Research Article

Prediction of Later-Age Concrete Compressive Strength Using Feedforward Neural Network

Thuy-Anh Nguyen, Hai-Bang Ly, Hai-Van Thi Mai, and Van Quan Tran

University of Transport Technology, Hanoi 100000, Vietnam

Correspondence should be addressed to Hai-Van Thi Mai; vanmth@utt.edu.vn

Received 26 June 2020; Revised 11 August 2020; Accepted 18 August 2020; Published 8 September 2020

Guest Editor: Hainian Wang

Copyright © 2020 Thuy-Anh Nguyen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate prediction of the concrete compressive strength is an important task that helps to avoid costly and time-consuming experiments. Notably, the determination of the later-age concrete compressive strength is more difficult due to the time required to perform experiments. Therefore, predicting the compressive strength of later-age concrete is crucial in specific applications. In this investigation, an approach using a feedforward neural network (FNN) machine learning algorithm was proposed to predict the compressive strength of later-age concrete. The proposed model was fully evaluated in terms of performance and prediction capability over statistical results of 1000 simulations under a random sampling effect. The results showed that the proposed algorithm was an excellent predictor and might be useful for engineers to avoid time-consuming experiments with the statistical performance indicators, namely, the Pearson correlation coefficient (R), root-mean-squared error (RMSE), and mean squared error (MAE) for the training and testing parts of 0.9861, 2.1501, 1.5650 and 0.9792, 2.8510, 2.1361, respectively. The results also indicated that the FNN model was superior to classical machine learning algorithms such as random forest and Gaussian process regression, as well as empirical formulations proposed in the literature.

1. Introduction

In the field of civil engineering, concrete plays a significant role and is the most important construction material [1–7]. Concrete is a kind of artificial stone material made by molding and solidifying an appropriate mixture of cement, water, coarse, and fine aggregate. To further enhance the properties of concrete, several chemical or mineral admixtures such as fly ash, blast furnace slag, silica fume, metakaolin, and superplasticizer could be added into the mixture. Compressive strength is the most important factor that reflects the quality of a concrete mix design [8]. The concrete compressive strength is usually characterized by estimating the values at 28 days. The compressive strength directly depends on the ingredients of concrete and their relative amount [9], making the procedure of selecting appropriate mixture components critical. Therefore, knowing the potential compressive strength of later-age concrete at the earliest possible time after curing, especially in the first hours, can be helpful for necessary adjustment.

In recent decades, the development of suitable tools to obtain the desired compressive strength has been the subject of numerous researches. Abrams’ law is the well-known concept in the field of civil engineering for calculating the compressive strength of concrete with two inputs, namely, cement and water. The results of Abrams’ law showed that an increase of compressive strength is inversely proportional to the water-to-cement ratio. Besides, the compressive strength of concrete also depends on the content and nature of the mix design, the curing condition, and the age of specimens [10–12]. Apart from Abrams’ law, Hasan and Kabir [8] and ACI 209-71 [13] have proposed empirical equations to determine the compressive strength of concrete at any age (Table 1). However, the empirical methods could not take into account all factors that affect the concrete compressive strength, such as supplementary cementitious materials or...
the mixture components. Therefore, these methods cannot correctly estimate the compressive strength of concrete using supplementary cementitious materials. Alternative approaches, thus, need to be investigated to predict the later-age compressive strength of concrete accurately.

Recently, development and application of machine learning techniques in the civil engineering sector have gained extensive attention [14]. Taking advantage of an extensive experimental database, machine learning algorithms exhibit the ability to simplify classical approaches, such as experimental or numerical simulation. During the last four decades, machine learning techniques have been proven to successfully solve many problems of civil engineering such as coastal engineering, construction engineering [15], earthquake engineering, geotechnical engineering, environmental engineering, structural engineering [16, 17], and materials science engineering [18].

Moreover, taking into account numerous input factors is an essential advantage of machine learning techniques. Therefore, machine learning technique appears to be a potential approach to predict the later-age compressive strength of concrete. Among the machine learning algorithms that are currently being utilized, the feedforward neural network (FNN) is the most popular model due to the structural flexibility, excellent prediction performance, and the availability of a significant amount of training algorithms [19]. With the ability to deal with complex functions, FNN is an excellent predictor in many studies, such as compressive strength of concrete [20] or bearing capacity of steel tube columns [21].

In this paper, the FNN algorithm was developed and used to predict the later-age compressive strength of concrete. For this aim, a database containing 190 samples, gathered from the available literature, was used to train and validate the FNN algorithm. The quality of the FNN model was assessed by common statistical measurements such as root-mean-square error (RMSE), mean absolute error (MAE), and the Pearson correlation coefficient (R). The paper is structured as follows: the introduction section is presented in Section 1. In Section 2, the machine learning feedforward neural network (FNN) is introduced and fundamental elements are focused. Section 3 presents the database used for the development of FNN, as well as the statistical information of input and output variables. Section 4 shows the results obtained by FNN and emphasizes the superior prediction accuracy by comparing with classical machine learning models such as random forest and Gaussian process regression. Finally, Section 5 provides some conclusions and perspectives.

### 2. Machine Learning Method

#### 2.1. Feedforward Neural Network

Feedforward neural network (FNN) is the most commonly used neural network in the literature [22]. It could be considered as one of the first and most successful machine learning algorithms. FNN could also be called a multilayer perceptron (MLP) or in a simple manner as neural networks. The FNN structure contains four main compositions: the input layer, output layer, one or more hidden layers, and the weights associated with each neuron in the hidden layer(s) [23]. In the input layer, it contains a certain number of neurons, which corresponds to the number of features in the dataset. The input layer receives the information related to the inputs and passes such information to the second layer, the hidden layer(s). The hidden layer is located between the input and output layers, which contains many neurons to apply the transformation from the input layer to the output layer. Each neuron in the hidden layer(s) is associated with a weight, representing the amplitude of a connection between two neurons. The output layer is the target of the problem or the feature that we want to predict [24].

The principle of the process of FNN can be summarized in 3 steps as follows: (i) multiplication of weights and inputs, where each feature in the input layer is multiplied by the weight associated; (ii) biases added: the product in the previous step is added, each with a bias; and (iii) activation step: an activation function is applied to map the features in the input layer to the output layer (Figure 1). Finally, the desired output is obtained by summing up all the results of the previous steps (for each neuron). The difference, or the error, of the predicted and actual outputs, is calculated and called a loss function, which measures the algorithm performance.

The effectiveness of FNN has been proven in many works [25, 26], especially in material modeling [20, 27]. Therefore, FNN with ten neurons is proposed to predict the later-age compressive strength of concrete in this study. Detail information on the FNN, as well as all equations governing the algorithm, could be found in the literature [28].

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Equation</th>
<th>Symbol</th>
<th>Time-dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abrams’ law</td>
<td>$S = (A/R^{W(C)})$</td>
<td>$S$: compressive strength; $A$, $B$: constant; $W/C$: water/cement ratio</td>
<td>No</td>
</tr>
<tr>
<td>Hasan and Kabir [8]</td>
<td>$S_n = (pD_n/Dn + q)$</td>
<td>$p$, $q$: constant; $n$: day; $D_n$: number of days</td>
<td>Yes</td>
</tr>
<tr>
<td>ACI 209-71 [13]</td>
<td>$S_{28} = (t/(a + bt))S_{28d}$</td>
<td>$a$, $b$: constant; $t$: number of days; $S_{28d}$: compressive strength at 28 days</td>
<td>Yes</td>
</tr>
</tbody>
</table>

#### 2.2. Performance Criteria

In this study, the performance of FNN is evaluated by common statistical measurements such as the Pearson correlation coefficient ($R$), root-mean-square error (RMSE), and mean absolute error (MAE). Typically, the values of $R$ are in the $[-1, 1]$ range, and better prediction accuracy is obtained when the absolute value of $R$ is close to 1. Oppositely, lower values of RMSE and MAE designate a
higher performance of the given machine learning algorithm. The expressions of the indicators are as follows [29, 30]:

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} (u_i - \bar{u}_i),
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (u_i - \bar{u}_i)^2},
\]

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (u_i - \bar{u}_i)^2}{\sum_{i=1}^{n} (u_i - \bar{u})^2},
\]

where \(\bar{u}_i\) is the predicted output value, \(u_i\) is the measured target value, and \(n\) is the number of samples.

### 3. Database Collection

In this study, a database containing 190 samples was collected from the literature [31–35]. The input parameters of the concrete compressive strength database include the content of cement (denoted as C), the content of blast furnace slag (denoted as BFS), the content of fly ash (denoted as FA), water content (denoted as W), superplasticizer (denoted as SP), aggregate content including coarse aggregates (denoted as CA) and fine aggregates (denoted as FA), and testing age (denoted as A). The considered output is the later-age compressive strength (LACS) of concrete.

For the statistical information of the database, Table 2 presents the min, average, max, median, and standard deviation values (denoted as SD) of all the parameters used in this study.

The database containing 190 samples is divided randomly into two parts, namely, the training part (70% of the total samples) and the testing part (30% of the remaining samples). The training part is used to construct the FNN model, whereas the testing subset was dedicated to the assessment of the accuracy of the model.

### 4. Results and Discussion

#### 4.1. Performance of FNN

The samples of the training part greatly affect the accuracy of the machine learning black box [36] so that the sampling technique is important in the construction phase of the model. A certain number of samples are randomly taken to be in the training part. Besides, there is systematic sampling, stratified sampling, or cluster sampling, belonging to the probability sampling technique. In this study, simple random sampling was chosen as it was an effective way to select samples. The performance of FNN is evaluated throughout the sampling technique in this section.

A number of 1000 simulations were performed in shuffling the indexes of the database, and 70% of data were taken to construct the training part. The results of 1000 simulations are given by the histograms of values \(R\) (Figure 2), RMSE (Figure 3), and MAE (Figure 4) for the training and testing parts. The values of \(R\) of the FNN algorithm showed a strong prediction capability of LACS as the minimum values of \(R\) were in the \(R = 0.7–1\) range. Thus, a strong correlation was observed between the predicted output given by FNN and the experimental results.

The training dataset exhibited greater accuracy than the testing dataset, which was helpful in preventing overfitting problems. A high concentration of \(R\) close to 1 was observed,
and $R = 0.94$ was the most frequent result obtained for the testing dataset.

Similarly, the RMSE and MAE values showed complementary information on the observations. The most frequent values of RMSE and MAE were 2.5 and 2.1 for the training dataset, whereas those of the testing dataset were 5 and 3.9, respectively. In all the cases, no extreme values were observed for the training dataset, showing that the prediction performance of FNN was satisfactory. Table 3 shows the values of $R$, RMSE, and MAE for 1000 simulations performed in this study, along with the standard deviation values associated with each case. The average values of $R$ were 0.9391 and 0.9350 for the training and testing datasets,

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Average</th>
<th>Max</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>$R$</td>
<td>0.8179</td>
<td>0.9391</td>
<td>0.9677</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>3.1481</td>
<td>4.1988</td>
<td>7.1676</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>2.4860</td>
<td>3.0472</td>
<td>3.8312</td>
</tr>
<tr>
<td>Testing</td>
<td>$R$</td>
<td>0.4379</td>
<td>0.9350</td>
<td>0.9781</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>2.7580</td>
<td>4.3342</td>
<td>11.5334</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>2.2113</td>
<td>3.2552</td>
<td>7.5907</td>
</tr>
</tbody>
</table>

SD = standard deviation.
respectively. Overall, FNN is a promising algorithm to predict the LACS of concrete.

4.2. FNN Typical Configuration Result. The results of a typical FNN configuration are presented in this section. Figure 5 shows the comparison between LACS results of the FNN model and the experimental data for the training dataset. The comparison shows that the predicted results were very close to the experimental results. Figure 6 shows that the FNN algorithm can successfully predict the experimental compressive strength for the testing dataset. The histograms of error with respect to the training and testing parts are displayed in Figures 7 and 8, respectively. The mean errors were 0.08317 for the training part and 1.3724 for the testing part. The standard deviations of error values were computed as 2.1566 and 2.5211 for the training and testing datasets, respectively. It can be seen that the accuracy of the training part was slightly superior to the testing part, but both results showed excellent accuracy.

Finally, the relationships of the actual and predicted data were given in the form of regression plots in Figure 9. The values of R were 0.9861 and 0.9792 for the training and testing datasets, respectively. Two equations related to such relationships were given in the graph, where "\( Y = 0.97X + 1.3 \)" was the best fit line for the training set and "\( Y = 0.98X - 0.4 \)" was the best fit linear regression for the testing set.
The FNN algorithm showed an excellent performance in predicting the LACS of concrete. It is worth noticing that the values of $R$ could reach 0.9792 with the use of a simple FNN structure FNN (8-10-1), representing one hidden layer and ten neurons in such hidden layer.

For comparison purpose, the performance of FNN algorithm was shown superior to other benchmark machine learning algorithms, such as random forest (RF) and Gaussian process regression (GPR). The detail and development of RF and GPR algorithms are described in [37, 38], respectively. The regression graphs for the training and testing datasets of RF and GPR algorithms are shown in Figure 10. A summarized performance of FNN, RF, and GPR algorithms is shown in Table 4.

Overall, the FNN algorithm exhibited the best prediction performance with respect to RMSE, MAE, and $R$, compared with RF and GPR algorithms. Therefore, the FNN algorithm appears as a promising numerical tool to predict the later-age compressive strength of concrete.

## 5. Conclusion

In this study, the FNN algorithm was introduced to predict the later-age compressive strength of concrete containing cement, fly ash, blast furnace slag, water, and aggregates. A number of 190 experimental results were collected to construct the FNN model. A number of 70% of data was randomly chosen for the training phase of FNN, and 30% of data was chosen for the validation phase of the constructed FNN model. To fully evaluate the performance of FNN, a number of 1000 simulations were performed using a random sampling technique, and the evaluation of FNN model was carried out by common statistical measurements.

The results showed that the FNN was a promising algorithm to predict the LACS of concrete, with the average values of 0.9350, 4.3342, and 3.2552 for $R$, RMSE, and MAE, respectively, for the testing part. The most accurate FNN model exhibited an excellent prediction performance with $R = 0.9792$, RMSE = 2.8510, and MAE = 2.1361. These results were shown superior to other classical machine learning algorithms, namely, RF and GPR. The results could help in constructing a reliable soft computing tool to predict accurately and quickly the later-age compressive strength of concrete. Once the tool is built, the forecast process would take a millisecond to estimate the compressive strength of the given mix design.

### Table 4: Performance analysis of FNN, RF, and GPR algorithms.

<table>
<thead>
<tr>
<th></th>
<th>FNN</th>
<th>RF</th>
<th>GPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training dataset</td>
<td>$\begin{array}{l} R \ RMSE \ MAE \end{array}$</td>
<td>$\begin{array}{l} 0.9861 \ 2.1501 \ 1.5650 \end{array}$</td>
<td>$\begin{array}{l} 0.9654 \ 3.6900 \ 2.7728 \end{array}$</td>
</tr>
<tr>
<td>Testing dataset</td>
<td>$\begin{array}{l} R \ RMSE \ MAE \end{array}$</td>
<td>$\begin{array}{l} 0.9792 \ 2.8510 \ 2.1361 \end{array}$</td>
<td>$\begin{array}{l} 0.9006 \ 6.2663 \ 5.0560 \end{array}$</td>
</tr>
</tbody>
</table>
Data Availability

The processed data are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References


