JOURNAL OF INTERNATIONAL CRISIS AND RISK COMMUNICATION RESEARCH
ISSN: 2576-0017
2024, VOL 7, NO S7

Sequential Machine Learning Model: Comparison of Optimizers for Concrete Compressive Strength Prediction

Bernilla Rodriguez[1], Ruiz-Pico, Angel Antonio[2]

1. Engineering Department, Universidad Católica Santo Toribio de Mogrovejo, Perú
2. Engineering Department, Universidad Tecnológica del Perú, Perú

***Corresponding author E-mail:aaruizpico@gmail.com**

Abstract
Concrete is the most widely used construction material today due to its exceptional ability to withstand compressive forces, commonly referred to as CS. Determining the value of CS for concrete involves conducting various tests, with the uniaxial compression test on concrete specimens being the most used, assessing strength at different time intervals. This research focuses on predicting the value of concrete's CS at 28 days using 16 predictive models based on artificial neural networks, each employing different optimizers, including Adam, Adamax, Rmsprop, and Nadam. Input data includes aggregate properties, cement type, and the proportions of its components, such as water, cement, and aggregates. The neural networks were built using Google's TensorFlow, with two hidden layers that vary in the number of neurons in the first hidden layer (13, 16, 19, 22) and 8 neurons in the second hidden layer. Training was conducted over 400 epochs.The comparison of optimizers reveals that Nadam exhibits the best performance in both the training and prediction stages, with R^2 values of 0.83 and 0.87 respectively. The increase in the number of input variables in the prediction results in higher accuracy in predicting CS.

1. Introduction

Undoubtedly, concrete's CS is a key mechanical characteristic in the construction industry. The durability and load-bearing capacity of a structure largely depends on this property[1]. To obtain a specific CS value for a construction project, a mix design must be performed, which involves the use of aggregates, cement, water, and sometimes additives[2]. In the list of the most consumed products on our planet, concrete ranks second, surpassed only by the vital resource of fresh water [3].

Cement is the preeminent synthetic material globally in building construction[4]. The cement industry plays a fundamental role in pursuing carbon neutrality, a goal aligned with the principles of the Paris Agreement for 2050. This purpose gains even more relevance considering the notable CO_2 emissions associated with global cement production, a subject that has been studied and a cause for concern, as documented by[5]. Furthermore, this production involves intensive use of natural resources and an increase in greenhouse gas emissions[6]. Approximately 50% of GHG emissions come from calcination. The presence of concrete in our everyday environment emphasizes the need to address its production and utilization sustainably and efficiently, with the aim of minimizing its environmental impact and ensuring durability.

CS of concrete is determined through various tests, which can be classified as either destructive or non-destructive. Examples of these tests include uniaxial compression, splitting compression, ultrasonic, Windsor, and sclerometer [7]. For the purposes of this research, the concrete specimens underwent destructive testing, specifically simple compression, which is a widely used method for evaluating concrete strength. This test involves gradually applying axial loads to the specimen until it fails, providing an accurate measure of compressive strength. Concrete is ideally expected to reach full strength in 28 days[8] . However, in practice, test results often show variations from the design values, which can pose problems in the construction industry as it may affect the safety and longevity of structures.

In addition to result variability, another relevant concern is the disposal of used test specimens. After testing, these specimens are often discarded outdoors, which is not only a waste of material but also contributes to environmental pollution. This environmental dilemma has driven the search for more efficient and eco-friendly approaches, to limiting CO_2 emissions[9].

The Deep Learning (DL) domain has recently undergone a remarkable revolution. This sector, a subset of the broader machine learning field, has facilitated the development of highly advanced predictive modelsand has enabled the development of predictive models exceptionally. These models not only offer precise and dependable predictions but have also opened the door to addressing intricate challenges that were once deemed insurmountable[10]. Artificial Neural Networks (ANNs) play a pivotal role in the realm of deep learning, as they are purpose-built to model and acquire hierarchical data representations. This intrinsic capacity empowers them to adeptly process exceptionally intricate information, enabling the execution of diverse tasks, including but not limited to pattern recognition, classification, regression, text generation, computer vision, and more.

In this context, ANNs emerge as an innovative solution. These complex algorithmic structures can identify patterns in highly intricate data. A recent study by[11]emphasized the potential of ANNs in predicting CS at 28 days in carbon nanotube-reinforced cement composites. This study is notable for considering the sample size, revealing its profound influence on CS at 28 days.

Furthermore,[12] emphasized the capacity of machine learning algorithms by employing XGBoost to predict compressive strength in clay calcined cement. Their approach utilized 14 input parameters, achieving a low absolute error of only 2.5 MPa. On the other hand, [13]used the Bayesian linear regression algorithm for prediction, having an extensive dataset of 162 mixed designs from articles published between 2000 and 2020, achieving acceptable accuracyand [14]used the Levenberg-Marquardt (LM) algorithm to develop the model. This was constructed based on the results derived from laboratory experiments analogous to those carried out in the present investigation.

The crucial limitation in the present investigation is that the values obtained for the CS of concrete were only at 28 days. In the construction field, it is essential to know the CS of concrete at earlier intervals, such as 3 days, or at other critical moments. This is essential to evaluate the ability of the structure to withstand additional loads, thus allowing optimization of time and associated costs. The ability to determine the concrete strength at early stages can be crucial to make quick and efficient decisions during the construction process; it is relevant to note that the data collected for the CS of concrete is in a specific range, varying between 150 kg/cm² and 400 kg/cm², as detailed in Table 2. This variability in values may introduce uncertainty in the overall applicability of the results and require additional considerations when extrapolating the findings to different conditions and concrete mixes in construction practice. In addition, it should be noted that environmental conditions, such as temperature, humidity, were not incorporated as input data, since the specimens were prepared under controlled laboratory conditions. This omission may influence the accuracy of the model when applied to different conditions, such as those present in the construction site, thus representing a significant gap in the research.

Taking all the above into account, this study will delve into the application of ANNs in predicting concrete compressive strength, considering the properties of coarse and fine aggregates, cement type and specific weight, and constituent quantities as input data, with CS as the sole output data. To accomplish this, a sequential model using the TensorFlow library will be implemented. The primary purpose of this research lies in predicting concrete CS, which translates into a notable reduction in environmental impact, by significantly reducing the need for costly physical tests and effectively contributing to material waste reduction, thereby promoting more efficient and environmentally sustainable construction practices. Considering all the above, this study will delve into the application of ANNs in the prediction of concrete compressive strength, considering as input data the properties of coarse and fine aggregates, the type and specific weight of cement and the amounts of constituent, with CS being the only output data. For this purpose, a sequential model will be implemented using the Python programming language and the TensorFlow library. The main objective of this research lies in the prediction of the CS of concrete, which translates into a significant reduction of the environmental impact by significantly reducing the need for costly physical testing and effectively contributing to the reduction of material waste.

1. Research method

1.1. Data collection

A quantitative methodology was employed in this research. A total of 465 mix design data used in the production of concrete were compiled, along with the results of the uniaxial compression tests performed on the specimens. To obtain these results, a hydraulic compression press was used. The data in this study were collected from laboratories located in northern Peru. The aggregates used in the study meet the quality requirements established by international standards. They were tested for suitability for use in concrete manufacture, complying with the standard. The aggregates are siliceous in nature and come from 9 local quarries.Table 3 and Figure 2 provide detailed information on the types of cement used in the study.The collected data includes, for cement: (1) quantity, (2) type and specific gravity; for water: (1) quantity; for fine aggregate: (1) quantity, (2) fineness modulus, (3) specific gravity, (4) loose dry weight, (5) compacted dry weight, (6) humidity percentage, (7) absorption percentage; and for coarse aggregate: (1) quantity, (2) fineness modulus, (3) nominal maximum size, (4)specific gravity, (5) loose dry weight, (6) compacted dry weight, (7) humidity percentage, (8) absorption percentage. The initial data storage was in an Excel file. After implementing the treatments outlined in section 2.1.1, the data will be transmitted to a database, specifically MySQL. Of the data collected, 90% will be used for training of the deep neural network, and the remainder will be used for testing purposes.

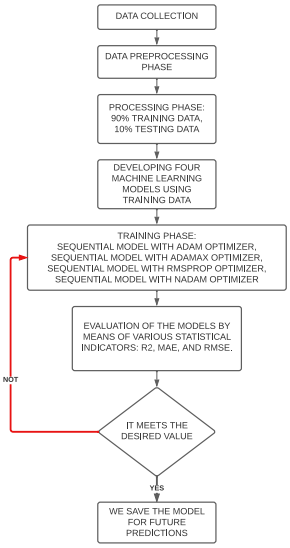


Figure 1 Research Flowchart

Table 1 briefly describes each parameter, its unit of measure, and its function, whether as a predictive model input or output; And Table 2 shows a detailed distribution of the amount of data collected based on different ranges of (CS) in kg/cm².

Table 1Characteristics of input and output elements.

Parameters	Units	Characteristics
Compressive strength (CS)	kg/cm²	Output
Cement (C)	Kg	Input
Cement type (CT)	-	Input
Specific gravity of cement (SGC)	kg/m³	Input
Water (W)	Lt	Input
Fine Aggregate (FA)	Kg	Input
Fineness modules of the fine aggregate (FM-FA)	-	Input
Specific gravity of fine aggregate (SG-FA)	kg/m³	Input
Loose dry weight of fine aggregate (LDW-FA)	kg/m³	Input
Compacted dry weight of fine aggregate (CDW-FA)	kg/m³	Input
Moisture content of fine aggregate (MC-FA)	%	Input
Absorption of fine aggregate (A-FA)	%	Input
Coarse aggregate (CA)	Kg	Input
Fineness modules of the coarse aggregate (FM-CA)	-	Input
Nominal maximum size (NMS)	Inch	Input
Specific gravity of coarse aggregate (SG-CA)	kg/m³	Input
Loose dry weight of coarse aggregate (LDW-CA)	kg/m³	Input
Compacted dry weight of coarse aggregate (CDW-CA)	kg/m³	Input
Moisture content of coarse aggregate (MC-CA)	%	Input
Absorption of coarse aggregate (A-CA)	%	Input

Table2 Number of data points in a specific CS range.

Number	CS range (kg/cm²)	Quantity of data
1	150 – 200	80
2	200 – 250	101
3	250 – 300	124
4	300 – 350	118
5	350 - 400	42
Total		465

1.1.1. Data processing

A preliminary treatment of the obtained data was performed before inputting it into the neural network. It was verified that the laboratory-obtained CS falls within the minimum value allowed by [15] concerning the design CS. All collected data meet the requirement imposed by[15].

The representation of cement types in the dataset was initially recorded as text, including cement type I, cement type HS, cement type MS, and cement type V. However, to process this data correctly, it needs to be converted into a numerical value.

To address this issue, the decision was made to assign a numerical value to each type of cement.

Table 3 Number of samples for mechanical testing of unit

Cement Type	Value
I	1
HS	2
MS	3
V	4

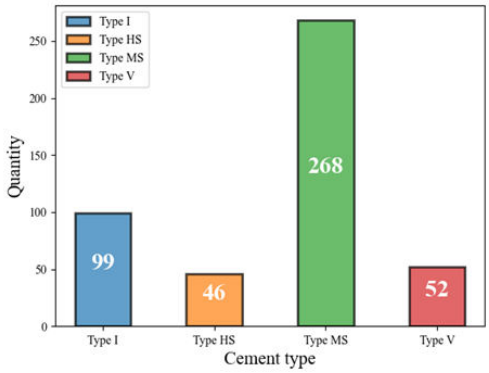


Figure 2 Data collected by type of cement.

The representation of the maximum nominal size (MNS) of coarse aggregate was in decimal values with the unit of measurement in inches, and the collected data includes two MNS values, 1/2 inch and 3/4 inch.

Table 4 Maximum nominal size

MNS (inch)	Value
1/2	0.5
3/4	0.75

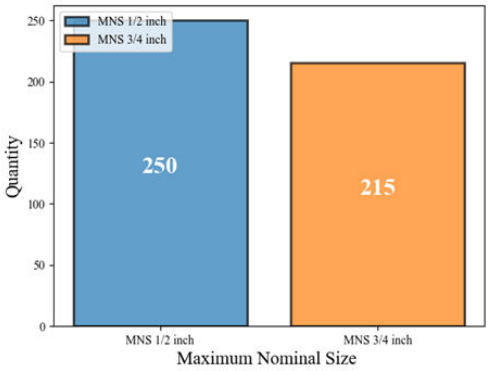


Figure 3 Data collected by MNS.

1.2. Neural networks creation

In recent years, Artificial Neural Networks (ANNs) have been widely used for the prediction of concrete CS[16]. These methods allow the creation of predictive models from known data and outcomes. This study describes the methodological approach used to develop the predictive model.

For the creation of the neural network, we will use TensorFlow, which is a library provided by Google for Machine Learning and is free to use. We imported the necessary libraries, such as Mysql.connector and TensorFlow.

In the TensorFlow library, we have a class called Sequential, with which we can create neural networks of Multilayer Perceptron (MLP) architecture.

The development of the Artificial Neural Network (ANN) was guided by experimental exploration[17]. Initially, the effectiveness of a single hidden layer was evaluated, however, it was determined that this configuration was not sufficient to capture the inherent complexity of the underlying relationships in the problem. On the other hand, by including more

than two hidden layers, unnecessary complexity was introduced. The final choice of two hidden layers was considered an optimal tradeoff, allowing the network to learn complex patterns without incurring overfitting, as explained in [18] [19], the inclusion of an excess of hidden layers or hidden neurons in the neural network increases the propensity to over-fitting, leading to over-generalization of the training data and, consequently, generating a higher error in the predictions. The selection of the number of neurons in each layer was based on literature review and considerations of the number of neurons in the input layer. The resulting ANN configuration consists of an input layer with 19 neurons. In addition, two hidden layers are implemented; the first one varies in the number of neurons (13, 16, 19 and 22), facilitating comparisons between optimizers, while the second hidden layer consists of 8 neurons. The output layer is reduced to a single neuron, culminating in a balanced network architecture tailored to the complexities of concrete compressive strength prediction.
Rectified Linear Unit (Relu) has been selected as the activation function in the hidden layers due to its ability to efficiently handle nonlinearity in the data. The Relu has proven to be effective in capturing complex patterns and contributes to faster convergence during training.
As for the loss function, Mean Squared Error (MSE) Loss has been chosen. This choice is based on the nature of our regression problem, where we seek to minimize the quadratic discrepancy between network predictions and actual compressive strength values.
This configuration has been designed to evaluate the performance of the neural network in predicting the CS of concrete at 28 days and determine which of the options for neurons in the first hidden layer, along with the type of optimizer, offers the best performance.

1.2.1. Multilayer Perceptron (MLP)

One of the distinct features of artificial neural networks is their capability to learn and approximate any function from the provided data. This enables them to comprehend complex relationships derived from input information. A specific neural network architecture, termed the Multilayer Perceptron (MLP), is comprised of an input vector, one or multiple hidden layers, and an output layer commonly referred to as the output vector[20].

1.2.2. Supervised learning

The dataset is labeled, and characteristics associated with each label are provided, with the aim of creating a model capable of making predictions with new data [21]. To predict CS, the CS values obtained in the laboratory for each mix design were labeled.

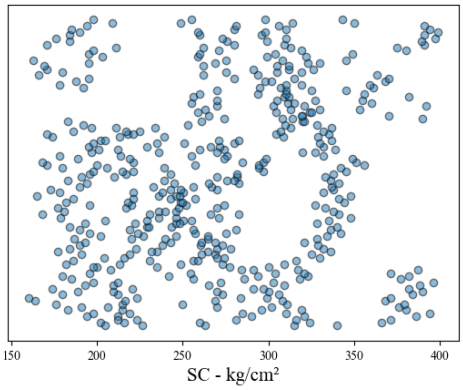


Figure 4 Dispersion of the CS

1.2.3. Optimizers

Algorithms used to adjust the weights or biases to minimize the loss function during the training process [22]. Adam, it is a first-order gradient-based optimization algorithm for stochastic objective functions. The method is suitable for non-stationary objectives [23]; Adamax, it is a variant of the Adam optimizer, based on first-order gradients. The difference from Adam is that it is based on the infinity norm, which means it considers the maximum magnitude of gradient values [23]; RMSprop, The optimization algorithm uses simple momentum, which aims to adjust the learning rate adaptively by maintaining a moving average of the squares of the previous gradients [24]; Nadam, This algorithm is the incorporation of Nesterov momentum into Adam. The inclusion of this momentum helps prevent oscillations and accelerates convergence [25].
To determine the best optimizer for the network, the trial-and-error technique will be employed, a strategy commonly used in the optimization of neural networks. The choice of optimizer plays a fundamental role in the model's performance, as it directly affects the process of adjusting the network's weights during training.
The trial-and-error technique involves experimenting with various available optimizers, such as Adam, Adamax, RMSprop, or Nadam, to assess their influence on convergence speed and prediction accuracy. Each optimizer has its characteristics and underlying algorithms that may be more effective for different types of data and problems. During this process, multiple training runs of the neural network are executed using different optimizers, and the results are compared based on metrics detailed in section 2.3. It's important to highlight that the choice of optimizer can significantly impact the efficiency of the training process and the network's ability to reach an optimal solution.

Therefore, this trial-and-error approach allows for identifying which of the available optimizers maximizes the network's performance and provides more accurate predictions.

1.2.4. Activation functions

Employed in every neuron of a neural network, activation functions introduce nonlinearity into the model, allowing the neural network to understand and represent intricate connections and patterns within the data.

$$f(x) = \frac{1}{1+e^{-x}} f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

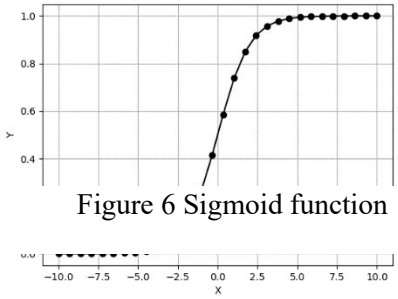
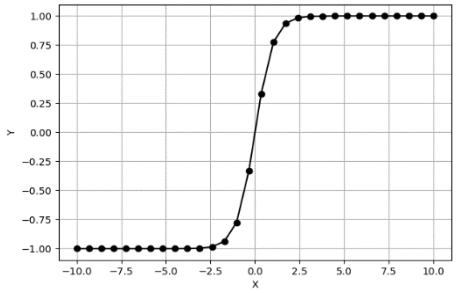


Figure 6 Sigmoid function



n

$$f(x) = \max(0, x)$$

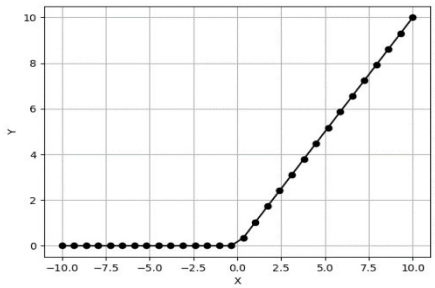


Figure 7 Relu function

1.2.5. Loss functions

Binary Cross Entropy Lossused in binary classification, the desired output is given as 0 and 1.

Positive Class, Y=1

$$Loss = -\log\left(Y_{pred}\right)$$

Negative Class Y=0

$$Loss = -\log\left(1 - Y_{pred}\right)$$

Categorical Cross Entropy Lossis employed in multiclass classification tasks.

$$Loss = -\sum_{i=1}^{N} Y_{truei} * \log\left(Y_{predi}\right)$$

Mean Squared Error Loss, used in regression problems, where the goal is to minimize the squared difference between the value predicted by the network and the actual value.

$$Loss = \frac{1}{n} * \sum_{i=0}^{n} (Y_i - Y_{predi})^2$$

1.3. Criteria for evaluation

The criteria used to evaluate the quality of the artificial neural network model include the coefficient of determination (R2), the mean absolute error (MAE), and the root mean square error (RMSE).

$$R^2 = 1 - \frac{\sum(y_i - x_i)^2}{\sum(y_i - \mu_y)} \quad (1)$$

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}} \quad (3)$$

2. Results and discussion

2.1. Training

A grand total of sixteen sequential models were created, with four models designated for each optimizer in the comparative analysis. These four models exhibited variations in the number of neurons within the initial hidden layer, offering a comprehensive evaluation of each optimizer's effectiveness with diverse neural network setups. Every neural network underwent training over a span of four hundred epochs.

2.1.1. Adam optimizer

Table 5 displays the Adam optimizer's efficiency during the training phase for predicting CS, with varying configurations of neurons in the 1st layer. The optimal configuration involved sixteen neurons in that layer, resulting in an R2, MAE, and RMSE of 0.74, 23.96 kg/cm², and 30.21 kg/cm², respectively.

Table 5Performance of the Adam Optimizer for CS Prediction

Input Layer	Hidden Layers		Output Layer		
Number of Neurons	Number of Neurons in the First Layer	Number of Neurons in the Second Layer	Number of Neurons	Criteria	Adam
19	13	8	1	R2	0.64
				MAE (kg/cm²)	29.13
				RMSE (kg/cm²)	35.92
19	16	8	1	R2	0.74
				MAE (kg/cm²)	23.96
				RMSE (kg/cm²)	30.21
19	19	8	1	R2	0.72
				MAE (kg/cm²)	24.92
				RMSE (kg/cm²)	31.80
19	22	8	1	R2	0.72
				MAE (kg/cm²)	25.05
				RMSE (kg/cm²)	31.91

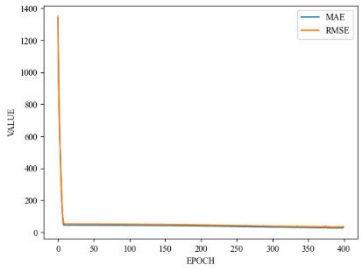


Figure 8Adam Optimizer - thirteen neurons.

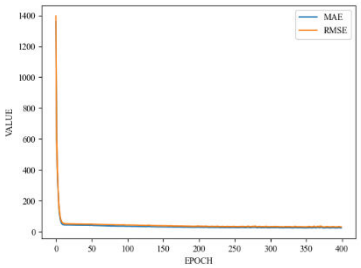


Figure 10Adam Optimizer - nineteen neurons.

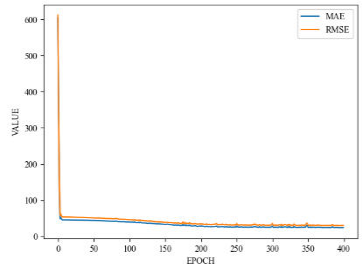


Figure 9Adam Optimizer - sixteen neurons.

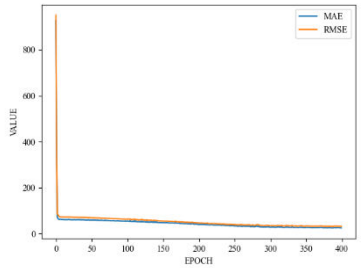


Figure 11Adam Optimizer - twenty-two neurons.

2.1.2. Adamax optimizer

Table 6 displays the Adamax optimizer's efficiency during the training phase for predicting CS, with varying configurations of neurons in the 1st layer. The optimal configuration involved sixteen neurons in that layer, resulting in an R2, MAE, and RMSE of 0.74, 24.15 kg/cm², and 30.51 kg/cm², respectively.

Table6Performance of the Adamax Optimizer for CS Prediction

Input Layer	Hidden Layers		Output Layer		
Number of Neurons	Number of Neurons in the First Layer	Number of Neurons in the Second Layer	Number of Neurons	Criteria	Adamax
19	13	8	1	R2	0.70
				MAE (kg/cm²)	26.36
				RMSE (kg/cm²)	32.98
19	16	8	1	R2	0.74
				MAE (kg/cm²)	24.15
				RMSE (kg/cm²)	30.51
19	19	8	1	R2	0.68
				MAE (kg/cm²)	26.76
				RMSE (kg/cm²)	33.98
19	22	8	1	R2	0.73
				MAE (kg/cm²)	25.01
				RMSE (kg/cm²)	31.19

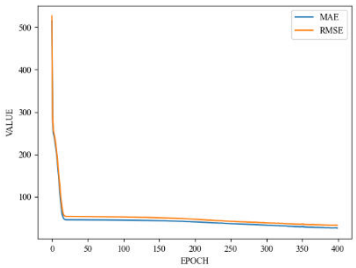


Figure 12Adamax Optimizer - thirteen neurons.

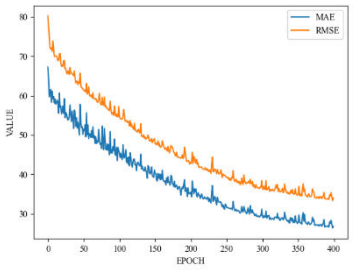


Figure 14Adamax Optimizer - nineteen neurons.

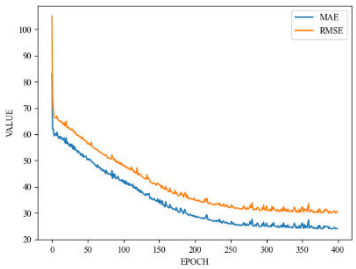


Figure 13Adamax Optimizer - sixteen neurons

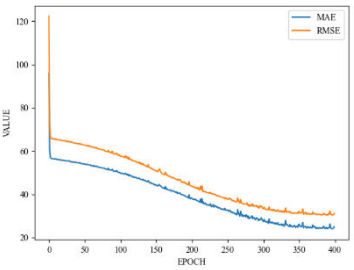


Figure 15Adamax Optimizer - twenty-two neurons.

2.1.3. Rmsprop optimizer

Table 7 displays the Rmsprop optimizer's efficiency during the training phase for predicting CS, with varying configurations of neurons in the 1st layer. The optimal configuration involved thirteen neurons in that layer, resulting in an R2, MAE, and RMSE of 0.74, 22.20 kg/cm², and 28.29 kg/cm², respectively.

Tabla7Performance of the Rmsprop Optimizer for CS Prediction

Input Layer	Hidden Layers		Output Layer		
Number of Neurons	Number of Neurons in the First Layer	Number of Neurons in the Second Layer	Number of Neurons	Criteria	Rmsprop
19	13	8	1	R2	0.78
				MAE (kg/cm²)	22.20
				RMSE (kg/cm²)	28.29
19	16	8	1	R2	0.75
				MAE (kg/cm²)	24.40
				RMSE (kg/cm²)	30.11
19	19	8	1	R2	0.73
				MAE (kg/cm²)	24.99
				RMSE (kg/cm²)	31.41
19	22	8	1	R2	0.66
				MAE (kg/cm²)	28.28
				RMSE (kg/cm²)	35.37

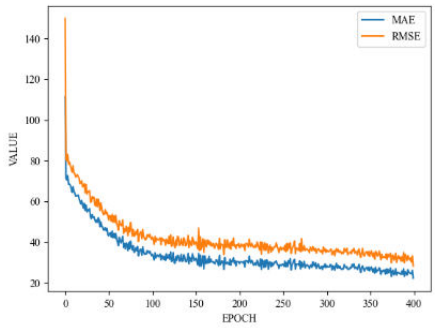


Figure 16 Rmsprop Optimizer - thirteen neurons.

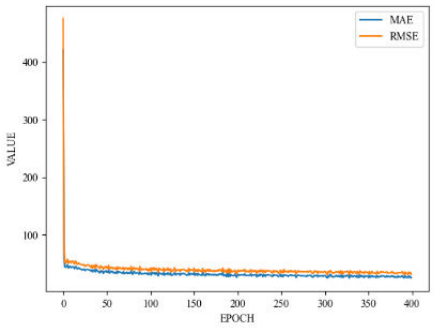


Figure 18Rmsprop Optimizer - nineteen neurons.

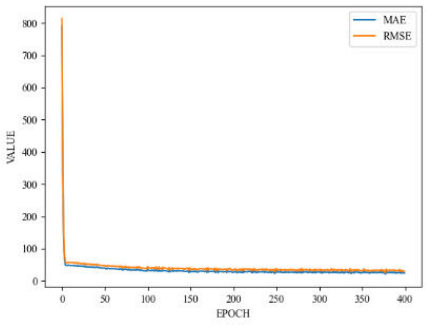


Figure 17Rmsprop Optimizer - sixteen neurons.

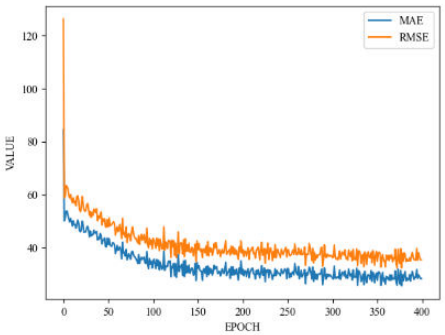


Figure 19Rmsprop Optimizer - twenty-twoneurons.

2.1.4. Nadam optimizer

Table 8 displays the Nadam optimizer's efficiency during the training phase for predicting CS, with varying configurations of neurons in the 1st layer. The optimal configuration involved nineteen neurons in that layer, resulting in an R2, MAE, and RMSE of 0.83, 20.62 kg/cm², and 28.98 kg/cm², respectively.

Table 8Performance of the Nadam Optimizer for CS Prediction

Input Layer	Hidden Layers		Output Layer		
Number of Neurons	Number of Neurons in the First Layer	Number of Neurons in the Second Layer	Number of Neurons	Criteria	Nadam
19	13	8	1	R2	0.79
				MAE (kg/cm²)	22.58
				RMSE (kg/cm²)	27.81
19	16	8	1	R2	0.72
				MAE (kg/cm²)	24.89
				RMSE (kg/cm²)	31.83
19	19	8	1	R2	0.83
				MAE (kg/cm²)	20.62
				RMSE (kg/cm²)	24.98
19	22	8	1	R2	0.72
				MAE (kg/cm²)	24.67
				RMSE (kg/cm²)	32.04

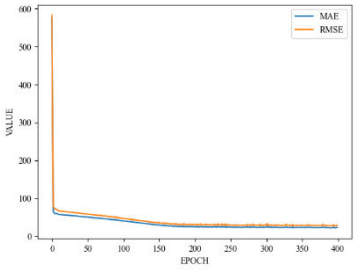


Figure 20 Nadam Optimizer - thirteen neurons

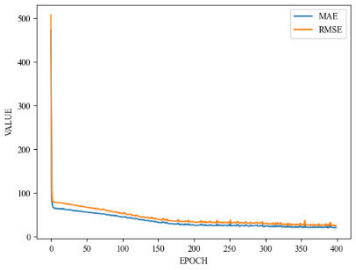


Figure 22Nadam Optimizer - nineteen neurons.

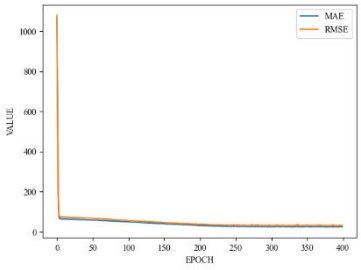


Figure 21Nadam Optimizer - sixteen neurons.

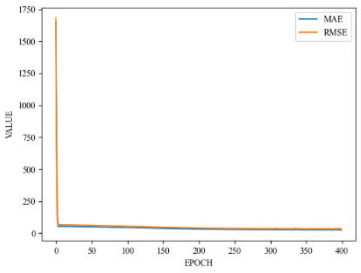


Figure 23Nadam Optimizer - twenty-two neurons.

Table 9 furnishes an overview of optimizer performance under different arrangements of neurons in the initial concealed layer.

The Nadam optimizer achieves superior performance when the number of neurons in its initial hidden layer is lower than the number of neurons in the input layer, whereas under these conditions, the Adam optimizer displays less satisfactory performance.

The Adamax optimizer exhibits superior performance in cases where the hidden layer contains a greater number of neurons than the input layer, whereas the Rmsprop optimizer yields less satisfactory outcomes in the same cases.

This analysis emphasizes the impact of the quantity of neurons in the initial concealed layer and the chosen optimizer on the neural network's overall efficiency.

The neural network configured with nineteen input neurons, nineteen and eight neurons in the hidden layers, one neuron in the output layer, and utilizing the Nadam optimizer, shows superior performance for this type of prediction during the training phase.

Tabla9Summary of Optimizer Performance in the Training Stage

Input Layer	Hidden Layers		Output Layer					
Number of Neurons	Number of Neurons in the First Layer	Number of Neurons in the Second Layer	Number of Neurons	Criteria	Adam	Adamax	RMSprop	Nadam
				R2	0.64	0.7	0.78	0.79
19	13	8	1	MAE(kg/cm²)	29.13	26.36	22.2	22.58
				RMSE(kg/cm²)	35.92	32.98	28.29	27.81
				R2	0.74	0.74	0.75	0.72
19	16	8	1	MAE(kg/cm²)	23.96	24.15	24.4	24.89
				RMSE(kg/cm²)	30.21	30.51	30.11	31.83
				R2	0.72	0.68	0.73	0.83
19	19	8	1	MAE(kg/cm²)	24.92	26.76	24.99	20.62
				RMSE(kg/cm²)	31.8	33.98	31.41	24.98
				R2	0.72	0.73	0.66	0.72
19	22	8	1	MAE(kg/cm²)	25.05	25.01	28.28	24.67
				RMSE(kg/cm²)	31.91	31.19	35.37	32.04

2.2. Prediction

For this phase, we worked with a dataset composed of 47 entries, representing 10% of the total. These data were carefully separated to ensure that the network had no prior knowledge of them. This strategy was essential to subject the models to an objective test. The absence of prior information allowed us to evaluate the predictive capacity of the neural networks against completely new data.

Table 10 provides a summary of optimizer performance in the prediction stage.

The neural network configured with nineteen input neurons, nineteen and eight neurons in the hidden layers, one neuron in the output layer, and utilizing the Nadam optimizer. This configuration has proven to be the most effective in the prediction stage, generating accurate results. The values of R^2, MAE, and RMSE confirm the superiority of this combination, reflecting Nadam's ability to optimize the neural network and provide highly reliable forecasts.

Tabla10Summary of optimizer performance in the prediction stage

Input Layer	Hidden Layers		Output Layer					
Number of Neurons	Number of Neurons in the First Layer	Number of Neurons in the Second Layer	Number of Neurons	Criteria	Adam	Adamax	RMSprop	Nadam
				R2	0.73	0.73	0.58	0.62
19	13	8	1	MAE(kg/cm²)	23.76	24.23	29.47	30.30
				RMSE(kg/cm²)	30.25	30.04	37.56	35.59
				R2	0.64	0.75	0.69	0.66
19	16	8	1	MAE(kg/cm²)	27.6	22.79	25.76	29.72
				RMSE(kg/cm²)	34.61	29.04	32.21	33.87
				R2	0.74	0.72	0.68	0.87
19	19	8	1	MAE(kg/cm²)	22.36	22.73	27.1	17.04
				RMSE(kg/cm²)	29.76	30.32	32.75	20.70
				R2	0.78	0.73	0.57	0.71
19	22	8	1	MAE(kg/cm²)	22.31	23.48	30.74	26.04
				RMSE(kg/cm²)	27.81	29.98	37.96	31.26

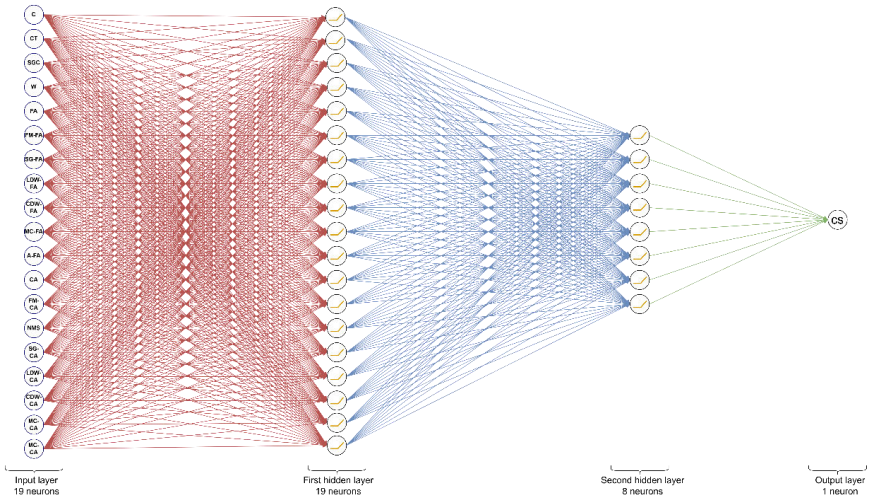


Figure 24Final neural network

2.3. Comparison with previous studies

Several authors have presented machine learning techniques to predict the CS of concrete, in most cases only presenting the mix proportions, such as water, cement, fine aggregate, coarse aggregate, and admixtures[26] [27] [28] [29] [30] [31] [32]; Not taking into account the physical and chemical properties of each one of them, for example,[32]presented six models, five of which present a large variation in the correlation factor in the training and testing phases (0.9744 - 0.7963), having in the input layer 6 variables such as coarse aggregate, fine aggregate, superplasticizer additive, silica fume, cement and water.

With the above studies we come to understand that both the external conditions and the physical and chemical properties of the concrete constituents influence the compressive strength of concrete. Because in the training phases the neural network can reach a correlation of the variables presented by the authors while in the testing or prediction phase of data never seen by the network, the prediction value decreases drastically, implying that more variables need to be entered in the input layer.

The study presented by [33]presents evaluation metrics similar to those presented in this research, where five models are evaluated to predict the CS of concrete at 28 days, including a neural network, where the architecture presented is two hidden layers with ten and five neurons in each one, obtaining as a result an R2 of 0.82 and RMSE 6.3 MPa equivalent to 64.24 kg/cm².

In the present investigation, 19 input variables were presented, including the type of cement and its specific weight, reaching an R2 of 0.83 in the training period and an R2 of 0.87 in the test phase.

2.4. Environmental impact

The models presented go beyond predicting the CS of concrete; they also seek to optimize mix designs to obtain specific strengths and, more critically, to reduce the carbon footprint incorporated in concrete. According to [34], when producing 1Tn of cement, approximately 900kg of carbon dioxide (CO_2) is released into the atmosphere, which highlights the importance of seeking more sustainable alternatives in the construction industry.

The methodology presented is intended to avoid the use of concrete specimens, which are used to obtain the CS of the concrete. These, once tested, are usually discarded in the open air, contributing to waste generation and environmental pollution.

By adopting this methodology, a more sustainable and environmentally friendly approach to the construction industry is encouraged, contributing to the reduction of CO_2 emissions associated with cement production and eliminating the generation of waste derived from traditional physical testing.

3. Conclusions

In this research, a total of sixteen artificial neural network models were developed and evaluated to predict concrete compression strength. Different datasets were used, consisting of nineteen input parameters related to aggregate characteristics, quantities, specific weight, type of cement used, and the required amount of water. These parameters were used to evaluate the presented neural networks.

The base neural network consisted of nineteen input parameters and two hidden layers. The first hidden layer had variations in the number of neurons to evaluate the optimizers and determine which one performed better for this type of data. The second hidden layer had a fixed value of eight neurons, and there was a single neuron in the output layer, representing the concrete compression strength.

The Nadam optimizer exhibited the best performance in the training stage, with a neural network configuration of nineteen input neurons, nineteen and eight neurons in the hidden layers, and a single neuron in the output layer, resulting in R^2, MAE, and RMSE values of 0.83, 20.56 kg/cm², and 24.98 kg/cm², respectively. In the prediction stage, the Nadam optimizer also demonstrated the optimal performance with a neural network configuration of nineteen input neurons, nineteen and eight neurons in the hidden layers, and a single neuron in the output layer, yielding R^2, MAE, and RMSE values of 0.87, 17.04 kg/cm², and 20.70 kg/cm², respectively.

Several authors have addressed this issue, focusing mostly on mixing ratios, but without considering the individual physical and chemical properties of the components, which has led to generalization problems and loss of accuracy in the testing phase of neural networks. In this study, 19 input variables were presented, which are detailed in Table 1, demonstrating an improvement in the accuracy of the predictions in compassion with previous research presenting fewer input variables.

A potential improvement was identified in the concrete CS prediction section as more input variables are included. For future research it is suggested to add the variables, specimen curing time, pouring temperature, relative humidity, slump, type of curing, and additional aggregate characteristics.

Declaration of competing interest

The authors declare that they have no any known financial or non-financial competing interests in any materialdiscussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

References

[1] F. Khademi, S. Mohammadmehdi Jamal, N. Deshpande and S. Londhe, "Predicting strength of recycled aggregate concrete using Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System and Multiple Linear Regression," *International Journal of Sustainable Built Environment,* pp. 355-369, 2016.

[2] F. Abanto, Tecnología del concreto, Lima: San Marcos E.I.R.L, 2009.

[3] P. J. M. Monteiro, S. A. Miller and A. Horvath, "Towards sustainable concrete," *Nature Materials,* pp. 698-699, 2017.

[4] H. A. Bengar, A. A. Shahmansouri, N. A. Z. Sabet, K. Kabirifar and V. W. Tam, "Impact of elevated temperatures on the structural performance of recycled rubber concrete: Experimental and mathematical modeling," *Construction and Building Materials,* 2020.

[5] G. Habert, S. A. Miller, V. M. John, J. L. Provis, A. Favier, A. Horvath and K. L. Scrivener, "Environmental impacts and decarbonization strategies in the cement and concrete industries," *Nature Reviews Earth & Environment,* pp. 559-573, 2020.

[6] T. L. Ávalos-Rendón, E. A. P. Chelala, C. J. M. Escobedo, I. A. Figueroa, V. H. Lara and L. M. Palacios-Romero, "Synthesis of belite cements at low temperature from silica fume and natural commercial zeolite," *Materials Science and Engineering,* pp. 79-85, 2018.

[7] A. C. 228, "Methods to Estimate Concrete," 2003.

[8] S. H. Kosmatka, B. Kerkhoff, W. C. Panarese and J. Tanesi, Diseño y control de mezclas de concreto, Illionis: Portland Cement Association, 2004.

[9] P. Visintin, T. Xie and B. Bennett, "A large-scale life-cycle assessment of recycled aggregate concrete: The influence of functional unit, emissions allocation and carbon dioxide uptake," *Journal of Cleaner Production,* pp. 1-18, 2020.

[10] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature,* pp. 436-444, 2015.

[11] Jinlong Yang, Yucheng Fan, Fan Zhu, Zhi Ni, Xili Wan, Chuang Feng and Jie Yang, "Machine learning prediction of 28-day compressive strength of CNT/ cement composites with considering size effects," *Composite Structures,* pp. 1-14, 2023.

[12] Yassine El Khessaimi, H. Youssef El , Agnés Smith, Claire Peyratout, Karim Tamine, Samir Adly and Moulay Barkatou, "Machine learning-based prediction of compressive strength for limestone calcined clay cements," *Journal of Building Engineering,* pp. 1-25, 2023.

[13] V. Toufigh and A. Jafari, "Developing a comprehensive prediction model for compressive strength of fly ash-based geopolymer concrete (FAGC)," *Construction and Building Materials,* pp. 1-17, 2023.

[14] Divyah Nagarajan, Thenmozhi Rajagopal and Neelamegam Meyappan, "A comparative study on prediction models for strength properties of LWA concrete using artificial neural network," *Journal of Construction,* pp. 103-111, 2019.

[15] A. 301S-10, Specifications for Structural Concrete, Farmington Hills: American Concrete Institute, 2010.

[16] E. M. Golafshani, A. Behnood and M. Arashpour, "Predicting the compressive strength of normal and High-Performance Concretes using ANN and ANFIS hybridized with Grey Wolf Optimizer," *Construction and Building Materials,* vol. 232, pp. 1-14, 2020.

[17] P. Chopra, R. K. Sharma and M. Kumar, "Prediction of Compressive Strength of Concrete Using Artificial Neural Network and Genetic Programming," *Advances in Materials Science and Engineering,* vol. 2016, p. 10, 2016.

[18] M. T. Hagan, H. B. Demuth and M. Beale, "Neural network design," *PWS Publishing Co,* 2014.

[19] T. Hastie, R. Tibshirani and J. Friedman, The elements of statistical learning: data mining, inference, and prediction,

Springer, 2010.

[20] T. Hamid Reza, O. Omid Lofti, S. Masoud Falahtabar and K. Saman Soleimani, "Prediction of combined effects of fibers and nano- silica on the mechanical properties of self- compacting concrete using artificial neural network," *Latin American,* pp. 1906-1923, 2014.

[21] I. Goodfellow, Y. Bengio and A. Courville, "Deep Learning," *MIT,* 2016.

[22] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz and Lu, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *arXiv,* 2016.

[23] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *ArXiv,* 2014.

[24] T. Tieleman and H. G., "Lecture 6.5 - RMSprop: Divide the gradient by a running average of its recent magnitude," 2012.

[25] T. Dozat, "Incorporating Nesterov Momentum into Adam," 2016.

[26] A. Öztaş, M. Pala, E. Özbay, E. Kanca, N. Çağlar and M. A. Bhatti, "Predicting the compressive strength and slump of high strength concrete using neural network," *Construction and Building Materials,* vol. 20, pp. 769-775, 2006.

[27] H.-G. Ni and J.-Z. Wang, "Prediction of compressive strength of concrete by neural networks," *Cement and Concrete Research,* vol. 30, pp. 1245-1250, 2000.

[28] A. U., "Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network," *Expert Systems with Applications,* vol. 38, pp. 9609-9618, 2011.

[29] O. A. Kabiru, T. O. Owolabi, T. Ssennoga and S. O. Olatunji, "Performance Comparison of SVM and ANN in Predicting Compressive Strength of Concrete," *IOSR Journal of Computer Engineering,* vol. 16, pp. 88-94, 2014.

[30] J.-S. Chou, C.-K. Chiu, M. Farfoura and I. Al-Taharwa, "Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques," *Journal of Computing in Civil Engineering,* vol. 25, pp. 242-253, 2011.

[31] Y. I.-C., "Modeling of strength of high-performance concrete using artificial neural networks," *Cement and Concrete Research,* vol. 28, pp. 1797-1808, 1998.

[32] M. F. Zarandi, I. Tu¨rksen, J. Sobhani and A. Ramezanianpour, "Fuzzy polynomial neural networks for approximation of the compressive strength of concrete," *Applied Soft Computing,* pp. 488-498, 2008.

[33] B. A. Young, A. Hall, L. Pilon, P. Gupta and G. Sant, "Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods," *Cement and Concrete Researc,* vol. 115, pp. 379-388, 2019.

[34] N. Mahasenan, S. Smith and K. Humphreys, "The cement industry and global climate change: current and potential future cement industry CO2 emissions," in *Greenhouse Gas Control Technologies - 6th International Conference*, Kyoto, Japan, 2002.