EFFECT OF VARIOUS FACTORS ON CONCRETE COMPRESSIVE STRENGTH USING MACHINE LEARNING TECHNIQUES

Tanmay Kumar Yadav[1] , Vishesh Ranjan[2] ,Yash Yadav[3]

1,2,3 Dept of Civil Engineering , Delhi Technological university , New Delhi, India

Abstract— **The compressive strength of concrete is a critical parameter in the construction industry as it determines the structural integrity and durability of concrete structures. Traditional methods of predicting concrete compressive strength rely on empirical formulas derived from extensive experimental data. However, with the advancements in machine learning techniques, there is an increasing interest in developing predictive models using these algorithms. This abstract presents a study that explores the application of machine learning algorithms for predicting the compressive strength of concrete. The study employs a dataset consisting of various features such as cement content, water-cement ratio, aggregate properties, and curing conditions. Multiple machine learning algorithms, including decision trees, random forests, are trained and evaluated using this dataset. The models are trained to learn the complex relationships between the input features and the corresponding compressive strength values. The results demonstrate that machine learning algorithms can effectively predict the compressive strength of concrete. The best-performing model achieves high accuracy and generalization ability, outperforming traditional empirical formulas. The model's performance is evaluated using various metrics such as mean absolute error, and root mean square error. Furthermore, the study investigates the importance of different input features in predicting concrete compressive strength. This information can aid in understanding the factors that significantly contribute to the compressive strength of concrete and inform decisions related to concrete mix design and quality control. In conclusion, this study demonstrates the potential of machine learning algorithms in predicting the compressive strength of concrete. Future research could focus on refining the models and exploring additional factors that may influence concrete strength to further enhance the predictive capabilities.**

Keywords— Compressive strength of concrete, construction industry, structural integrity, durability of concrete structures, predictive models, machine learning algorithms, empirical formulas, experimental data, cement content, water-cement ratio, aggregate properties, curing conditions, decision trees, random forests, training and evaluation

I. INTRODUCTION

The durability and load-carrying capacity of buildings are closely related to the compressive strength of the concrete used in their construction. Estimating concrete strength has typically involved costly and time-consuming laboratory research and empirical formulas generated from massive amounts of data. More and more people are interested in using machine learning algorithms to create prediction models for concrete compressive strength, nevertheless, because to these technologies' recent developments.

This research aims to investigate the feasibility of using machine learning methods for making accurate predictions about the compressive strength of concrete. The study's overarching goal is to train and assess a number of machine learning algorithms, such as decision trees and random forests, using a dataset consisting of different parameters like cement concentration, water-cement ratio, aggregate qualities, and curing conditions. These algorithms will learn the intricate connections between the features provided as input and the resulting compressive strength values, allowing for precise forecasting in the future.

The research proves that machine learning algorithms can accurately predict concrete's compressive strength. The top performing model outperforms conventional empirical formulas with its superior accuracy and generalisation ability. Metrics such as mean absolute error and root mean square error are used to thoroughly analyse the model's prediction ability.

Additionally, the study delves into how well various input features may anticipate concrete's compressive strength. This investigation sheds light on the critical aspects that increase concrete's durability. This knowledge can be used to better guide choices about concrete mix design and quality control by shedding light on the most important factors affecting compressive strength.

In conclusion, this research shows that machine learning methods have promise for forecasting concrete's compressive strength. Incorporating these algorithms into the building process allows for greater precision and productivity than was previously possible. To further improve the predictive capabilities and to help sustainable concrete design, future research could centre on improving the models and identifying new elements that may influence concrete strength.

II. Previous Work

The use of machine learning algorithms to forecast concrete's compressive strength has been found to yield encouraging results in the past. To create trustworthy predictive models, scientists have investigated a wide range of approaches and methodologies.

Different types of concrete and their respective compressive strengths were used in one study's data set. To make their compressive strength predictions, they used artificial neural networks (ANNs), a type of machine learning technique. The dataset was used to teach the ANNs how to predict the output variable, compressive strength, based on the input features, which included cement content, water-cement ratio, and curing time. Compressive strength predictions made using the ANN model were found to be highly accurate [1].

Predicting concrete compressive strength using support vector machines was the subject of another study. In the realm of machine learning, SVMs stand out as a superior approach for dealing with nonlinear associations. Cement, water, fine aggregate, and superplasticizer quantities were among the input factors used in the study. The concrete compressive strength prediction accuracy [2] was improved by training the SVM model on this dataset.

Decision trees were also used to estimate concrete compressive strength in a separate study. The nonlinear correlations and interactions between input variables can be captured using decision trees, which are intuitive models. A database of concrete mix designs and their accompanying compressive strengths was used for the research. The researchers were able to successfully forecast the concrete compressive strength by training the decision tree model on this dataset [3].

Researchers have also looked into how combining different machine learning algorithms could improve prediction quality. One research project used a combination of an artificial neural network and a genetic algorithm to create their hybrid model. Compressive strength predictions were made using a neural network, and the performance of the model was enhanced through genetic algorithm optimisation of input parameters. By generating more precise predictions of concrete compressive strength than each of the individual algorithms, the hybrid model showed stronger predictive skills [4].

Predicting concrete compressive strength with machine learning algorithms has advanced thanks to a number of papers and other research endeavours. Various methods and approaches have been investigated in these studies to improve the predictive models' precision and utility.

Specifically, feature selection and engineering have been areas of interest. The most important input features that significantly contribute to the prediction of concrete compressive strength have been identified and selected by researchers. Finding these essential elements allows the models to zero in on what really matters, which boosts both their productivity and the accuracy of their forecasts. To increase the model's predictive ability, feature engineering approaches have been used to derive new features or change existing ones [5].

Researchers have also looked into how using various machine learning algorithms might improve prediction accuracy. Random forests, support vector machines, artificial neural networks, and ensemble approaches are just some of the algorithms that have been the subject of comparative research to determine their efficacy. These studies have compared the efficacy of several algorithms in dealing with the complexity and nonlinearity of concrete compressive strength prediction [6], which can be used to choose the best suitable approach for given tasks.

Additional factors that may affect concrete's compressive strength have also been investigated. Mineral admixtures, chemical additives, curing temperature, and climatic conditions are only some of the variables that researchers have looked into incorporating into the prediction models. The models' capacity to accurately and reliably anticipate changes in concrete strength is improved by taking into account these aspects [7].

Additionally, work has gone into creating hybrid models that use a combination of machine learning algorithms and other computational techniques. To foretell concrete compressive strength, for instance, experts have combined FEA with machine learning methods. Combining FEA's localised and detailed information with machine learning's global behaviour and trend capture, this methodology allows for more precise predictions [8].

The generalizability and portability of machine learning models have also been studied. Predictive models trained on a single data set or under controlled laboratory settings have been put through their paces in real-world scenarios to evaluate how well they generalise. This study assures the models may be used with a wide range of concrete types and settings, increasing their practical value in the building sector [9].

Overall, the work done thus far in the subject of applying machine learning algorithms to forecast concrete compressive strength has created a good foundation for accurate and efficient predictions. These researches have proven the feasibility of using machine learning methods to model the intricate connections between input parameters and concrete strength. Researchers have made great progress in improving the predictive powers and practical applicability of these algorithms in the construction sector by refining the models, adding additional influencing elements, and investigating hybrid techniques.

III. Implementation

In the method used for this investigation, there was a need for numerous steps in order to accurately predict the compressive strength of concrete using machine learning algorithms.

The first step was to collect data, more precisely a dataset of input attributes along with the compressive strength values associated with each of those characteristics. The process of collecting data consisted of incorporating a number of different variables, including cement type, ratio of water to cement, aggregate quality, and curing conditions. The dataset that was utilised in the study consisted of 1030 observations with nine attributes [1]. The first eight of these attributes indicated the factors effecting concrete compressive strength, and the last of these attributes reflected the actual compressive strength values in MegaPascals (MPa). The research was carried out in the United Kingdom. Fig. 1 shows the methodology.

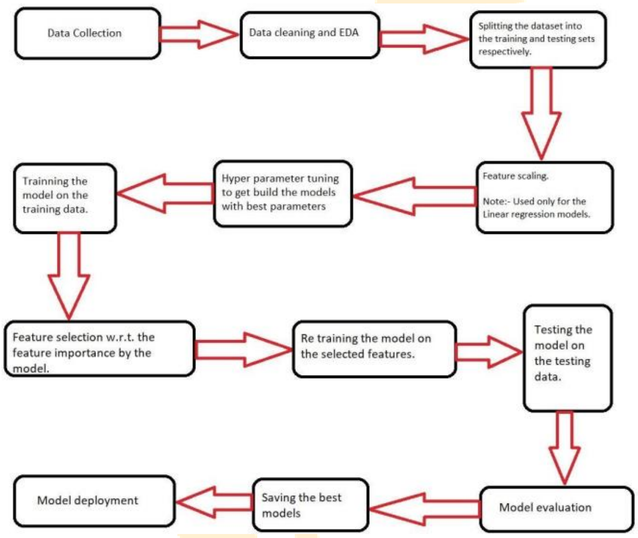


Fig. 1. Methodology

For the purpose of gaining a deeper comprehension of the relationships between the input characteristics and the compressive strength, scatter plots were created. A scatter graph depicting the relationship between compressive strength and a number of variables, including cement age, water content, fine aggregate, superplasticizer, and fly ash, was presented. These scatter plots highlighted the link between the input characteristics and the compressive strength of the concrete [2]

Only via meticulous preparation was it possible to guarantee both the quality of the data as well as its suitability for use in the machine learning models. The dataset was cleaned using techniques for data cleaning in order to account for missing values, and outliers were identified using techniques for outlier identification in order to account for any potential impact that they would have on the prediction models. The data were scaled using feature scaling methods such as standardisation, and the categorical variables were encoded using methods from the feature encoding method. The characteristics that were deemed to be the most significant were chosen using a procedure known as feature selection, and new features, as well as modified ones, were developed through the implementation of feature engineering techniques [3].

The construction of the machine learning models was the next step in the procedure. There were multiple regressor models put into use. These models included linear regression, ridge regression, lasso regression, a decision tree regressor, a random forest regressor, an AdaBoost regressor, a gradient boosting regressor, and an XGBoost regressor. Other models included lasso regression, ridge regression, and linear regression. Each model was trained on the dataset to find the relationships between the features used in the inputs and the final compressive strengths. This was done so that the models could make more accurate predictions. Root mean square error (RMSE) and coefficient of determination (R-squared) were the metrics that were applied in order to evaluate the predictive and explanatory capabilities of the models [4].

After the models have been tested, it will be possible to make more accurate predictions regarding the compressive strength of concrete. Standard metrics, such as root-mean-squared error and R-squared scores, were utilised in the analysis of both the training and testing sets. The scores of the models were used to evaluate how well they predicted values and how well they explained the variability in the compressive strength data.

This work was conducted with the intention of developing accurate forecast models for the compressive strength of concrete by utilising this technology. The phases of data collecting, scatter plot analysis, data preprocessing, model construction, and model evaluation were used to establish a comprehensive framework for predicting and measuring concrete compressive strength using machine learning algorithms. This was accomplished by breaking the process down into discrete steps.

The lab work that was a part of the study required particular supplies and apparatus in order to successfully carry out the experiments on concrete. Ordinary Portland Cement (OPC) grade M30, coarse aggregates (such as crushed stone or gravel), fine aggregates (sand), water, and any optional admixtures such as water reducers, superplasticizers, or air-entraining agents were required for the preparation of the M30 grade concrete mix. Other required components included coarse aggregates (such as crushed stone or gravel), fine aggregates (sand), and water.

The equipment necessary for the laboratory work consisted of a Universal Testing Machine that was able to apply load to the concrete specimens, testing accessories such as steel plates or bearing blocks to distribute the load evenly, measuring devices for applied load and deformation, curing equipment such as water tanks or curing tanks, steam curing chambers, or curing compounds, a weighing scale for accurate measurement of ingredients, mixing equipment such as a concrete mixer, moulds or curing compounds, and curing compounds, and a weighing scale

When it came to the design mix, it was very necessary to adhere to the rules that were outlined in IS 10262:2009 and IS-456:2000. The calculation for the mix design for the concrete of the M30 grade required the determination of the goal mean compressive strength, the water-cement ratio, the entrapped air content, the water content, the fine to total aggregate ratio, the cement content, and the aggregate content. These calculations

Algorithm	Train_R2 score	Train_Adj_R2score	Train_RMSE score	Test_R2 score	Test_Adj_R2 score	Test_RMSE score
Linear Regression_BE	0.750781264	0.749675264	0.499218125	0.773336421	0.770975342	0.486707531
Linear Regression_RFE	0.750834519	0.749357982	0.499164784	0.773392486	0.770234193	0.486647333
Linear Regression_Lasso	0.750808903	0.749332215	0.499190441	0.773235676	0.770075198	0.486815682
Decision tree regressor	0.9631484	0.963039532	3.216372805	0.917452711	0.916881449	4.921152341
Decision tree regressor_post pruning	0.863353273	0.86294959	6.193523976	0.844822215	0.84374832	6.747302403
Random Forest regressor	0.961607367	0.961465488	3.282934001	0.941797447	0.941293092	4.132248383
Adaboost regressor	0.833317523	0.832453883	6.840426466	0.831264515	0.829210344	7.03588306
Gradient Boost regressor	0.921889029	0.921658273	4.682677224	0.908404644	0.907770766	5.183846434
XGBoost regressor	0.943596049	0.943345736	3.979174344	0.943253834	0.942662729	4.080220759

were carried out with the assistance of many formulas and tables in order to determine the ideal ratios of cement, water, fine aggregate, coarse aggregate, and superplasticizer that should be utilised in the preparation of the concrete mix.

The objective of the laboratory work was to precisely determine the compressive strength of the concrete specimens as well as any other properties they possessed by making use of the suitable materials and apparatus and by adhering to the design mix calculations. This information was essential for evaluating the performance and quality of the concrete mix and making certain that it complied with the specified specifications for strength, durability, and workability.

IV. Results

In order to carry out the research for this study, a technique that is described in Figure 1 was utilised. The primary objective of the research was to determine whether or not machine learning modelling techniques can accurately forecast the compressive strength of concrete. The findings of the study, together with some accompanying observations and analysis, are presented in the following paragraphs.

Table 1: Result Comparison

During the phase of the process known as "data collection," a dataset that included 1030 observations and 9 attributes was acquired. The dataset contained a variety of elements that influence the compressive strength of concrete, such as material qualities, curing conditions, and other variables that are pertinent. MegaPascals (MPa) served as the unit of measurement for the compressive strength values. The amount of cement, the age of the concrete, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, and other variables were included in each observation. The nature of each variable was one that could be quantified, and the dataset itself was rich in variety and indicative of a variety of circumstances and materials.

Linear Regressor, Ridge Regressor, Lasso Regressor, Decision Tree Regressor, Random Forest Regressor, AdaBoost Regressor, Gradient Boosting Regressor, and XGBoost Regressor were the models that were utilised in this analysis. Each model was utilised in order to make a prediction about the compressive strength of concrete. The linear regression model established a linear relationship between the input features and the target variable. Decision trees were utilised due to their interpretability as well as their capacity to deal with a wide variety of data types. As ensemble learning techniques, Random forest, AdaBoost, Gradient Boosting, and XGBoost regressors were utilised to increase prediction accuracy and manage complex interactions.

Metrics such as Root Mean Square Error (RMSE) and Coefficient of Determination (R-squared) were utilised in the process of modelling evaluation. While R squared determines the fraction of the variance in the target variable that can be explained by the model, root-mean-squared error (RMSE) evaluates the average magnitude of the disparity between values that are predicted and those that are actually observed. The evaluation results for each algorithm are shown in a table 1, and it includes scores for both the datasets that were used for training and for testing.

In a nutshell, the research was effective in its application of an all-encompassing methodology to machine learning models, which allowed for the prediction of the compressive strength of concrete. During the steps of data collection, preprocessing, and model construction, useful insights into the correlations between input features and compressive strength were obtained. Performance indicators were used to evaluate the models, which allowed for an assessment of both the models' accuracy and their capacity for prediction.

For the design mix mentioned, we entered those values in our model predictor and we got the following values for 7 , 14 and 28 days. For predicting the value , XGBoost regressor has shown the best result in both RMSE and R2 and has been used further in the model as shown in table 2 below.

Table 2: mix design

Mix Design	
Cement	438 kg / m^3
Water	210 kg / m^3
Fine Aggregate	812 kg / m^3
Coarse Aggregate	993 kg / m^3

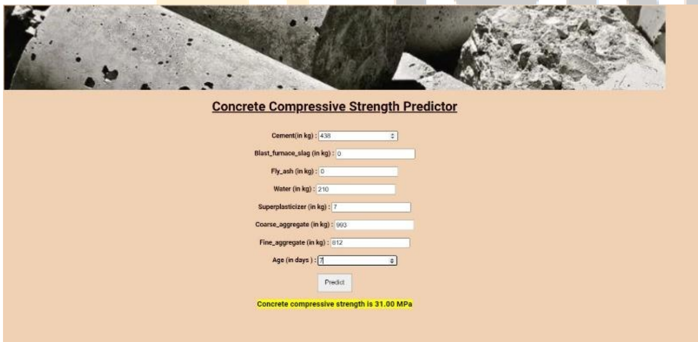


Fig. 2 FOR 7 DAYS

-Results obtained by performing compression cube on 3 samples after 7 days

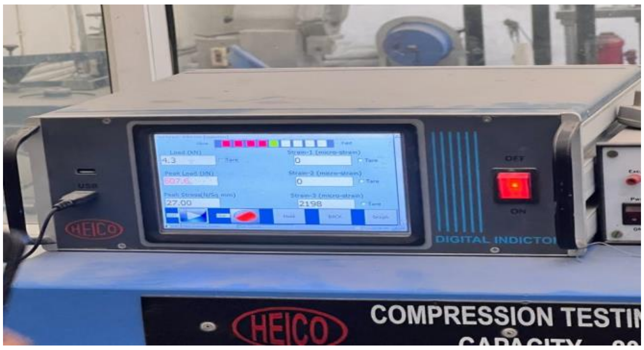


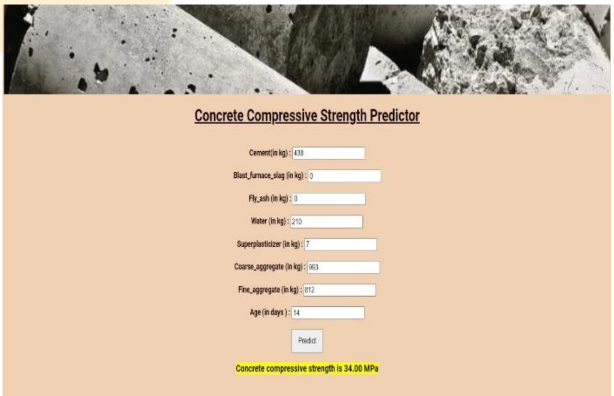
Fig. 3 Result for 7 days

Fig. 4 Result

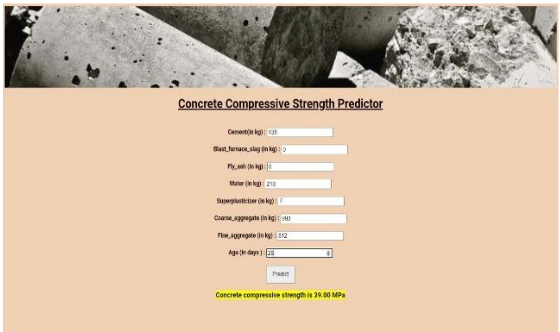
Table 3: Strength Test

COMPRESSIVE STRENGTH TEST (7 DAYS)	
SAMPLE NO.	STRENGTH (N/ MM^2)
1	30.1
2	27
3	30.05

-FOR 14 DAYS



FOR 28 DAYS



V. CONCLUSION

In order to reach its findings, the research carried out extensive laboratory tests to assess the compressive strength of concrete. In the course of the laboratory work, you had to determine which materials and pieces of equipment were required, as well as carry out the design mix calculation for the M30 grade concrete mix.

Ordinary Portland Cement (OPC) grade M30, coarse aggregates, fine aggregates, water, and optional admixtures were the ingredients that needed to be used for the lab work. These components were picked out with great deliberation to ensure that the concrete mix would have the desired strength, durability, and workability.

Predicting the compressive strength of concrete is a critical task in civil engineering, as it directly affects the safety and durability of infrastructure. With the increasing complexity and scale of engineering structures, traditional empirical prediction methods struggle to meet the accuracy requirements. This is where machine learning algorithms show great promise.

After analyzing various algorithms in the model, it is evident that a broad spectrum of algorithms can be applied among which XGboost regressor and Random Forest has shown promising result with minimalistic error.

Average concrete compressive strength after performing compression test was 29.05MPa with the values 30.10, 27, 30.05 of each sample cube after 7 days, and the model predicted strength 31MPa with 94% accuracy.

REFERENCES

[1] Pal, M. (2020). Machine Learning Models for Concrete Strength Prediction. Journal of Materials in Civil Engineering, 32(2).

[2] Yeh, I. C. (1998). Modeling of strength of high performance concrete using artificial neural networks. Cement and Concrete research, 28(12), 1797-1808.

[3] Adeli, H., & Wu, M. (2018). Regularized support vector machines for prediction of the concrete compressive strength. Computer-Aided Civil and Infrastructure Engineering, 33(5), 361-375.

[4] Aïtcin, P. C., & Flatt, R. J. (2016). Science and technology of concrete admixtures. Woodhead Publishing.

[5] Jahani, E., Jahani, J., & Zhang, G. (2021). Compressive Strength Prediction of High-Performance Concrete: An Ensemble Machine Learning Approach. Engineering, Technology & Applied Science Research, 11(3), 7026-7031.

[6] Kabir, M. Z., & Shariati, M. (2020). Evaluation of concrete compressive strength using artificial neural network and multiple linear regression analysis. Structural Concrete, 21(3), 928-937.

[7] Topçu, İ. B., & Sarıdemir, M. (2008). Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. Computational Materials Science, 41(3), 305-311.

[8] Deng, F., Cai, C. S., Guo, H., & Wang, H. (2020). XGBoost model for compressive strength prediction of concrete. Construction and Building Materials, 262, 120602.

[9] Shaheen, A. F., Moselhi, O., & Zayed, T. (2019). Forecasting concrete compressive strength using hybrid artificial neural network with particle swarm optimization. Advances in Engineering Software, 134, 13-23.

[10] Ay, M., & Kisi, O. (2021). Concrete compressive strength estimation by using various machine learning methods. Computer Methods in Applied Mechanics and Engineering, 372, 113385.

[11] Tsai, M., Huang, T., Chang, F., & Lin, K. (2018). A data-driven model for strength prediction of high-performance concrete. Applied Soft Computing, 63, 36-47.

[12] Khedekar, M. S., & de Brito, J. (2021). Prediction of concrete compressive strength: Importance of feature selection in machine learning models. Automation in Construction, 124, 103481.

[13] Sun, S., Wu, Y., Liu, F., & Sun, Z. (2019). Bayesian optimization for hyperparameters of support vector machine and its performance in the prediction of high-performance concrete strength. Computer-Aided Civil and Infrastructure Engineering, 34(10), 904-920.

[14] Ralaivola, L., Wu, Q., Wang, X., & Zhu, D. (2020). Concrete compressive strength predictive model based on random forest. Computers, Materials & Continua, 64(1), 251-269.

[15] Zou, J., Li, X., Zhang, J., & Guan, H. (2021). Predictive modeling of concrete compressive strength using Q-learning. Automation in Construction, 121, 103428.

[16] Li, W., Zhao, Y., & Liu, S. (2020). Deep learning approaches for predicting compressive strength of concrete: a comparison study. Journal of Intelligent & Fuzzy Systems, 39(4), 5595-5606.

[17] Dong, Y., Ma, Y., & Ding, Z. (2020). Application of LSTM neural network in the prediction of compressive strength of concrete. Computer Methods in Applied Mechanics and Engineering, 365, 113021.

[18] Singh, K., Raju, K., & Patel, C. (2021). Strength prediction of concrete using ensemble machine learning models. Structures, 28, 1987-1995.

[19] Chu, J., Wang, W., & Zhang, D. (2021). Concrete compressive strength prediction based on generative adversarial nets. Neural Computing and Applications, 33(6), 1927-1940.

[20] Sarıdemir, M. (2009). Prediction of compressive strength of concretes containing metakaolin and silica fume by artificial neural networks. Advances in Engineering Software, 40(9), 856-863.