concrete when delivered at the job site and W/C, FA/CA and C/A ratios. The output variable was the 28-day compressive strength of concrete. This study conducted feature importance analysis to determine which input variables or features may be the most or least relevant to the target or output variable. Feature selection was likewise performed to eliminate the least relevant features to improve the model prediction accuracy. Finally, the proposed ANN model was compared with other researchers' work.

2. Methodology

2.1 Data Collection

Datasets are necessary to construct, train and test the proposed ANN model. To obtain these data, the following activities were conducted: 1) formulation of concrete mixtures; 2) making and curing of concrete specimens; 3) conducting quality tests. A non-air-entrained normal concrete without admixtures and supplementary cementitious materials such as fly ash was used. Only one type of cement was utilized – Republic Type I Ordinary Portland cement with a specific gravity of 3.15 (Table 1).

Table 1. Materials and their inherent properties

Materials	Properties	
Cement	Republic Type I Ordinary Portland cement	
Specific gravity	3.15	
Fine Aggregate	First batch	Second batch
Specific gravity	2.51	2.42
Fineness modulus	3.06	3.33
Moisture content (%)	7.83	7.67
Absorption values (%)	3.33	5.30
	Maximum aggregate size	
Coarse aggregate (mm)	9.5	19
	(3/8 in.)	(3/4 in.)
Specific gravity	2.42	2.6
Moisture content (%)	5.28	2.98
Absorption values (%)	4.30	2.83

River sand and rounded aggregates from the Mandulog River, Philippines were used as fine aggregates and coarse aggregates, respectively. These were supplied by J-Con Material Testing Center, Iligan City. Their properties, such as specific gravity, fineness modulus, moisture content and absorption values were also provided by the supplier (Table 1). Lastly, potable water was used for casting all the concrete specimens.

The formulation of concrete mixtures was based on the absolute volume method as discussed in American Concrete Institute (ACI) 211.1-91 (ACI Committee 211, 2002). The 36 different concrete mix designs were derived based on their maximum aggregate size, slump and cement factors (Figure 1). The cement factor refers to the quantity of cement (bags) required for 1 m^3 of concrete. The details of the proportions for concrete mixtures are shown in Table 2 including the amount of cement, fine and coarse aggregates, water, slump, maximum aggregate size, air content, W/C ratio and fineness modulus of fine aggregates.
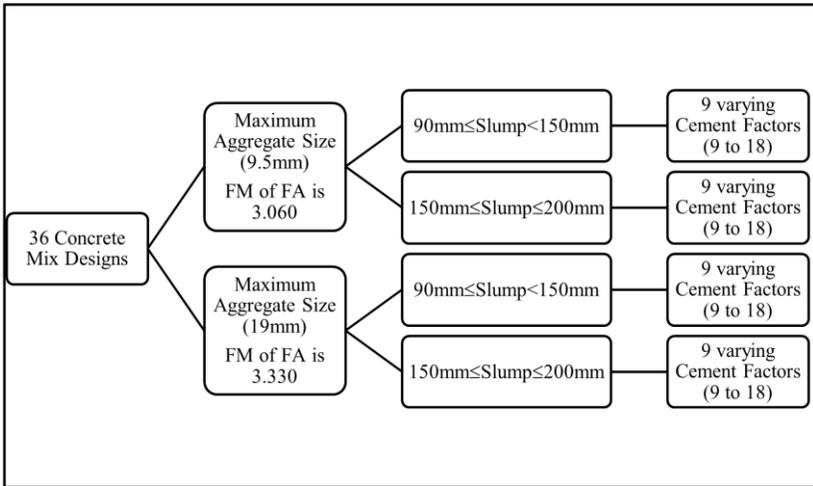
Figure 1. Derivation of 36 concrete mix designs

The mixing of concrete was based on the formulated 36 different concrete mix designs. The slump of concrete was measured following ASTM C143/C143M-20 (2020) to check its consistency and workability. The temperature of the concrete was also determined by using a temperature measuring device (Eagletech TP101, Aukey International Ltd., China) inserted into the concrete as per ASTM C1064/C1064M-17 (2017). The concrete specimens were made in 150 x 300 mm (6 x 12 in.) plastic cylinder molds (Figure 2). The ER was measured using a probe (Mancio *et al.*, 2010)

after the fresh concrete was placed into the plastic cylinder molds. The probe was vertically immersed in fresh concrete and centrally located within the plastic cylinder mold to measure the ER of fresh concrete within 30 min after mixing and before the initial setting of the cement. Two measurements were made as a standard procedure (Mancio *et al.*, 2010), and the computed average was used as one of the input variables for the ANN model. After removing the probe, the cylinder mold vibrated again using a rubber mallet. After 24 h of initial curing, the plastic cylinder mold was removed, and appropriate identification was placed in each specimen. Specimens were placed in a curing pond filled with tap water and cured at room temperature (24±2 °C). The making and curing of the specimen followed the ASTM C192/C192M-19 (2019). The compressive test was performed on the 28th day of curing as per ASTM C39/C39M-21 (2021). The compressive test was conducted in the J-Con Material Testing Center, a Department of Public Works and Highways (DPWH)-Bureau of Research and Standards (BRS) accredited testing laboratory.

Table 2. Details of concrete mix proportions

Mix no.	Slump (mm)	Max. size of aggregate (mm)	% air content	W/C ratio	FM of FA	Mix proportion of concrete (kg)			
						Water	Cement	FA	CA
1	130	19	0.02	0.3143	3.06	13.20	42.00	48.00	61.20
2	160	19	0.02	0.3316	3.06	13.00	39.20	51.30	61.20
3	163	19	0.02	0.3539	3.06	13.80	39.00	57.55	65.60
4	161	19	0.02	0.3973	3.06	13.35	33.60	56.10	61.20
5	170	19	0.02	0.3815	3.06	11.75	30.80	58.50	61.25
6	150	19	0.02	0.4161	3.06	11.65	28.00	60.90	61.25
7	130	19	0.02	0.4135	3.06	11.00	26.60	62.10	61.22
8	131	19	0.02	0.4564	3.06	11.50	25.20	63.30	61.25
9	182	19	0.02	0.4622	3.06	11.00	23.80	64.50	61.20
10	190	19	0.02	0.3119	3.06	13.10	42.00	46.80	61.25
11	150	19	0.02	0.3189	3.06	12.50	39.20	49.20	61.20
12	190	19	0.02	0.3434	3.06	12.50	36.40	51.60	61.20
13	150	19	0.02	0.3134	3.06	10.53	33.60	54.00	61.20
14	195	19	0.02	0.3571	3.06	11.00	30.80	56.45	61.25
15	240	19	0.02	0.3912	3.06	11.50	29.40	57.65	61.25
16	185	19	0.02	0.3232	3.06	9.05	28.00	61.25	61.35

Table 2 continued.

17	172	19	0.02	0.3571	3.06	9.50	26.60	65.45	61.40
18	178	19	0.02	0.3770	3.06	9.50	25.20	66.65	61.40
19	132	9.5	0.03	0.2865	3.33	14.44	50.40	45.30	51.20
20	141	9.5	0.03	0.3038	3.33	14.46	47.60	47.60	51.20
21	165	9.5	0.03	0.3237	3.33	14.50	44.80	48.30	50.40
22	154	9.5	0.03	0.3333	3.33	14.00	42.00	50.60	50.40
23	195	9.5	0.03	0.3304	3.33	12.95	39.20	52.95	50.40
24	130	9.5	0.03	0.3201	3.33	11.65	36.40	55.25	50.40
25	100	9.5	0.03	0.3720	3.33	12.50	33.60	57.60	50.40
26	125	9.5	0.03	0.4091	3.33	12.60	30.80	59.90	50.40
27	90	9.5	0.03	0.4464	3.33	12.50	28.00	62.20	50.40
28	143	9.5	0.03	0.3175	3.33	16.00	50.40	40.90	50.40
29	157	9.5	0.03	0.3151	3.33	15.00	47.60	43.20	50.40
30	130	9.5	0.03	0.3348	3.33	15.00	44.80	45.55	50.40
31	160	9.5	0.03	0.3452	3.33	14.50	42.00	47.85	50.40
32	164	9.5	0.03	0.3546	3.33	13.90	39.20	50.20	50.40
33	166	9.5	0.03	0.3984	3.33	14.50	36.40	52.50	50.40
34	142	9.5	0.03	0.3720	3.33	12.50	33.60	54.85	50.40
35	100.7	9.5	0.03	0.4367	3.33	13.45	30.80	57.15	50.40
36	140	9.5	0.03	0.5071	3.33	14.20	28.00	59.50	50.40

Figure 2. Plastic cylinder mold

A total of 144 datasets from 36 different concrete mix designs were obtained from the experiment. The datasets were arranged in a format of eight input variables and one output variable. The input variables used for the proposed ANN model were not only limited to the derived ratios of concrete constituents and their properties such as W/C ratio, fine aggregate-coarse aggregate (FA/CA) ratio, cement-aggregate (C/A) ratio, fineness modulus (FM) of fine aggregates and maximum aggregate size (mm); the results from field quality tests such as slump (mm), concrete temperature (°C), and the ER of the fresh concrete (Ω-m) were also considered. The output variable was the 28-day compressive strength of concrete. The W/C ratio, FA/CA ratio, C/A ratio, and the total amount of aggregates, A, were computed using Equations 1, 2, 3 and 4, respectively. The range of input and output variables used in the proposed ANN model is shown in Table 3.

$$W/C = \dfrac{mass\ of\ water\ (kg)}{mass\ of\ cement\ (kg)} \tag{1}$$

$$FA/CA = \dfrac{amount\ of\ FA}{amount\ of\ CA} \tag{2}$$

$$C/A = \dfrac{amount\ of\ cement}{total\ amount\ of\ aggregates} \tag{3}$$

$$A = FA + CA \tag{4}$$

Table 3. The range of input and output variables used in the ANN model

Name of the variables	Data used in training and testing the model	
	Minimum	Maximum
Input variables		
ER (Ω-m)	3.7217	11.0768
Slump (mm)	90	240
Max. aggregate size (mm)	9.5	19
Fineness modulus of fine aggregate	3.06	3.33
Temperature (°C)	26.40	35.10
W/C ratio	0.2865	0.5071
FA/CA ratio	0.7641	1.2341
C/A ratio	0.1893	0.5520
Output variable		
28-day compressive strength	17.4	45.63

2.2 Data Preparation

To have a better understanding of the datasets, data visualization was applied in this study. A histogram is a graphical representation of the distribution of each variable in the datasets. The purpose of a histogram is to estimate whether a variable has a normal distribution, is skewed, or even has an exponential distribution. It also detects the possible outliers in the datasets. A density plot is another way of getting a quick idea of the distribution of each variable in the datasets. It is a smooth distribution curve drawn on top of each histogram. A normal distribution, also referred to as Gaussian distribution, has a symmetrical bell-shaped curve. Many machine learning algorithms perform better when the variables in the datasets have a Gaussian probability distribution (Brownlee, 2020b). Non-normal data imposes a great challenge in developing a robust predictive model; hence, data visualization such as histograms and density plots were employed to facilitate the choice of techniques to improve the robustness of the predictive model.

Correlation indicates the strength of association between two variables (Fernando and Walters, 2021). The strength of the relationship is decided by the weight of the correlation coefficient, which ranges from +1 to -1. A weight of ±1 specifies a great degree of association between two variables. The connection between the two variables is weaker as the correlation coefficient goes towards 0. There are several varieties of correlation coefficients, and this study used Pearson's coefficient, which is calculated using Equation 5.

$$P_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\left[n \sum x_i^2 - (\sum x_i)^2 \right]} \sqrt{\left[n \sum y_i^2 - (\sum y_i)^2 \right]}} \tag{5}$$

where P_{xy} is the Pearson's correlation coefficient between x and y; n is the number of observations; and x_i and y_i are the values of x and y for *ith* observation, respectively.

The correlation matrix is a table displaying the correlation coefficients between two variables. It provides an easier way to understand the relationship between variables, which is useful to know because some machine learning algorithms like linear and logistic regression can have poor performance if there are highly correlated input variables.

2.3 ANN

2.3.1 Selection of ANN Architecture

The ANN models have hyperparameters that need to be set to customize the model specific to the datasets used such as the number of hidden layers, the number of neurons per hidden layer activation functions, optimization algorithm, batch size and scaling algorithm. The process of finding the best possible combination of hyperparameters is called hyperparameter optimization or hyperparameter tuning. This study administered GridSearchCV to automate the tuning process. GridSearchCV searched all different hyperparameter combinations defined by the authors, as shown in Table 4, using the Cross-Validation method (Brownlee, 2018a). The k-Fold Cross-Validation was used to find suitable hyperparameters. It ensured that the model performed uniformly even at the different subgroups within the whole dataset. In this study, the k-fold was equal to five, which meant that the datasets were split into 80% training and 20% testing (validation) for each fold.

Table 4. Hyperparameters that are considered during grid searching

Hyperparameters	
Activation function	Hyperbolic tangent (*Tanh*)
	Rectified linear unit (ReLu)
Optimization algorithm	Stochastic gradient descent
	Adaptive moment estimation
Batch size	4, 8, 16, 32
Scaling algorithms	Standardization
	Normalization (Min-max scaling)

2.3.2 Data Scaling

The scale of input and output variables is an important factor in improving the stability and performance of the ANN model (Brownlee, 2019a). As shown in Table 3, the input variables have different units of measurement and different scales. To avoid problems related to the poor performance of the ANN model during the learning process, data scaling was applied before the input layer. Data normalization or also known as minimum-maximum scaling and standardization were both employed during the hyperparameter tuning.

Normalization rescales the range of the variable so that all values are within the range of 0 and 1, while standardization rescales the distribution of a variable so that the mean of the observed variable is 0 and the standard deviation is 1. Data scaling was applied after the datasets were split into training and testing sets as discussed in the next section.

2.3.3 Calibration of the Proposed ANN Model

The proposed ANN model was calibrated based on the experimental results to investigate its accuracy. A multi-layer feed-forward network with a back-propagation learning rule was used using Python. The datasets were sorted into training and testing sets. Generally, the ANN model is trained using known data, referred to as training sets, and its performance is validated using unknown data or first seen data, known as testing sets. About 70% of the datasets were fed to ANN for training, and the remaining 30% were used for testing the trained model. The data scaling was applied only after splitting the datasets to avoid information leakage on testing sets which neglected the purpose of having testing sets. Before starting the training, values for epoch, batch size, learning rate and momentum rate were decided. The training process will run for a fixed number of iterations through the training sets called the epoch. The initial weights were randomly initialized and were updated every batch. The batch size refers to the number of training datasets considered by the model within an epoch before weights are updated. A constant learning rate was used, which was equal to 0.001. It refers to the increment at each iteration and controls how fast the model is customized to the problem. The models were trained until they converged at 800 epochs using stochastic gradient descent (SDG) algorithm. The SDG is an optimization algorithm used to explore viable sets of weights the model may use to make good predictions. A momentum rate of 0.9 was utilized to expedite the optimization process. The error for those predictions was calculated using a loss function, and the value calculated by the loss function was referred to as loss.

The goal of the optimization algorithm is to minimize the loss or the difference between the actual and predicted data by changing the weights. Large weights can signify that the network has overfitted the training datasets and will likely perform poorly when making predictions on new data (Brownlee, 2018b). A solution to this problem is to impose weight regularization to change the learning algorithm to encourage the network to keep the weights small, thus reducing the overfitting of the training datasets and improving the model's generalization. Weight regularization was imposed in this study to penalize

the model during training by adding the weight size as a penalty term to the loss function. The weight size was calculated using two main approaches, L1 and L2. The L1 calculated the sum of the absolute values of the weights, while L2 computed the sum of the squared values of the weights. The regression model that uses L1 regularization is called lasso regression, and the model that uses L2 regularization is called ridge regression. L2 regularization can also deal with multicollinearity problems. It can be used to estimate the significance of the input variables and based on that; it can penalize the insignificant input variables. Finally, the performance of the ANN model was validated by presenting the remaining 30% of the testing set to the model. The prediction performance of the ANN model was evaluated by the following statistical parameters: 1) correlation coefficient (*R*); 2) absolute fraction of variance (R^2); 3) mean absolute error (*MAE*); and 4) mean square error (*MSE*) which are expressed by Equations 6, 7, 8 and 9, respectively.

$$R = \frac{\sum_{i=1}^{n} (t_i - \bar{t})(o_i - \bar{o})}{\sqrt{\left[n\sum t_i^2 - (\sum t_i)^2\right]\left[n\sum o_i^2 - (\sum o_i)^2\right]}} \tag{6}$$

$$R^2 = \frac{(n\sum t_i o_i - \sum t_i \sum o_i)^2}{\left[n\sum t_i^2 - (\sum t_i)^2\right]\left[n\sum o_i^2 - (\sum o_i)^2\right]} \tag{7}$$

$$MAE = \frac{1}{n}\left[\sum_{i=1}^{n} |t_i - o_i|\right] \tag{8}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (t_i - o_i)^2 \tag{9}$$

where t_i is the target value; o_i is the output value; n is the total number; and $\bar{t}$ and $\bar{o}$ are the average value of target and output values, respectively.

The R measures the strength of the relationship between the predicted and the actual values (Fernando and Walters, 2021). As expressed by Equation 6, the formula returns a value between -1 and 1, and the value closer to 1 or -1 indicates a strong correlation. The R^2 is the proportion of the variance in the output variable that is foreseeable from the input variables (Asteris *et al.*, 2021). R^2 ranges from 0 to 1, where 0 implies that the input variables do not explain the variation of the output variable, while 1 suggests that the input variables elucidate the variation in the output variable. The MAE is the mean of all the absolute errors between the predicted value and the actual value (Ahmad *et al.*, 2020). The MSE informs how close a regression line is between

the predicted and the actual values by quantifying the errors between the two values and squaring them (Ahmad *et al*., 2020).

2.4 Feature Importance Analysis

Feature importance analysis was performed using permutation feature importance (PFI) to evaluate the significance of the input variables or features on the output variable. The PFI measures the increase in the model's prediction error after permuting the feature. It tells which features a model relies on the most. A feature is important if shuffling its values increases the prediction error, which indicates that the model relied on the feature for prediction. A feature is unimportant if shuffling its values does not change the prediction error, which shows that the model ignored the feature for the prediction (Brownlee, 2020a).

The computation of PFI was performed after the authors obtained a trained ANN model with acceptable performance. The training datasets, T, were used to estimate the ANN model's predictive power using R^2. The model's performance was recorded as baseline performance, P, on the datasets, T. To calculate the importance of a particular feature, f, the column for that feature is temporarily removed from the testing datasets, T. The new model was fitted on the reduced datasets using GridSearchCV as discussed in section 2.3.1, and then its feature columns were shuffled. Afterward, the new model's performance, $P_{i,f}$, on reduced and shuffled datasets, $T_{i,f}$, was computed. The preceding three steps were repeated multiple times, $i = 1, \ldots, I$. The difference between the baseline performance, P, and the average performance of the new models calculated the feature importance score, FI_f, for the feature f, as expressed in Equation 10 (Fisher *et al*., 2019).

$$FI_f = P - \frac{1}{R} \sum_{i=1}^{I} P_{i,f} \tag{10}$$

The feature importance scores can also be used for feature selection. Feature selection is a procedure of reducing the number of input variables when developing a predictive model to reduce the computational cost of modeling and improve the model's performance (Brownlee, 2019b). The comparison of ANN models without and with feature selection was performed in this study to verify this conjecture. Their performance was evaluated in terms of the selected statistical parameters.

3. Results and Discussion

3.1 Descriptive Statistics of Data

As stated earlier, the datasets were collated and presented through visualization such as histograms and density plots (Figure 3). As shown in Figure 3, most of the variables, if not all, did not follow a normal distribution. The normal distribution, also called Gaussian, has a symmetrical bell-shaped curve. To emphasize, the maximum aggregate size and fineness modulus variables had the least likely to be considered normal. That was mainly because there were only two maximum aggregate sizes used in this study, which were 9.5 and 19 mm.
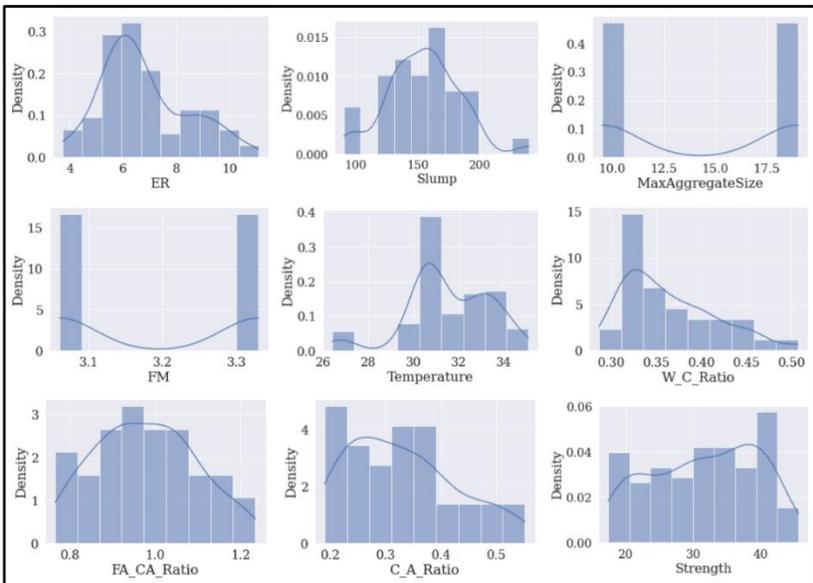
Figure 3. Histogram and density plots of each variable in the dataset

Moreover, there were two batches of fine aggregates used in this work. Although they were from the same source, the test showed that these batches of fine aggregates had slightly different fineness modulus (3.06 and 3.33). Since machine learning algorithms perform better when numerical values have a normal distribution, this study employed power transforms like Box-Cox and Yeo-Johnson transformations to make the probability distribution of the variables more Gaussian.

Figure 4 presents the correlation matrix of the datasets used in this study. A correlation matrix is a table that displays the correlation coefficients between variables. Each cell in the correlation matrix indicates the correlation between two variables.
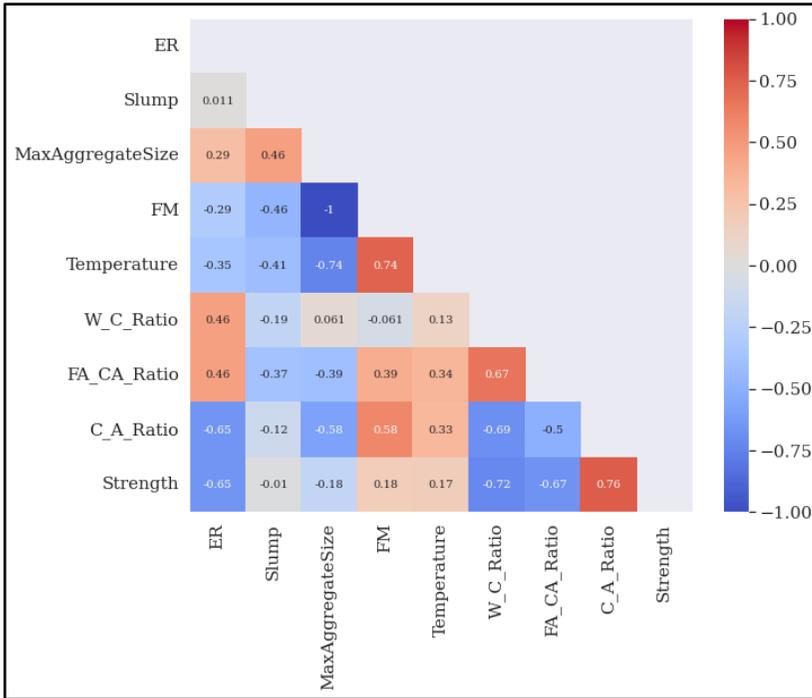

Figure 4. Correlation matrix

It was found that input variables ER and FA/CA ratio showed marginal correlation ($|r| > 0.5$) to the 28-day compressive strength, while input variables W/C Ratio and C/A Ratio exhibited strong correlation ($|r| > 0.7$) to the 28-day compressive strength of concrete. It is important to note that a high correlation between input and output variables is a good indication of better predictability. However, when two or more input variables are highly correlated ($|r| > 0.7$), it indicates a presence of multicollinearity, which is present among input variables FM, temperature, and maximum aggregate size. Some machine learning algorithms, such as linear and logistic regression, can have poor performance if multicollinearity is present in the data. However, Aggarwal *et al*. (2014) discovered that when using the concrete constituents in concrete compressive strength prediction, multicollinearity is most likely present, and the use of ridge regression can circumvent the problem (Aggarwal *et al*., 2014;

Chong *et al*., 2021). Another way of dealing with multicollinearity is to drop some redundant variables or by simplifying the model by using feature selection techniques. Both ridge regression and feature selection were applied in this study, which were discussed in sections 2.3.3 and 2.4, respectively.

3.2 Parametric Configuration of ANN Model

The GridSearchCV was administered to automate the selection of the best hyperparameter combinations. GridSearchCV generated a long list of the results after multiple training sessions, but only the top five most optimal architectures are presented in Table 5.

Table 5. ANN models generated during grid searching

Model	Hidden layers	Activation function	Optimization algorithm	Batch size	Scaling algorithm	R^2	Rank
without Feature Selection							
MO-1	12, 6	*Tanh*	SDG	8	Standardization	0.8926	1
MO-2	12, 6	*Tanh*	SDG	32	Standardization	0.8901	2
MO-3	18, 6	*Tanh*	SDG	32	Standardization	0.8882	3
MO-4	16, 6	*Tanh*	SDG	32	Standardization	0.8775	4
MO-5	14, 6	*Tanh*	SDG	8	Standardization	0.8740	5
with Feature Selection							
MW-1	7, 6	*Tanh*	SDG	8	Standardization	0.9063	1
MW-2	14, 6	*Tanh*	SDG	16	Standardization	0.9054	2
MW-3	12, 6	*Tanh*	SDG	8	Standardization	0.8969	3
MW-4	7, 7, 6	*Tanh*	SDG	8	Standardization	0.8883	4
MW-5	13, 12, 6	*Tanh*	SDG	8	Standardization	0.8791	5

The ANN model exhibiting the highest R^2 is the best model which can accurately predict the 28-day compressive strength of delivered concrete at the job site. As shown in Table 5, MO-1 had the highest R^2 value of 0.8926; therefore, it was selected as the final model for the prediction of the 28-day compressive strength of delivered concrete. The final ANN model has two hidden layers with 12 and six neurons in the first and second hidden layers, respectively. Meanwhile, the *Tanh* function was used as an activation function throughout the layers, and SDG was used as the optimizer. To find the best architecture, the authors tried two scaling techniques: 1) normalization and 2)

standardization. However, the best model (MO-1) only employed the standardization method to rescale the input variables into the desired value ranges. As mentioned earlier, feature selection was performed in this study. Some of the input variables were removed based on their importance scores which are further discussed in section 3.3. The removal of input variables resulted in a new fitted model using GridSearchCV. As shown in Table 5, MW-1 had the highest R^2 value of 0.9063; therefore, it was selected as the final model for the prediction of the 28-day compressive strength of delivered concrete. The final ANN model had two hidden layers with seven and six neurons in the first and second hidden layers, respectively. It also had the same activation function, optimization algorithm, batch size, and scaling algorithm as the final model without feature selection.

3.3 Feature Importance Analysis Results

The result of the feature importance analysis (Table 6) indicates that the C/A ratio was the most significant followed by the FA/CA ratio, ER, W/C ratio, slump and the temperature of delivered concrete.

Table 6. Feature importance scores

Rank	Feature	Score
1	C/A ratio	0.53803
2	FA/CA ratio	0.20456
3	ER	0.11503
4	W/C ratio	0.09468
5	Slump	0.03797
6	Temperature	0.01420
7	Maximum aggregate size	0.00001
8	FM	0.00000

3.3.1 Effect of C/A Ratio

This study found that the compressive strength of concrete was related to the C/A ratio (Figure 7a). The results are similar to Poon and Lam (2008), which revealed that when the A/C ratio decreased, the compressive strength of concrete increased. This can be explained by the decrease in the W/C ratio when the cement content was increased given constant workability. As established by the well-known Abram's model (Abrams, 1927), if the W/C

ratio minimizes, the compressive strength of the concrete escalates. Moreover, if the volume of aggregates remains constant and the water content is the same for the constant workability, the rise in cement content means a stronger paste for the expanse of aggregates (Mishra, 2013). Hence, when the C/A ratio increases, the compressive strength of concrete increases as well.

3.3.2 Effect of FA/CA Ratio

The volume of the concrete was mostly occupied by fine and coarse aggregates (Table 2), which explained why the two most significant features of the model had something to do with aggregates. Misha (2013) underscored that the compressive strength of concrete decreases when the FA/CA ratio increases. This was validated in the present study (Figure 7b); when the FA/CA ratio increases, it requires more paste due to the increased overall aggregate surface area (Mishra, 2013). Hence, water demand is expected to increase to make more paste which can cause the W/C ratio to increase as well. Thus, the compressive strength of concrete will likely decrease due to the increase in the W/C ratio.

3.3.3 Effect of ER

The relationship between ER of fresh concrete and the actual W/C ratio of concrete was established by Mancio *et al*. (2010), and their strong direct correlation was also confirmed by Neha and Rajeev (2015). In their study, they varied the W/C ratio while keeping the water content constant, which indicated that an increase in the W/C ratio means a lower cement content. These further demonstrated that fewer conductive ions were released by the reduced quantity of cement per unit volume of concrete; hence, there is a lower conductivity or higher resistivity (Mancio *et al*., 2010; Neha and Rajeev, 2015). Meanwhile, Obla *et al*. (2018) related the ER of fresh concrete to the W/C ratio of concrete by keeping the paste volume constant and by increasing or decreasing the paste volume. The paste volume was calculated as the sum of the volumes of the cement and the mixing water content. Their results were presented in three scenarios (Obla *et al*., 2017). First, at a constant paste volume, as the W/C ratio decreased, the water content decreased, while the cement content also increased. This led to the more conductive ions in the mixture with lesser conductive pathways. As a result of these competing effects, as the W/C ratio increases, the ER of fresh concrete may increase or decrease depending on which of the factors is dominant. Second, when the W/C ratio is increased while keeping the cement content constant, the water

content increases, and the paste volume also increases. The higher mixing water content will provide greater conductive pathways consequently lowering the fresh concrete's electrical resistivity. Third, when the W/C ratio is increased by reducing the amount of cement in the mixture and while keeping the water content fixed, the paste volume decreases. Hence, as the W/C ratio is increased, the electrical resistivity is also increased, and this was referred to as the "pore solution" effect in Obla *et al*.'s (2017) discussion, which supported the results of Mancio *et al*. (2010).

In this study, the paste volume and the cement content were varied. The paste volume varied from 0.0175 to 0.0320 m^3, while the cement factor varied from 9 to 18. On the other hand, the water content varied from 9.05 to 16 L; however, as demonstrated in Figure 5, the variation in the water content was lesser than the variation of cement content making it almost negligible.
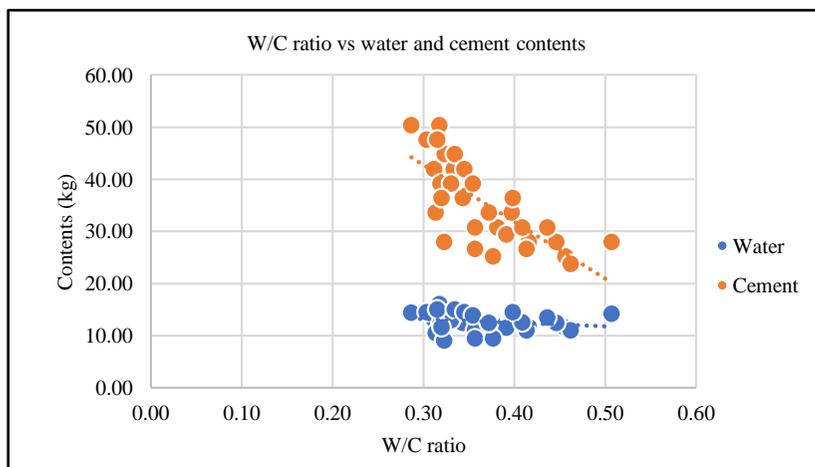
Figure 5. W/C ratio versus water and cement contents

As presented in Figures 5 and 6, the experimental data indicated that an increase in the W/C ratio was chiefly caused by a decrease in cement content. Although the water content also varied, the variation was lesser than the cement content, and it can be considered almost constant. Furthermore, as the W/C ratio increased, the paste volume decreased, which had the same observation from scenario three discussed by Obla *et al*. (2017). Therefore, as the W/C ratio is increased due to the "pore solution" effect, the ER of fresh concrete is likewise increased. Since the W/C ratio is inversely proportional to the compressive strength of concrete, that means the ER of fresh concrete increased as its compressive strength decreased (Figure 7c).
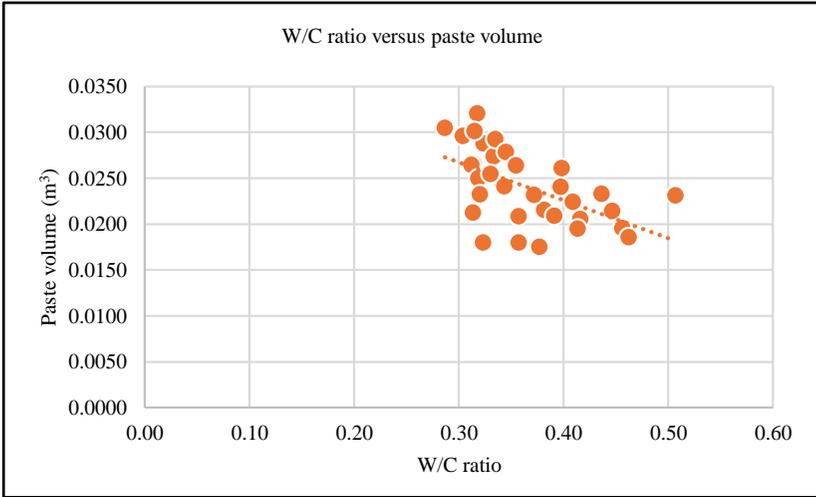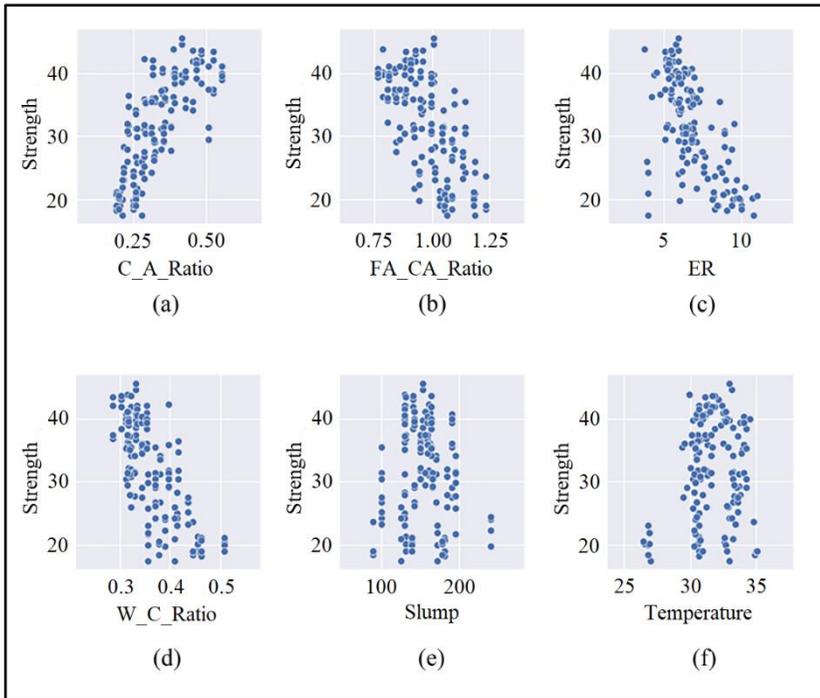
Figure 6. W/C ratio versus paste volume

Figure 7. Scatter plot: (a) C/A ratio vs. strength, (b) FA/CA vs. strength, (c) ER vs. strength, (d) W/C ratio vs. strength, (e) slump vs. strength, (f) temperature vs. strength

3.3.4 Effect of W/C Ratio

The findings of this study (Figure 7d) are identical to the well-known Abram's model (Abrams, 1927). The results showed that the compressive strength of concrete decreased as the W/C ratio increased. Briefly, the W/C ratio was inversely proportional to the compressive strength of concrete. Although the findings of this study are similar to Abram's model, it is significant to note that the compressive strength of concrete was not only determined by the W/C ratio but was also affected by the other factors as confirmed by several researchers (Pengcheng *et al*., 2020; Muliauwan *et al*., 2020; Silva *et al*., 2020; Xu *et al*., 2021; Lin and Wu, 2021).

3.3.5 Effect of Slump

It can be observed from Figures 4 and 7e that slump had a poor correlation to the compressive strength of concrete. However, it can still affect its compressive strength indirectly. Concrete mix with a low slump is more difficult to work with and is not suitable for some applications preventing the concrete from consolidating properly after it is poured. Entrapped air is removed during the consolidation process by vibration or compaction. Therefore, when the concrete is not properly consolidated, the trapped pockets of air in the concrete decrease the ability of concrete to withstand compression. On the other hand, when the slumps are above average, the strength and durability of the concrete reduce as well. The concrete mix weakens when too much water is added. Generally, when concrete has too much water, concrete aggregates tend to segregate and settle at the bottom. So, when the concrete hardens, it is not a homogeneous mixture. Hence, admixtures are introduced to concrete mixtures to achieve a higher slump so that the quality of concrete is maintained. However, it is noteworthy that even when the concrete mix has no admixtures, the concrete mix with a given W/C ratio can give different slump measurements depending on the aggregates' gradation.

3.3.6 Effect of Temperature

As shown in Figure 4, the temperature had a higher correlation to compressive strength than slump. Yet, it has a lower importance score (Table 6), which means the slump was more relevant to the compressive strength predictive model than temperature. The correlation coefficient between temperature and compressive strength was -0.35, which can be considered a poor correlation. It can also be observed in Figure 7f that the scatter plot for temperature, and

compressive strength was vertically concentrated rather than sloping. The study found that between a concrete temperature of 26.40 to 35.10 °C, the temperature had no direct effect on the compressive strength of concrete.

Nevertheless, when the temperature of concrete increases, the rate of the hydration process also increases; hence, the concrete will gain strength rapidly. However, the final strength of the concrete that is kept at a higher temperature will be lower due to the less well-structured and more porous form of hardened concrete paste resulting from a faster rate of hydration reaction (Patel, 2018). Temperature test on the delivered concrete is already a common practice at the job site; hence, the concrete temperature was considered as an input variable rather than the ambient temperature.

Moreover, the feature importance scores were also used as the basis for feature selection. A threshold of 1e-5 was used, meaning a score of less than that was eliminated. As shown in Table 6, the input variables maximum aggregate size and fineness modulus had feature importance scores of 0.00001 and 0.00000, respectively. Hence, these two input variables were eliminated from the model. As mentioned earlier, feature selection is a process of reducing the number of input variables when developing a predictive model to reduce the computational cost of modeling and improve the performance of the model (Brownlee, 2019b). The results of the comparison between ANN models without and with feature selection are discussed in section 3.4.

3.4 ANN Model's Performance

The performance of the ANN models without and with feature selection is presented in Table 7. Even without feature selection, the model had already shown good performance in predicting the 28-day compressive strength values. The R and R^2 for training sets were found as 0.9448 and 0.8926 implying that there was a strong correlation between the actual and predicted values in both cases. The testing sets were fed to the model, and the results showed that R and R^2 were improved (Table 7). The results validated the model performance in the training sets. The MAE for training was 2.0749, which stands for the average difference between the actual and predicted values. This was improved in the testing set with an MAE value of 1.9908. The MSE for training sets was 6.5638 indicating that the ANN model over forecasted the 28-day compressive strength averagely by 6.5638 MPa for training. The MSE was decreased to 5.9914 in the testing sets, which showed that the model performs well with unseen data.

On the contrary, it can be seen in the results (Table 7) that the ANN model with feature selection outperformed the ANN model without feature selection in all four statistical parameters. The *R* values of the ANN model with feature selection were increased by 0.76 and 1.69% in training and testing sets, respectively, compared with the model without feature selection. Furthermore, it was also found that the MSE values for training and testing sets were decreased by 0.8381 and 1.8882 MPa, respectively. Thus, the results showed that feature selection helped improve the prediction accuracy of the model. The visual plots are shown in Figures 8 and 9, which present the relationship between the experimental data and the predicted data in the training and testing stage of the ANN model with feature selection, respectively. It supported that the ANN model with feature selection can predict the 28-day compressive strength of delivered concrete with great precision.

Table 7. Performance of models without and with feature selection

ANN models	Statistical parameters			
	R	*R²*	*MAE*	*MSE*
Without feature selection				
Training Set	0.9448	0.8926	2.0749	6.5638
Testing Set	0.9499	0.9023	1.9908	5.9914
With feature selection				
Training Set	0.9520	0.9063	1.9275	5.7257
Training Set	0.9660	0.9331	1.6421	4.1032

The best model that can predict the 28-day compressive strength of delivered concrete at the job site is the model with the least values of MAE and MSE and the highest values of *R* and *R²* (Ahmad *et al*., 2021). From the engineering point of view, the ANN prediction model is considered accurate and robust when the *R*-value is above 95% (Tavares *et al*., 2020). In this study, the ANN model with feature selection was selected as the best model with an architecture defined in Figure 10. As discussed earlier, the ANN model originally had eight neurons in the input layer. It was reduced to six neurons after feature selection which represents the top six most significant input variables, namely C/A ratio, FA/CA ratio, ER, W/C ratio, slump and the temperature of the delivered concrete. Furthermore, the ANN model had two hidden layers with seven and six neurons in the first and second hidden layers, respectively. Then, one neuron in the output layer refers to the 28-day compressive strength of delivered concrete at the job site.
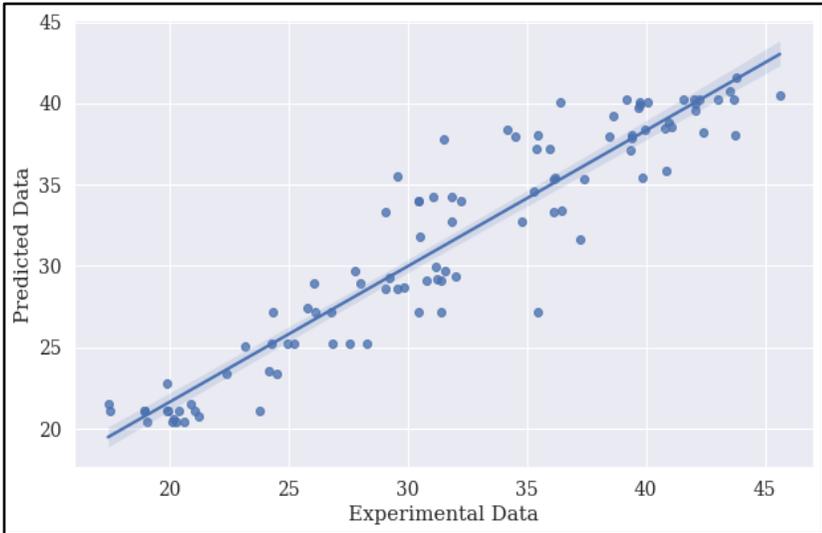
Figure 8. Correlation of experimental and predicted data in the training stage
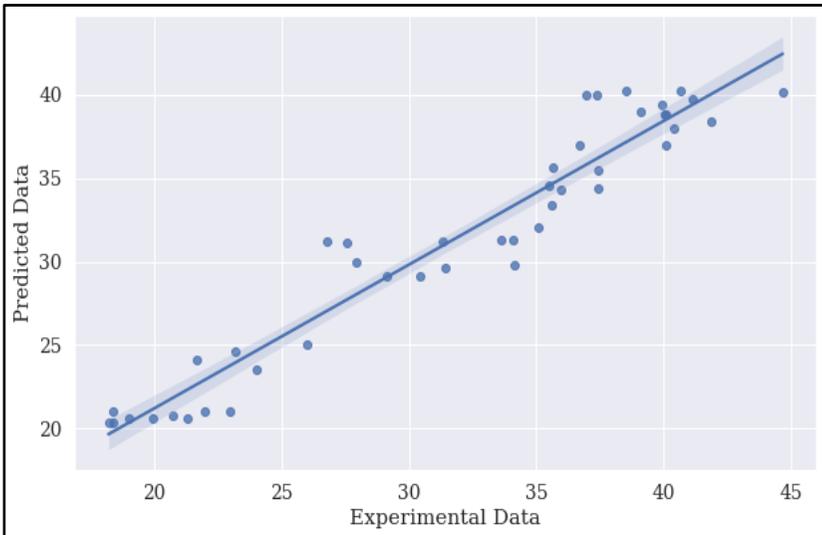

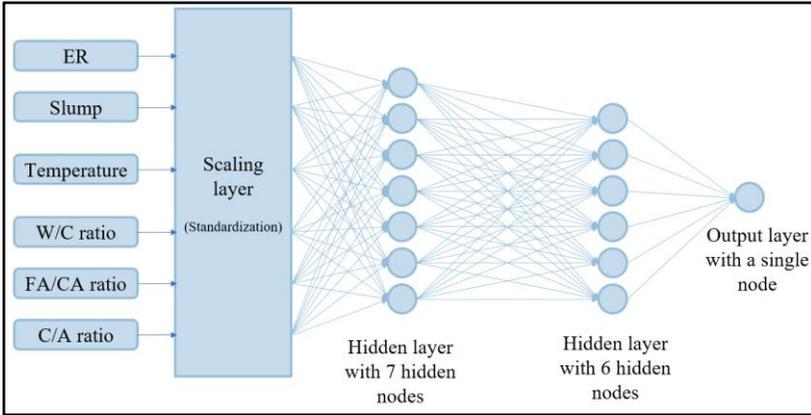Figure 9. Correlation of experimental and predicted data in the testing stage

Figure 10. Proposed ANN model

3.5 Comparison with Other Studies

There is no general empirical formula to compute the compressive strength of concrete. In engineering practice, the compressive strength of concrete is measured through a compression test of cylindrical concrete specimens after 28 days of curing (National Ready Mixed Concrete Association, 2014). The compressive strength of concrete is computed by dividing the maximum load resisted by the specimen during the test by the average cross-sectional area as described in ASTM C39/C39M-21 (2021). For example, the 150 mm – diameter cylindrical concrete specimen failed at 638.45 KN during compression testing. The computed compressive strength was 36.13 MPa, as shown below using Equation 11.

$$f'c = \frac{F}{A} \tag{11}$$

The sample computation is the following:

$$f'c = \frac{(638.45 \times 1000)N}{\frac{\pi}{4}(150\ mm)^2} = 36.13\ MPa$$

The conventional way of determining the compressive strength of concrete is time-consuming. Hence, to resolve this problem, conventional statistical methods and machine learning techniques were explored by several researchers to predict the 28-day compressive strength of concrete (Muliauwan *et al.*, 2020; Silva *et al.*, 2020; Xu *et al.*, 2021). The results of the

current study have been compared with the studies conducted by Jee *et al*. (2004), Derousseau *et al*. (2019) and Xu *et al*. (2021) (Table 8). Jee *et al*. (2004) predicted the 28-day compressive strength of in-situ concrete using multiple linear regression (MLR) analysis. Although the study aimed to predict the compressive strength of in-situ concrete, it did not capture the variation caused by construction practices and variable environmental conditions in developing the model, which implies that the model was proposed for laboratory-sampled concrete instead of in-situ concrete. Derousseau *et al*. (2019) concluded that the predictive models trained by laboratory-obtained data were not reliable when used for predicting the compressive strength of field-placed concrete. However, it can be significantly improved by employing hybrid data, which is a combination of laboratory-obtained and field-specific data. Xu *et al*. (2021) used a genetic algorithm (GA) to comprehensively evaluate the various influencing factors from the 4M1E perspective and used random forest (RF) as a modeling algorithm to predict the compressive strength of concrete based on the selected influential factors. Although the approach improved the model accuracy, the collection of 20 selected variables was difficult, which made the method more tedious. On the other hand, the current study focused on capturing the variability of water content and W/C ratio caused by variable environmental conditions and other uncertainties encountered during the transportation of concrete from the laboratory or batching plant to the construction or job site to predict the compressive strength of delivered concrete. In Table 8, the proposed model of the current study had the second-highest R^2.

Table 8. Comparative analysis

Sources	Method	R^2	Rank
Current study	ANN	0.9331	2
Xu *et al*. (2021)	GA-RF	0.9645	1
Jee *et al*. (2004)	MLR	0.6700	3
Derousseau *et al*. (2019)	RF	0.5100	4

4. Conclusion

The *R* values of the ANN model with feature selection were 0.9520 and 0.9660 for training and testing sets, respectively. From the engineering standpoint, these results, which were above 95%, demonstrated that the ANN model could

predict the 28-day compressive strength of delivered concrete with high accuracy and robustness. After the feature importance analysis, the C/A ratio was found to be the most influential feature of the model followed by FA/CA ratio, ER, W/C ratio, slump, the temperature of delivered concrete, maximum aggregate size and FM. Applied in this study, feature selection improved the prediction accuracy of the model. The *R* values of the ANN model with feature selection were increased by 0.76 and 1.69% in training and testing sets, respectively, compared with the model without feature selection. Furthermore, it was also revealed that the MSE values for training and testing sets were decreased by 0.83812 and 1.88814 MPa, respectively. Lastly, the six selected variables during feature selection were the C/A ratio, FA/CA ratio, ER, W/C ratio, slump and the temperature of delivered concrete.

5. Acknowledgement

6. References

Abrams, D. (1927). Water-cement ratio as a basis of concrete quality. ACI Journal Proceedings, 23, 452-457.

Aggarwal, R., Kumar, M., Sharma, R., & Sharma, M.K. (2014). Reliability based design optimization of concrete mix proportions using generalized ridge regression model. International Journal of Science and Engineering, 8. https://doi.org/10.12777/ij se.8.1.26-37

Ahmad, A., Elchalakani, M., Elmesalami, N., El Refai, A., & Abed, F. (2021). Reliability analysis of strength models for short-concrete columns under concentric loading with FRP rebars through artificial neural network. Journal of Building Engineering, 42(102497). https://doi.org/10.1016/j.jobe.2021.102497

Ahmad, A., Plevris, V., & Khan, Q. (2020). Prediction of properties of FRP-confined concrete cylinders based on artificial neural networks. Crystals, 10(811). https://doi.or g/10.3390/cryst10090811

American Concrete Institute (ACI) Committee 211. (2002). Standard practice for selecting proportions for normal, heavyweight, and mass concrete (ACI 211.1-91) (reapproved 2009). Michigan, United States: American Institute of Concrete.

American Society for Testing and Materials (ASTM) C39/C39M-21. (2021). Standard test method for compressive strength of cylindrical concrete specimens. West Conshohocken, PA: ASTM International.

American Society for Testing and Materials (ASTM) C143/C143M-20. (2020). Standard test method for slump of hydraulic-cement concrete. West Conshohocken, PA: ASTM International.

American Society for Testing and Materials (ASTM) C192/C192M-19. (2019). Standard practice for making and curing concrete test specimens in the laboratory. West Conshohocken, PA: ASTM International.

American Society for Testing and Materials (ASTM) C1064/C1064M-17. (2017). Standard test method for temperature of freshly mixed hydraulic-cement concrete. West Conshohocken, PA: ASTM International.

Asteris, P., Skentou, A., Bardhan, A., Samui, P., & Lourenco, P. (2021). Soft computing techniques for the prediction of concrete compressive strength using non-destructive tests. Construction and Building Materials, 303(124450). https://doi.org/10.1016/j.conbuildmat.2021.124450

Brownlee, J. (2018a). A gentle introduction to k-fold cross-validation. Retrieved from https://machinelearningmastery.com/k-fold-cross-validation/

Brownlee, J. (2018b). Use weight regularization to reduce overfitting of deep learning models. Retrieved from https://machinelearningmastery.com/weight-regularization-to-reduce-overfitting-of-deep-learning-models/

Brownlee, J. (2019a). How to use data scaling improve deep learning model stability and performance. Retrieved from https://machinelearningmastery.com/how-to-improve-neural-network-stability-and-modeling-performance-with-data-scaling/

Brownlee, J. (2019b). How to choose a feature selection method for machine learning. Retrieved from https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/

Brownlee, J. (2020a). How to calculate feature importance with Python. Retrieved from https://machinelearningmastery.com/calculate-feature-importance-with-python/

Brownlee, J. (2020b). How to use power transforms for machine learning. Retrieved from://machinelearningmastery.com/power-transforms-with-scikit-learn/

Chong, B.W., Othman, R., Putra Jaya, R., Mohd Hasan, M.R., Sandu, A.V., Nabiałek, M., Jeż, B., Pietrusiewicz, P., Kwiatkowski, D., Postawa, P., & Abdullah, M.M.A.B. (2021). Design of experiment on concrete mechanical properties prediction: A critical review. Materials, 14, 1866. https://doi.org/10.3390/ma14081866

Derousseau, M., Laftchiev, E., Kasprzyk, J., Rajagopalan, B., & Srubar, W., III, (2019). A comparison of machine learning methods for predicting the compressive strength of field-placed concrete. Construction and Building Materials, 228, 116661. https://doi.org/10.1016/j.conbuildmat.2019.08.042

Fernando, J., & Walters, T. (2021). Correlation coefficient. Retrieved from https://www.investopedia.com/terms/c/correlationcoefficient.asp

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. Journal of Machine Learning Research, 20(177), 1-81.

Jee, N., Yoon, S., & Hongbum, C. (2004). Prediction of compressive strength of in-situ concrete based on mixture proportions. Journal of Asian Architecture and Building Engineering, 3, 9-16. https://doi.org/10.3130/jaabe.3.9

Lin, C.-J., & Wu, N.-J. (2021). An ANN model for predicting the compressive strength of concrete. Applied Sciences, 11(9), 3798. https://doi.org/10.3390/app11093798

Mancio, M., Moore, J., Brooks, Z., Monteiro, P., & Glaser, S. (2010). Instantaneous in-situ determination of water-cement ratio of fresh concrete. ACI Materials Journal, 107, 586-592.

Mishra, G. (2013). Factors affecting strength of concrete. Retrieved from https://theconstructor.org/concrete/factors-affecting-strength-of-concrete/6220/

Muliauwan, H., Prayogo, D., Gaby, G., & Harsono, K. (2020). Prediction of concrete compressive strength using artificial intelligence methods. Journal of Physics: Conference Series, 1625, 012018. https://doi.org/10.1088/1742-6596/1625/1/012018

Neha, Y., & Rajeev, C. (2015). Investigation of fresh concrete for quality control purpose using resistivity probe method. International Journal for Innovative Research in Science and Technology, 2, 77-81.

National Ready Mixed Concrete Association (NRMCA). (2014). CIP-35 testing compressive strength of concrete. Maryland, United States: National Ready Mixed Concrete Association (NRMCA).

Obla, K., Hong, R., Sherman, S., Bentz, D., & Jones, S. (2018). Relating the electrical resistance of fresh concrete to mixture proportions. Advances in Civil Engineering Materials, 7(1). https://doi.org/10.1520/ACEM20170126

Obla, K., Lobo, C., Hong, R., & Sherman, S. (2017). Evaluation of ASTM standard practice on measuring the electrical resistance of fresh concrete. Retrieved from https://www.nrmca.org/research_engineering/Documents/FreshResistanceWCMReportSep17.pdf

Patel, H. (2018). 11 factors affecting strength of concrete. Retrieved from https://gharpedia.com/blog/factors-that-affect-strength-of-concrete/

Pengcheng, L., Xianguo, W., Hongyu, C., & Tiemei, Z. (2020). Prediction of compressive strength of high-performance concrete by random forest algorithm. IOP Conference Series: Earth and Environmental Science, 552, 012020. https://doi.org/10.1088/1755-1315/552/1/012020

Poon, C.S., & Lam, C.S. (2008). The effect of aggregate-to-cement ratio and types of aggregates on the properties of pre-cast concrete blocks. Cement and Concrete Composites, 30, 283-289. https://doi.org/10.1016/j.cemconcomp.2007.10.005

Ren, Q., Ding, L., Dai, X., Jiang, Z., & De Schutter, G. (2021). Prediction of compressive strength of concrete with manufactured sand by ensemble classification and regression tree method. Journal of Materials in Civil Engineering, 33, 04021135. https://doi.org/10.1061/(ASCE)MT.1943-5533.0003741

Silva, P., Moita, G., & Arruda, V. (2020). Machine learning techniques to predict the compressive strength of concrete. Revista Internacional de Métodos Numéricos Para Cálculo y Diseño En Ingeniería, 36, 1-9. https://doi.org/10.23967/j.rimni.2020.09.008

Tavares, D., Ribeiro, D., Yanagi Junior, T., Lacerda, W., Tiradentes, E., Teixeira, R., & Garcia, H. (2020). Computational intelligence applied in the prediction of the compressive strength of Portland cement concrete. Congresso Brasileiro de Automática, 2(1). https://doi.org/10.48011/asba.v2i1.1073