Erudite Journal of Engineering Technology and Management Sciences
Vol. 5, No. 4, December 2025, pp. 06-09
Journal homepage: http://www.ejetms.com

Machine Learning Techniques to Predict the Compressive Strength of Metakaolin Blended Concrete

Uppara Raghu Babu [1*], J.Mukesh [2], B.Harshavardhan[2], K.Sudheer Kumar[2], S.Sohaib Ali[2]

[1]Professor, Civil Engineering Department, Srinivasa Ramanujan Institute of Technology, Ananthapuramu, Andhra Pradesh, India.
[2] Under Graduate Student, Civil Engineering Department, Srinivasa Ramanujan Institute of Technology, Ananthapuramu, Andhra Pradesh, India.

Received: 21 July 2025
Accepted: 19 September 2025

Keywords:
Supplementary Cementitious Materials, Compressive Strength, Silica Fume, Prediction Models, Artificial Neural Networks, Cement Replacement, Classification and Regression Tree, Gene Expression Programming.

ABSTRACT

Testing the mechanical properties of concrete is costly and time-consuming, often requiring up to 28 days. This study proposes machine learning models to predict the compressive strength of metakaolin- blended concrete, reducing the need for extensive testing. Techniques such as CART, ANN, and GEP were applied using seven predictors: cement, metakaolin, fine aggregate, coarse aggregate, water, superplasticizer, and curing age. Model performance was evaluated with R^2, RMSE, MSE, and MAPE, and validated using 10-fold cross-validation. Results show that machine learning effectively predicts compressive strength at different curing periods, offering a reliable alternative to traditional testing.

1. INTRODUCTION

The compressive strength of concrete is its most crucial mechanical property. Determining this strength is a lengthy and expensive process, typically taking 28 days to test in a lab. The final strength is influenced by various factors, including the type of cement, water-to-cement ratio, aggregate size, and chemical admixtures.

The Focus of This Study

The current research aims to address the time-consuming and labor-intensive nature of testing concrete compressive strength. The study uses machine learning models—specifically, Artificial Neural Networks (ANN), Classification and Regression Trees (CART), and Gene Expression Programming (GEP)—to predict the compressive strength of concrete containing Metakaolin. The study will also incorporate factors like the size and shape of the test specimens to improve the accuracy of these predictions.

Supplementary Cementitious Materials (SCMs)

Concrete is one of the most widely used construction materials globally, but its production, particularly the manufacturing of cement clinker, is a major source of carbon dioxide (CO_2) emissions. To reduce this environmental impact, the construction industry is increasingly using Supplementary Cementitious Materials (SCMs).

One such SCM is Metakaolin (MK), a highly reactive pozzolanic material obtained by heating pure kaolin clay. MK can partially replace cement, leading to several benefits: Improved Strength: MK enhances the compressive strength and other mechanical properties of concrete. Reduced Permeability: It helps reduce the porosity of concrete, increasing its durability. Lower Emissions: Using MK reduces the need for cement, which in turn lowers CO_2 emissions.

2. LITERATURE

C.S. Poon et al. (2005) investigated the compressive strength, chloride diffusivity, and pore structure of high-performance concrete incorporating metakaolin (MK) and silica fume (SF). Tests were conducted at water-to-binder ratios of 0.3 and 0.5. Results showed that MK and SF improved compressive strength, reduced chloride penetrability, and decreased porosity. The study highlighted that interfacial porosity strongly influences durability, offering key insights into the microstructure– performance relationship of MK- and SF-based concretes.

M.Saridemir et al (2008), observed that ANN models can accurately predict the compressive strength of concrete containing metakaolin and silica fume at different ages without conducting any experiments. Two different multilayer artificial neural network architectures were developed namely ANN-I and ANN-II. In ANN-I model, one

hidden layer was selected. The discussion highlights the benefits of using artificial neural networks in civil engineering applications to solve complicated problems. The study suggests that ANN has been widely used in modelling various human activities in civil engineering applications.
R.M. Ferreira et al (2015), identified that adding metakaolin to concrete can improve its strength, durability properties, and chloride penetration resistance. His study concludes that the durability performance of concrete with metakaolin replacements improves not only due to the increase in resistance to chloride ingress but also due to the rate at which this decrease takes place over time.

3. METHODOLOGY

Data Collection

Experimental data on metakaolin concrete from previous studies is used to develop prediction models at different ages. This chapter outlines the collected data and adopted methodology.

Table 1: Metakaolin data

Name	Cement Type	Pozzolan	Input	Output	Data points	Ref
C.S. Poon et al	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	24P	[4]
M. Saridemir	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	64P	[3]
R.M. Ferreira	OPC	MK	AGE, C, MK, FA, CA, WC, SP	CS	55P	[7]

Machine Learning Techniques for Prediction Models

To develop compressive strength prediction models, machine learning techniques such as Artificial Neural Networks (ANN), Decision Tree, Classification and Regression Tree (CART), and Gene Expression Programming (GEP) are adopted. The models are developed using selected dependent and independent variables.

Artificial Neural Network (ANN)

Artificial Neural Networks are computational models inspired by the human brain. They consist of interconnected neurons arranged in input, hidden, and output layers.
Forward propagation: Input data passes through layers, multiplied by weights and biases, then processed by an activation function (e.g., log-sigmoid, tan-sigmoid) to generate output probabilities.

Back propagation: Errors between predicted and actual outputs are sent backward to adjust weights, minimizing error through repeated training cycles.

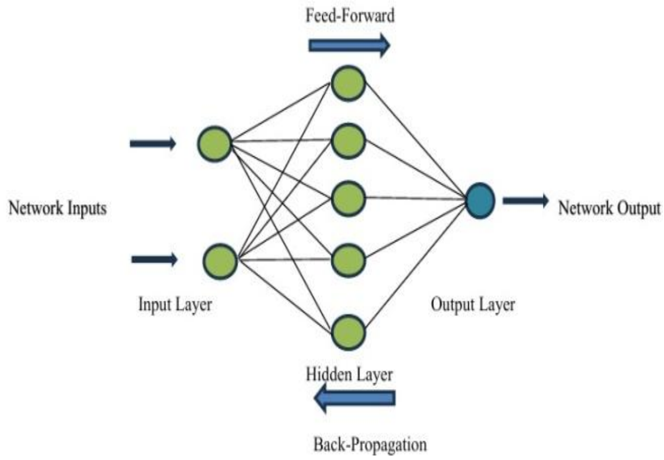

Figure 1: Neural Network Architecture

In this research, the connection between input value and output value can be expressed as

$$Y_i^{a+1} = F\left(\sum_b W_{ib}^a X_b^a\right)$$

Classification and Regression Tree

Classification and Regression Tree (CART) analysis is a tree-building technique that partitions data into regions based on predictor variables, enabling classification and regression tasks. The process involves four steps: tree building through recursive splitting, stopping the growth of the maximal tree, pruning to remove overfitting, and selecting the optimal tree. Impurity measures such as the Gini index and least-squared deviation guide the splitting process, while terminal and non-terminal nodes define the tree structure. To evaluate CART performance, statistical indices such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) are employed. These metrics provide insights into prediction accuracy and error distribution, making CART a robust and interpretable model for data analysis.

Gene Expression Programming (GEP)

Gene Expression Programming (GEP), introduced by Ferreira, is an evolutionary algorithm that automatically generates computer programs or models. Unlike Genetic Algorithms (GAs) and Genetic Programming (GP), GEP represents solutions as fixed-length linear chromosomes, which are later expressed as expression trees (ETs). These trees adapt by changing their size and structure, similar to biological evolution. GEP consists of five main components: 1. Function set, 2. Terminal set, 3. Fitness function, 4. Control parameters, 5. Termination condition The algorithm applies genetic operators such as crossover, mutation, and rotation to evolve better solutions. The chromosomes are encoded as strings with a head (containing functions and terminals) and a tail (only terminals). They are later converted into Karva expressions (K-expressions) and then into expression trees (ETs) for evaluation. A typical GEP process involves: Creating an initial population, Translating chromosomes into ETs, Executing ETs and evaluating fitness, Selecting the best individuals

4. EXPERIMENTAL RESULTS

Artificial Neural Network For this development of Artificial Neural Network, we have taken 742 datapoints. In that 742 datapoints we have taken 520 (70%) datapoints are used for training purpose and 222 (30%) datapoints are used for testing purpose. In this data, the inputs taken as age, cement, metakaolin, fine aggregate, course aggregate, water content, super plasticizers, SiO2, Al2O3, cao. Output is the Compressive strength of Metakaolin concrete. The formula for calculating R-squared is:

R-squared = 1-(SSres/SStot)

Classification and Regression Tree
From the available data a decision tree is formed in classification and regression tree analysis. A decision tree is a tree-like model of decisions and their possible consequences.

Table 2: Observations from the CART analysis

Statistics	Training	Test
R-squared	0.93	0.83
Root mean squared error (RMSE)	5.5264	8.6034
Mean squared error (MSE)	30.5412	74.0186
Mean absolute deviation (MAD)	4.0651	5.7907
Mean absolute percent error (MAPE)	0.0848	0.1217

Gene Expression Programming
In the GEP model, we used total of 405 datapoints. In this 284 datapoints are used for training the model and remaining 121 datapoints are used for testing the model.

Table3: Observations from GEP analysis

Statistics	Training	Test
R - Squared	0.812	0.82
Root mean squared error (RMSE)	174.74	170.00
Mean squared error (MSE)	30535.87	28901.20
Mean absolute error (MAE)	139.42	139.43
Relative absolute error (RAE)	10.07	10.89

5. CONCLUSION

Compressive strength is the most important property of concrete, but its laboratory determination takes 28 days. To overcome this, prediction models are needed. In this study, the compressive strength of Metakaolin blended concrete was predicted using ANN, CART, and GEP models with input parameters such as cement, aggregates, water, fly ash, superplasticizer, and age. Among the models, GEP showed the highest accuracy with an R^2 value of 0.82, indicating its effectiveness in predicting compressive strength.

Limitations
• The present work has been carried out by considering specific input variables and the data collected from existing literature.
• The models will be valid for the concrete age of 7 to 90 days only and for metakaolin admixture.

REFERENCES

1. S. Safarzadegan Gilan, H. Bahrami Jovein, and A. A. Ramezanianpour, "Hybrid support vector regression - Particle swarm optimization for prediction of compressive strength and RCPT of concretes containing metakaolin," Constr. Build. Mater.,vol. 34, pp. 321-329, 2012, doi:10.101016/j.conbuildmat.2012.02.038.
2. O. O. Akin, A. Ocholi, O. S. Abejide, and J. A. Obari, "Prediction of the Compressive Strengt of Concrete Admixed with Metakaolin Using Gene Expression Programming," Adv. Civ. Eng., vol. 2020,2020, doi: 10.1155/2020/8883412.
3. M. Saridemir, "Prediction of compressive strength of concretes containing metakaolin and silica fume by artificial neural networks," Adv. Eng. Softw.. vol. 40, no. 5, pp. 350-355, 2009, doi: 10.1016/j.advengsoft.2008.05.002.
4. C. S. Poon, S. C. Kou, and L. Lam, "Compressive strength, chloride diffusivity and pore structure of high performance metakaolin and silica fume concrete," Constr. Build. Mater., vol. 20, no. 10, pp. 858-865, 2006, doi: 10.1016/j.conbuildmat.2005.07.001.
5. P. Muthupriya, K. Subramanian, and B. G. Vishnuram, "Prediction of Compressive Strength and Durability of High-Performance Concrete By Artificial Neural Networks," Int. J. Optim. Civ. Eng., vol. 1, pp. 189-209, 2011.
6. R. M. Ferreira, J. P. Castro-Gomes, P. Costa, and R. Malheiro, "Effect of metakaolin on the chloride ingress properties of concrete," KSCE J. Civ. Eng., vol. 20, no. 4, pp. 1375-1384, 2016, doi: 10.1007/s12205-015-0131-8.
7. S. Wild, J. M. Khatib, and A. Jones, "Relative strength, pozzolanic activity andcement hydration in superplasticised metakaolin concrete," Cem. Coner. Res., vol. 26, no. 10, pp. 1537 1544, 1996, doi: 10.1016/0008-8846(96)00148-2.
8. K. B. Park, R. S. Lin, Y. Han, and X. Y. Wang, "Model-based methods to produce greener metakaolin composite concrete," Appl. Sci., vol. 11, no. 22, 2021, doi: 10.3390/app112210704.
9. A. A. Ramezanianpour and H. Bahrami Jovein, "Influence of metakaolin as supplementary cementing material on strength and durability of concretes," Constr. Build. Mater., vol. 30, pp. 470-479, 2012, doi: 10.1016/j.conbuildmat.2011.12.050.
10. J.M. Justice, L.H. Kennison, B.J. Mohr, S.L. Beckwith, L.E. McCormick, B. Wiggins, Z.Z. Zhang, and K.E. Kurtis, "Comparison of Two Metakaolins and a Silica Fume Used as Supplementary Cementitious Materials"
11. Varinder Kumar Bansal, Maneek Kumar, Prem Pal Bansal, Ajay Batish, "Predicting compressive strength of concrete containing ternary combination of industrial byproducts as partial replacement of cement and fine aggregate of

cement and fine aggregates using ann and anfis" doi.org/10.21172/1.113.07.
12. Vikas Srivastava, Rakesh Kumar. Agarwal V.C. Mena P.K-Effect of silica Fume and Metakaolin combination on concrete" doi:10.6088/ijcser.00202030017.
13. Jian-Tong Ding and Zongjin Li, "Effects of Metakaolin and Silica Fume on Properties of Concrete". [14] Hoang Nguyen, Thanh Vu, Thuc P. Vo, Huu-Tai Thai, "Efficient machine learning models for prediction of concrete strengths" https://doi.org/10.1016/j.conbuildmat.2020.120950