International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Predicting Rupture Strength of Concrete Using Machine Learning Techniques

Satvik B.Walikar[1], Amit K. Yadawadi[2], Pruthviraj Shinde[3], Sankarshan S. Joshi[4], Sheetal B. Patil[5], Santosh B. Chikkbagewadi[6],

[1]Dept. of CSE S. G. Balekundri Institute of Technology. Belagavi, Karnataka, India. satvikwalikar16@gmail.com
[2]Dept. of CSE S. G. Balekundri Institute of Technology. Belagavi, Karnataka, India. amityadawadi@gmail.com
[3]Dept. of CSE S. G. Balekundri Institute of Technology. Belagavi, Karnataka, India. pruthvishinde9449@gmail.com
[4]Dept. CSE S. G. Balekundri Institute of Technology. Belagavi, Karnataka, India. sankarshanjoshi9@gmail.com
[5]Dept. of CSE S. G. Balekundri Institute of Technology. Belagavi, Karnataka, India. Sheetaladi13@gmail.com
[6]Dept. of CSE S. G. Balekundri Institute of Technology. Belagavi, Karnataka, India. santoshhc@gmail.com
DOI: https://doi.org/10.55248/gengpi.4.523.42366

ABSTRACT—

Concrete's rupture strength is a crucial component in determining its structural integrity and longevity. In this research, we look into the uses of an algorithms that utilise machine learning to predict the compressive strength of concrete based on a number of different factors, such as the type of cement used, the water-cement ratio, the curing time, and the age of the concrete. Utilising different regression techniques, such as decision tree, random forest regression, and linear regression, we gathered data on concrete samples and their respective compressive strengths. The data was prepared by being divided into training and testing sets, choosing important features, and scaling the data as needed. We then used the training data to train the random forest model, and performance was improved by optimising the hyperparameters. The model was evaluated using various metrics, including mean squared error (MSE), accuracy score, root mean squared (RMS), and root absolute error (RAE). The findings demonstrate that the random forest regressor method is a useful tool for the construction sector since it can predict concrete's compressive strength with high accuracy.

Keywords—Compressive strength, Concrete, Random forest regressor, Python.

I. INTRODUCTION

In the present system, civil engineers find the compressive strength of concrete by using a compressive testing machine (CTM). It's one of the methods to find the compressive strength; it requires a machine to find the strength of concrete. Compression tests are performed to characterise the behaviour of a material under compressive loading.

To assess the various qualities of the material being tested, pressure is applied to a specimen during the test using compression plates or unique tools mounted on a universal testing machine. Manual testing of concrete strength involves the casting and curing of concrete specimens, followed by destructive testing in a laboratory. It's a lengthy process because we need to create a cube by using various elements like Cement, Fly ash, Blast-Furnace, Super-Plasticizer, Fine-Aggregate, Coarse-Aggregate, Water, Age(days). And we should wait for at least 7 or more days for curing and here we need manpower for the creation of the concrete cube and for other work. This method is time-consuming. labour-intensive and expensive. It is not suitable for real-time monitoring of concrete strength in construction projects.

Based on the concrete's mix design, the suggested approach use machine learning to forecast its compressive strength. The shortcomings of the current system are overcome by this technique. Compared to manual testing, it is quicker, less expensive, and labor-intensive. It can offer real-time monitoring of concrete strength in building projects, enabling prompt interventions and mix design changes. By enhancing the machine learning model with more data, the predictions' accuracy can be steadily increased. In order to train the model and obtain the desired results, we are taking into account a number of characteristics, including cement, fly ash, blast furnace, superplasticizer, fine aggregate, coarse aggregate, water, and age (days). Here, all the process is done in the system, so there is no need for the creation of cubes and all. And it's cost-efficient and easy to use.

A. Related Work

Before starting with the project a brief study was made on the currently existing systems in this domain. The work carried out by notable research scholars is studied in depth and its excerpt is documented here in study of the literature.

VeereshKarikatti [1] defines an artificial neural network (ANN) as a machine learning model that creates connections between inputs and outputs through interconnected data networks called neurons. Artificial neural networks can process complex data and representative data like the human brain. It has a neuron structure similar to the structure of the human brain. The weight of connected neurons determines the neural network's ability to process information effectively.

Vimal Rathakrishnan [2] used the Python programming language and the "PyCaret" library on the Google Collab platform for data analysis and modelling. A brief summary of the step-by-step process for training, optimizing and validating BML (augmented machine learning) models to predict compressive strength: vi. Model Evaluation: Compare models based on metrics such as R2, RMSE, MAE, MSE, RMSLE, and MAPE. The algorithm with the best performance is selected.

Ashraf Shaqadan [3] uses a method that includes data collection, pre-processing, model selection, optimization, model validation, model evaluation and analysis. / Advert. Random Forest (RF) requires special configuration. First, a set of predictive variables (m) is selected to identify tree nodes. Also, recursively select the number of trees to load. RF provides a performance measure against out-of-bag (OOB), the default example in the boot training program. The residual and pseudo R2 can also be calculated for the OOB state. In addition, RF assigns different significance scores to each estimator.

Muhammed Amir Shafiq [4] defined neural networks as a combination of high-power handling and nonlinear adaptive filters called neurons to solve various problems. In a multilayer feedforward neural network (MFNN), there is no feedback connection between the layers. The access layer takes the weighted objects and forwards them to the first hidden layer. There can be many hidden layers in MFNN and the output of the last hidden layer becomes the input of the output layer. In this paper, a backpropagation MFNN was used and neurons were trained using scaled conjugate gradient (SCG) and Levenberg-Marquardt (LM) functions.

Zhi Wan [5] described four machine learning models: linear regression (LR), support vector regression (SVR), cloud gradient boosting (XGBoost), and artificial neural network (ANN). The material compresses the strength of the rock. Based on the composition and age of the four ML models, linear regression with an R-square below 0.90 performed the worst, while the other three ML models all had an R-square above 0.

Hai-Van [6] uses an RF algorithm was proposed to estimate the compressive strength of concrete containing GGBFS with . Collected 453 test samples to build the RF model. The database work is divided into two parts, 70% training data and 30% test data are used for the validation phase of the established RF model. Ground Granulated Blast Furnace Slag (GGBFS) was used as an additional cementitious material in Portland Concrete. GGBFS is a product of glass granular material formed after molten blast furnace slag is rapidly cooled with water. GBGFs can replace 35-65% of Portland Cement in the Rock. Fix strong rock and persist by creating a denser matrix using GGBFS, switch to Portland Cement.

Ayaz Ahmad [7] uses AdaBoost is supervised machine learning that uses community learning to train multiple models. Model is referred to as weak learners working together to achieve a common goal. AdaBoost adjusts the weights for each sample, giving incorrect samples more weight. Its aim is to reduce prejudice and inequality in educational care. An infinite number of decision trees are used during training. An initial model is created and samples are not classified as significant. Only these examples are used to train the next model. This process is repeated until a weak learner is formed.

II. METHODOLOGY

Here, we are predicting the rupture strength of concrete using machine learning approaches. Similar types of processes are used by all machine learning algorithms as the block diagram below illustrates.

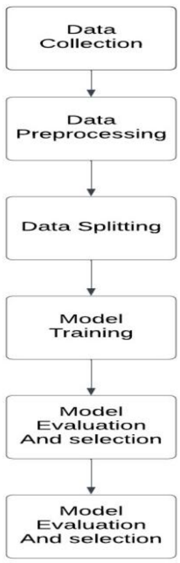
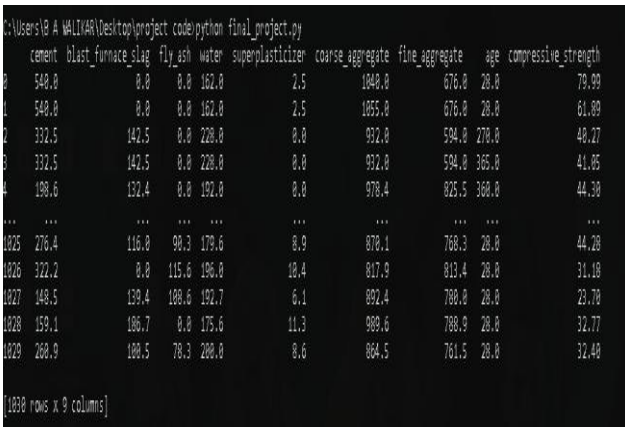


Fig. 1 Block diagram

First, we gather the data that will be used to train the model. This is the very first and most important step. The experiment or a reputable online respiratory service should gather the dataset. The information must meet our requirements. The sample dataset that we are using to train and test the model is depicted in the figure below.



C:\Users\B A WALIKAR\Desktop\project code>python final_project.py

	cement	blast_furnace_slag	fly_ash	water	superplasticizer	coarse_aggregate	fine_aggregate	age	compressive_strength
0	540.0	0.0	0.0	162.0	2.5	1040.0	676.0	28.0	79.99
1	540.0	0.0	0.0	162.0	2.5	1055.0	676.0	28.0	61.89
2	332.5	142.5	0.0	228.0	0.0	932.0	594.0	270.0	40.27
3	332.5	142.5	0.0	228.0	0.0	932.0	594.0	365.0	41.05
4	198.6	132.4	0.0	192.0	0.0	978.4	825.5	360.0	44.30
...	...	...	...	...	...	...	...	...	...
1025	276.4	116.0	90.3	179.6	8.9	870.1	768.3	28.0	44.28
1026	322.2	0.0	115.6	196.0	10.4	817.9	813.4	28.0	31.18
1027	148.5	139.4	108.6	192.7	6.1	892.4	780.0	28.0	23.70
1028	159.1	186.7	0.0	175.6	11.3	989.6	788.9	28.0	32.77
1029	260.9	100.5	78.3	200.0	8.6	864.5	761.5	28.0	32.40

[1030 rows x 9 columns]

Fig. 2 Sample dataset

By using the pandas library in python we can load CSV file.

In the process of building a predictive model, the second step involves preparing the dataset for analysis through data preprocessing, which includes cleaning the dataset and handling missing values, outliers, and categorical variables.

The Simple Imputer technique offered by scikit-learn can be used to manage missing values, which can happen for a variety of causes. The SimpleImputer() method, which is based on a specified strategy like mean, median, or mode, replaces the NaN (Not a Number) values with a designated placeholder.

To be used in the prediction model, categorical variables, which indicate discrete categories or groups, must be transformed into numerical values. Techniques like label encoding and one-hot encoding can be used to accomplish this.

The SimpleImputer() method can be customised with an appropriate approach to fill in the missing values, making it an effective tool for handling missing value. This step ensures that the dataset is complete and appropriate for analysis.

In general, data preparation is a crucial stage in creating a predictive model since it makes sure that the dataset is precise, comprehensive, and appropriate for the algorithms employed in the model. Data splitting, the third stage, divides the pre-processed dataset into training sets and testing sets. The models will be trained using the training set, and their performance will be assessed using the testing set.

The training of models is the fourth phase. Utilising the Python scikit-learn module, we train the models on the training data. In this instance, we train machine learning models such as random forest regressors, both linear regression and the decision tree.

Model selection and evaluation is the fifth phase. Use measures like the model's score (as calculated using the score() method), mean squared error, root mean square error (RMSE), and mean absolute error to assess the trained model's performance on the testing set. We will choose the best-fit model based on the score and the outcomes.

Sixth step is Model Deployment once the model has been trained and evaluated, save it to a file using the pickle module in Python. Use the model in a web application or other production environment to make predictions based on fresh data.

These type of stages we'll take to put our model into practise, and they serve as base for all ML model building processes.

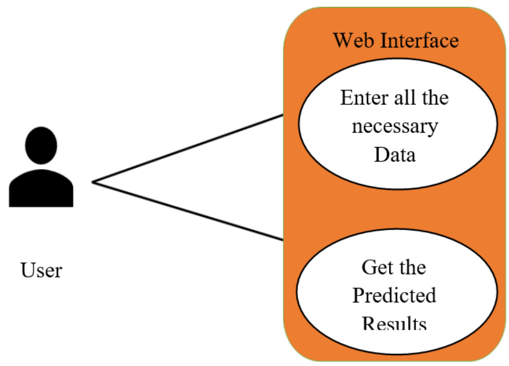


Fig.3 Use case diagram

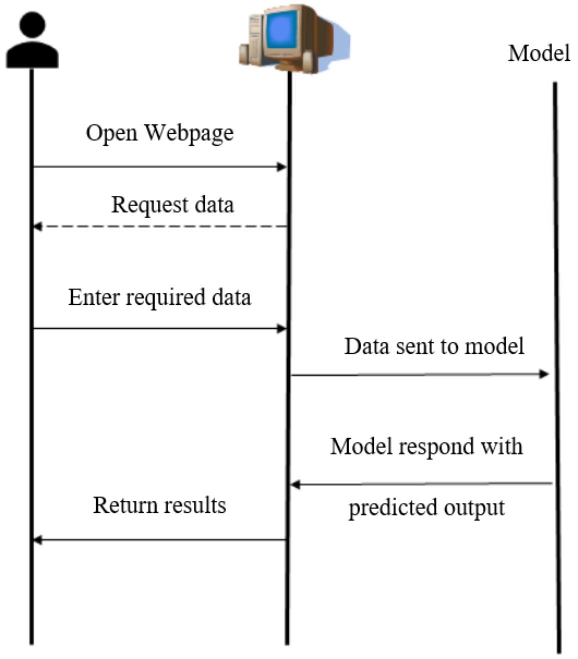


Fig.4 Sequence diagram

III.EXPERIMENTAL RESULTS

We are used 7 components for training that are Cement, Slag, Flyash, Superplasticizer, Coarseaggregate, Water, Fineaggregate, Age(No. of Days for curing).

All are in kg/m^3 unit.

We found that the Random Forest Regressor (RFR), which outperformed both the Decision Tree Regressor (DTR) and Linear Regression (LR) models, had the highest score of the three regression models, coming in at 0.892. The results of the models are displayed in the following graph.

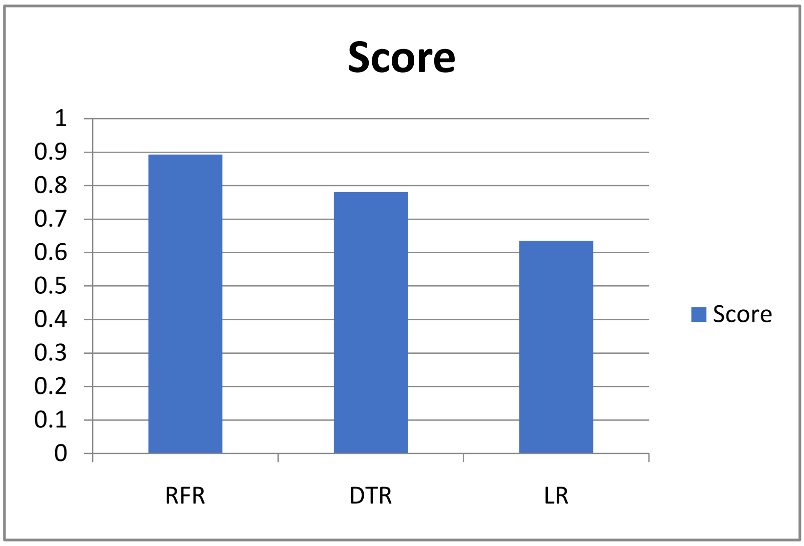


Fig. 5 Score of RFR DTR and LR model

The RFR model exhibited the lowest mean squared error (MSE) of 27.57 when the performances of the three regression models, Random Forest Regression (RFR), Decision Tree (DT), and Linear Regression (LR), were compared. These findings suggest that for the particular dataset utilised in the study, the RFR model may be better suited for outcome prediction than the DT and LR models. The mean squared error (MSE) of the models is displayed in the following graph.

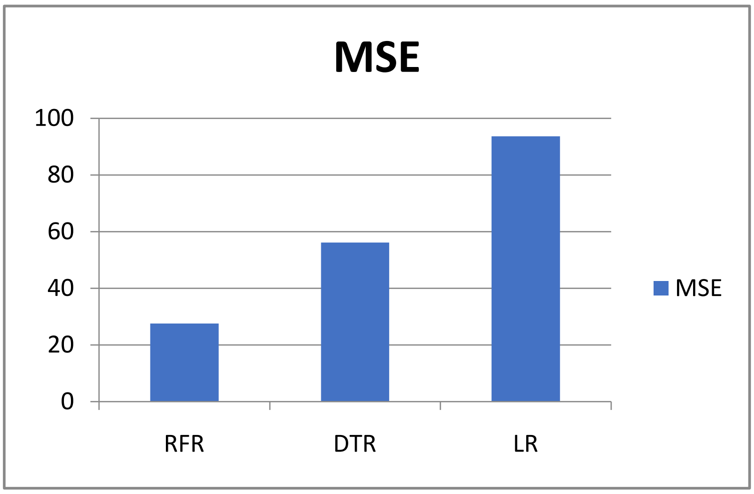


Fig. 6 Mean Squared Error (MSE) of RFR DTR and LR

Among the three models evaluated, the Random Forest Regressor (RFR) had the lowest mean absolute error (MAE) that is 3.60. When compared to the Decision Tree (DT) and Linear Regression (LR) models, indicating that it outperformed the other models in terms of accuracy. The following chart shows the Mean Absolute Error (MAE) of the Models.

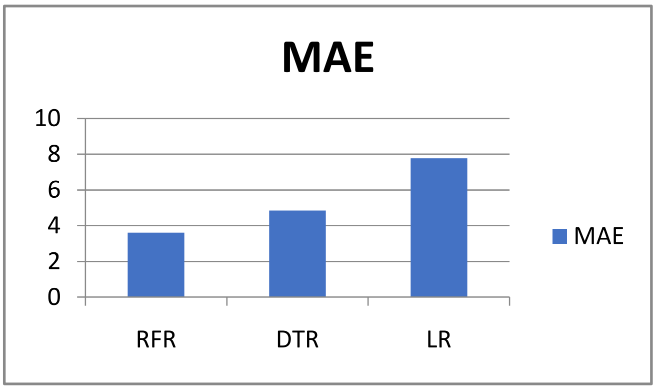
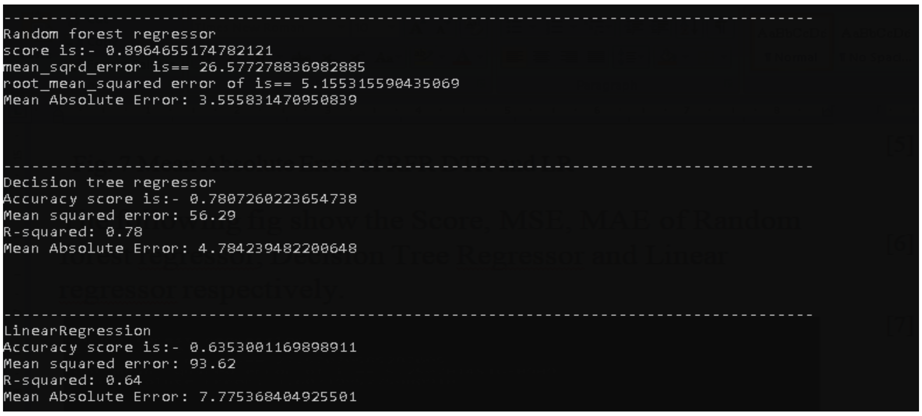


Fig. 7 Mean Absolute Error of RFR DTR and LR

The following figure show the Score, MSE, MAE of Random Forest Regressor, Decision Tree Regressor and Linear Regressor respectively.



```
--------------------------------------------------------------------------------
Random forest regressor
score is:- 0.8964655174782121
mean_sqrd_error is== 26.577278836982885
root_mean_squared error of is== 5.155315590435069
Mean Absolute Error: 3.555831470950839


--------------------------------------------------------------------------------
Decision tree regressor
Accuracy score is:- 0.7807260223654738
Mean squared error: 56.29
R-squared: 0.78
Mean Absolute Error: 4.784239482200648


--------------------------------------------------------------------------------
LinearRegression
Accuracy score is:- 0.6353001169898911
Mean squared error: 93.62
R-squared: 0.64
Mean Absolute Error: 7.775368404925501
```

Fig. 8 Score, MSE, MAE of RFR DTR and LR

Based on the evaluation of the Score, Mean Squared Error, R-squared, and Mean Absolute Error metrics, the analysis indicates that the Random Forest Regressor outperforms the other two models. Therefore, it is recommended to select the Random Forest Regressor for our case.

After thorough analysis and evaluation of the performance metrics, the results indicate that the Random Forest Regressor model outperformed the other two models (Decision Tree and Linear Regression). This conclusion was based on a comparison of evaluation metrics, such as Score, Mean Squared Error, R-squared, and Mean Absolute Error. The Random Forest Regressor approach has been found to be the best fit for our unique situation as a consequence.

After selecting the Random Forest Regressor model, we proceeded to save it for future use. To enable non-experts in data analysis and civil engineering to use the model, we developed a user-friendly web interface. This interface allows users to input relevant variables and obtain a reliable prediction of the target variable using the Random Forest Regressor model.

The web interface was designed to accommodate individuals with varying levels of civil engineering knowledge. Thus, the input variables are labelled in plain language and categorized accordingly. Additionally, we provided clear instructions on how to use the web interface and interpret the results. Overall, this approach allows for accurate and efficient predictions while enabling non-experts to leverage the power of machine learning in civil engineering applications.

To illustrate, we have included some snapshots of the web interface and the results page.

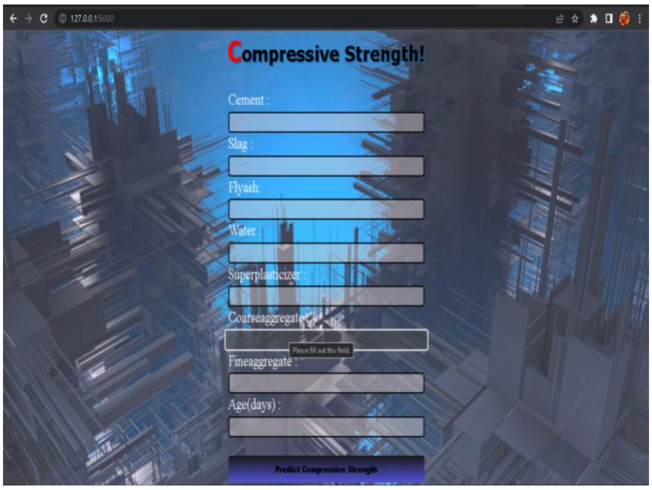


Fig.9 Web user interface for getting data

Once the user inputs all the necessary variables into the web

Once the user has entered all the necessary input values, the data will be transmitted to the backend where the model is located. The prediction is then made using the Random Forest Regressor model, and the resulting compressive strength of concrete is displayed to the user in the results tab.

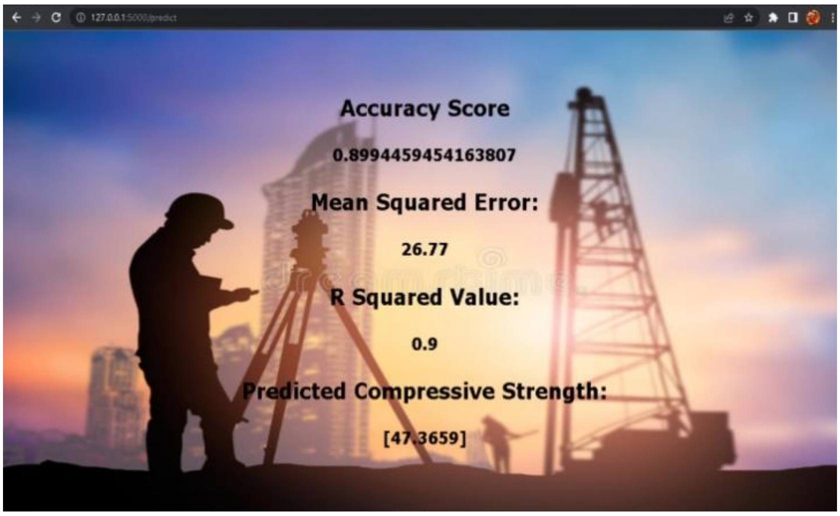


Fig.10 Result page showing the predicted compressive strength

IV. Conclusion

This project demonstrates the effective application of machine learning technique in civil engineering. By implementing the Random Forest Regressor model and developing a user-friendly web interface, we have provided a practical and efficient solution for predicting the rupture strength of concrete. The outcomes show that the Random Forest Regressor model outperformed the other two models (Decision Tree and Linear Regression) based on evaluation metrics such as Score, Mean Squared Error, R-squared, and Mean Absolute Error. By creating a user-friendly web interface, we have enabled non-experts in data analysis and civil engineering to leverage the power of machine learning to obtain accurate predictions. Additionally, to ensure simplicity of use for those with varied degrees of civil engineering experience, the interface was built with simple instructions and identified input variables.This project highlights the potential of machine learning in civil engineering and serves as a valuable contribution to the field.

REFERENCES

[1] Veeresh.Karikatti1 ,M.V.Chitawadagi , Ishwaragouda S. Patil ,Sanjith J , Mahesh Kumar C L ,Kiran B. M , "ANN Model for Predicting Compressive Strength of Alkali Activated Slag ConcreteCured at Environmental Temperature", NeuroQuantology Volume 20, issues 1 June 2022, doi:10.14704/nq.2022.20.6.NQ22044.

[2] Vimal Rathakrishan,Salmi a bt.Beddu& Ali Najah Ahmed Vimal, "Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms", ResearchGate, 2022, doi: 10.1038/s41598-022-12890-2.

[3] Ashraf Shaqadan "Prediction of concrete mix strength using random forest model." International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 22 (2016) pp. 11024-11029.

[4] Muhammad Amir Shafiq "Predicting the Compressive Strength of Concrete using Neural Network and Kernel Ridge Regression", Reasearch Gate, 2016 doi: 10.1109/FTC.2016.7821698.

[5] Zhi Wan, *Yadingxu and Brankosavija, "On the Use of Machine Learning Models for Prediction of Compressive Strength of Concrete:Influence of Dimensionality Reduction on the Model Performance", ResearchGate Materials 2021, 14, 713, doi:10.3390/ma14040713.

[6] Hai-Van Thi Mai, Thuy-Anh Nguyen , Hai-Bang Ly and Van Quan Tran "Prediction Compressive Strength of Concrete Containing using Random Forest Model". Hindawi Volume 2021.

[7] Ayaz Ahmad, Waqas Ahmad, Fahid Aslam, PanuwatJoyklad "Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques". Elsevier, 2021.

[8] Suhaila Khursheed, J. Jagan1, Pijush Samui and Sanjay Kumar2 "Compressive strength prediction of fly ash concrete by using machine learning techniques". ResearchGate, 2021, doi10.1007/s41062-021-00506-z.

[9] Hongwei Song, Ayaz Ahmad , Furqan Farooq , Krzysztof Adam Ostrowski "Compressive strength prediction of fly ash concrete by using machine learning techniques". ResearchGate, Materials, 2021, doi:10.3390./ma14040794.

[10] Ayaz Ahmad, Furqan Farooq, Pawel Niewiadomski, Krzysztof Ostrowski ''Prediction of Compressive Strength of Fly Ash Based Concrete Using Individual and Ensemble Algorithm." 2021 ResearchGate, Materials, doi:10.3390/ma14040794

[11] Mohammed Hmood Mohana "The determination of ground granulated concrete compressive strength-based machine learning models". ResearchGate, Vol8, No.2, June 2020 pp.1011-1023.

[12] Hao Ling, Chunxiang Qian, Wence Kang, Chengyao Liang a, Huaicheng Chen "Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment." ResearchGate, 2019.