**Application of Supervised Learning Method for the Prediction of Shear and Flexural Strength of Reinforced Recycled Aggregate Concrete Beam**

##### **Nandita Suresh Kamath (2018B4A20868P, BITS Pilani, Pilani Campus)**

Under the Guidance of Dr. Subhasis Pradhan

1. **ABSTRACT**

Shear failure remains as a very prominent reason that needs to be taken into account in the design of concrete beams as it is critical in the design of structural members. To prevent such detrimental situations, the beams are designed with reinforcements for additional support. These beams are called reinforced cement concrete beam and are more capable of sustaining vertical load and preventing bending. Due to the low tension, high compression features of concrete, steel **reinforcement** is added in **concrete** to improve its **tensile resistance.** The shear failure of recycled aggregate concrete (RAC) beams even more dangerous due to the inferior nature of recycled coarse aggregate (RCA). This study aims to develop an effective Machine Learning model to accurately estimate the shear strength of RAC beams by analyzing existing data on experiments conducted on RAC beams to be of help to civil engineers. to help predict shear strength.

1. **INTRODUCTION – TYPES OF REINFORCED CONCRETE BEAM**
2. Based on the on percentage of reinforcement present, RCC beams maybe classified into;

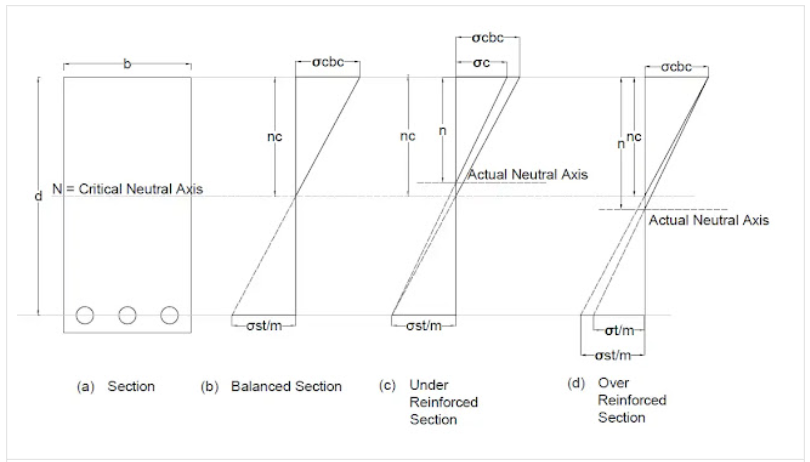


Figure 1

### **Balanced Sections**

The stress in concrete and steel in a balanced-reinforced section reaches its permitted value at the same time. This means that when a load is applied to the beam, both the compressive and tensile zones yield simultaneously. The breakdown is rapid, and there is minimal warning of distress in tension failure since the concrete will crush and the tensile steel will yield at the same time. Balanced steel is the percentage of steel that corresponds to this section, and critical neutral axis nc is the neutral axis. Steel-reinforced concrete moment-carrying parts should typically be under-reinforced so that users of the structure are alerted to approaching collapse.

### **Under Reinforced Section**

The percentage of steel provided in an under reinforced section is lower than in a balanced section. As a result, it is less expensive. The neutral axis will shift higher, nc > n. An under-reinforced section is one in which the steel stress achieves its maximum allowed value before the concrete. This means that as the bending moment on the beam increases, the tension steel will give while the concrete will not reach its ultimate failure point. An “under-reinforced” concrete yield in a ductile manner when the tension steel yields and stretches, revealing a considerable deformation and warning before its ultimate failure. This feature ensures the safety of this form of reinforcement.

### **Over Reinforced Section**

The percentage of steel provided in an over reinforced section is higher than in a balanced section. As a result, the true neutral axis moves downward, n>nc. In this part, concrete stress reaches its maximum allowable amount, although steel is not fully stressed. Concrete is brittle and breaks down quickly, making it dangerous. The over reinforced portion is uneconomical since steel is significantly more expensive than concrete. A bending force causes a beam to bend, resulting in a minor curvature.

1. Depending on the placement of reinforcement, RCC beams are classified as;

1. A singly reinforced beam is where the concrete element is only reinforced near the tensile face and the reinforcement, called tension steel, is designed to resist the tension.

2. A doubly reinforced beam is one in which the concrete element is reinforced near the compressive face in addition to the tensile reinforcement to aid the concrete resist compression. Compression steel is the name for this type of reinforcement. If the engineer limits the section's dimensions, further reinforcement is required when the compression zone of the concrete is insufficient to resist the compressive moment (positive moment).

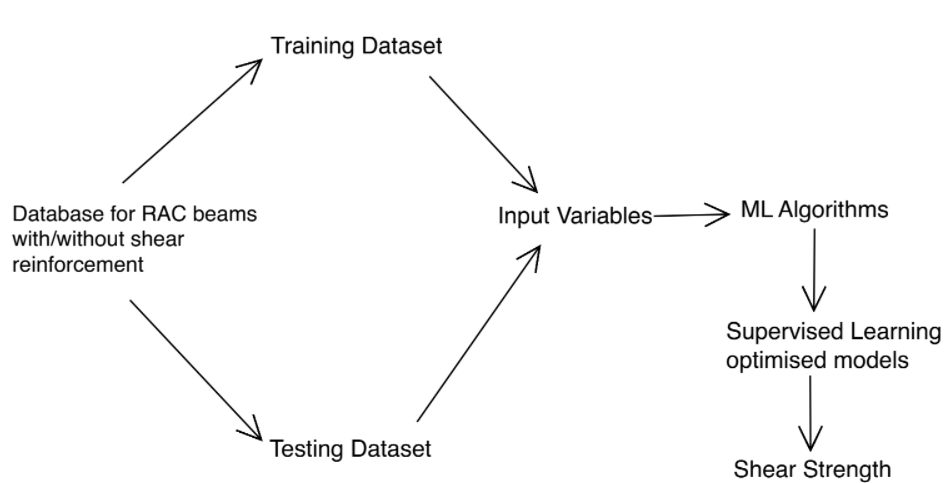
1. Depending on the supporting system, RCC beams may be classified into the following;
2. **A simple concrete beam is one with a single span supported at one end and no constraint at the support.**
3. **A continuous beam is one that is supported by more than two supports. It can be a single beam with intermediate supports of smaller beams for lengthy spans between columns or walls, or a single continuous beam for the entire length of the construction with intermediate column or wall supports.**
4. **A beam having two spans with or without constraint at the two extreme ends is known as a semi-continuous beam.**
5. **Cantilever beams are supported on one end and project beyond the support or wall on the other.Simple Concrete Beams** refers to the beam having a single span supported at its end, lacking a restraint at the support.
6. A continuous beam is one that is supported by more than two supports. It can be a single beam with intermediate supports of smaller beams for lengthy spans between columns or walls, or a single continuous beam for the entire length of the construction with intermediate column or wall supports.
7. A beam having two spans with or without constraint at the two extreme ends is referred to as a semi-continuous beam.
8. Cantilever beams are supported on one end and protrude beyond the support or wall on the other.
9. T – Beam/L – Beam- Most reinforced concrete constructions have monolithic slabs and beams. As a result, the beam is part of the floor system. A portion of the slab bends along with the beam when it bends. As a result, T beams operate as intermediate beams in a floor system, whereas L beams act as end beams. Flanged beams are those that have a piece of the slab acting alongside the beam to resist compressive pressures. [1]
10. **LITERATURE REVIEW**

Over a decade has been spent researching computational approaches for estimating the shear strength of RC beams. To forecast the strength of RC deep beams, Cheng and Cao [2] used multivariate adaptive regression splines, whereas Gandomi [3] used Genetic Programming and simulated annealing. To anticipate the strength of FRP RC beams without stirrups, Kara [4] used gene expression programming. Perez [5] used genetic programming to predict beam strength in the absence of shear reinforcement. Zhang [6] used numerical solutions and mechanism equations to forecast prestressed RC member shear behavior and failure, as well as the shear strength of one-way slabs without stirrups.

It has been noted from the literature review that there is a critical need for ML research in the shear strength estimation in RAC beams with and without shear reinforcement. Such a data-driven approach can lead to an efficient predictive model for practical as well as research use.

1. **MODEL DEVELOPMENT AND WORK FLOW**

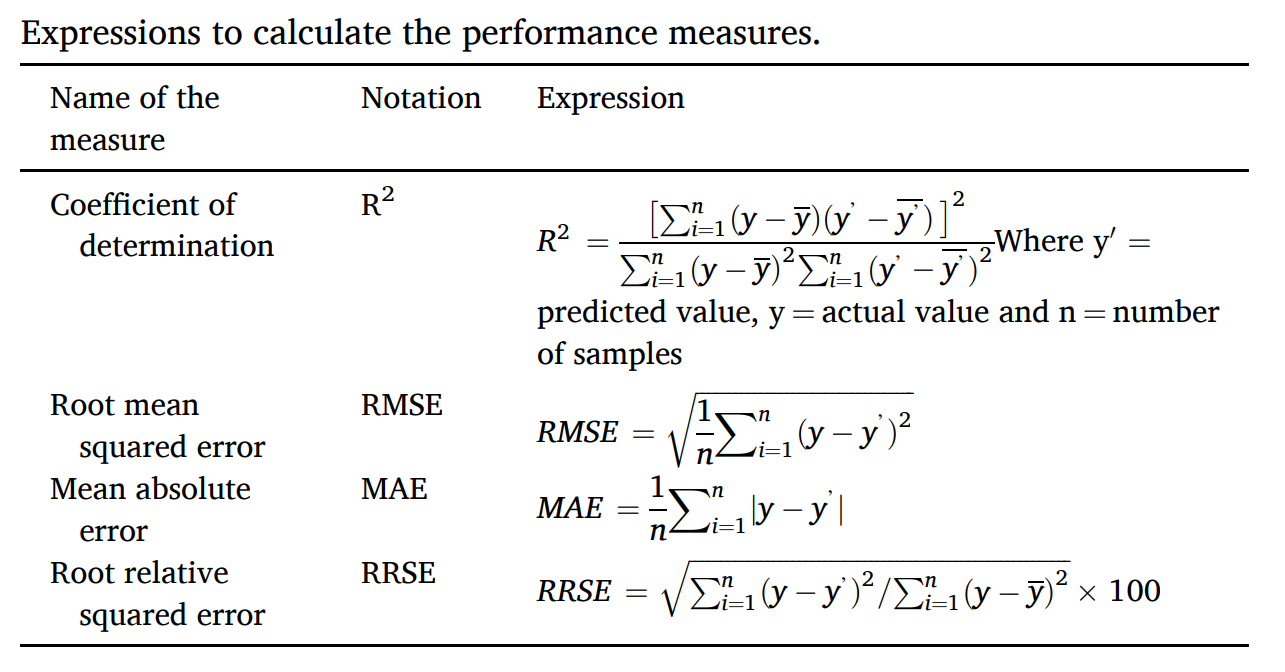
The shear strength of RAC beam is predicted with the help of a predictive Machine Learning model. Independent predictor variables are used to model the result of the target variable, shear strength. As this target variable is given in the dataset, a supervised ML Regression model is implemented. This ML model is developed after compilation, exploration and cleaning of the data. The language used to develop this model is Python and the online notebook used to code and deploy the model is Kaggle. The project is approached by defining the problem statement, analysing the dataset in python using data exploration tools, identifying predictor and target variables, dividing the sample space into training and testing set and trying multiple regression algorithms to deploy the best model in production. The dataset is applied on numerous models such as Multiple Linear Regression model, Decision Tree, AdaBoost, Support Vector Machine and Random Forest to find a suitable model with maximum accuracy. The models are analysed using the eleven to twelve input features and the model producing maximum accuracy for the given dataset is optimised. As seen in the figure, the dataset of RAC beams with and without shear reinforcement are divided into training and testing sets and ML models are trained on its input parameters to build a predictive model.

****

*Figure 2 – Framework of the proposed ML model*

1. **PERFORMANCE EVALUATION**

The mathematical efficiency of a model is measured by calculating how well it fits a curve. These are quantified by parameters such as coefficient of determination R2, root mean square error (RMSE), and mean absolute error (MAE). The R2 predicts how efficiently the proposed model can predict the original data. The RMSE is associated with the learning process of the ML model and is the cost function. RMSE and MAE are associated with accuracy and goodness of fit. The mathematical formulae used determine each of these measures are summarized below. A good predictor model is indicated by a R2 value close to 1.00. [7]

****

*Table 1 – Mathematical expressions for performance measures*

1. **DATABASE USED**

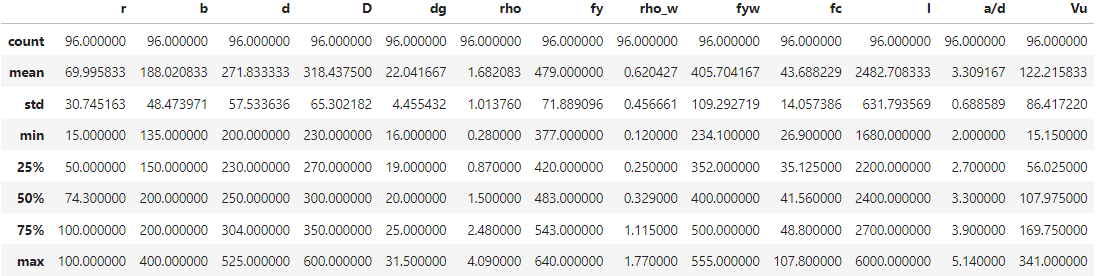
This study utilizes the basic concepts of a supervised ML model and uses RAC beam properties such as replacement ratio, beam size, aggregate size, longitudinal reinforcement ratio, effective span and shear span as predictor variables to forecast shear strength and construct a predictive model to help civil engineers save time and money on cumbersome experiments in the design and construction process. The dataset given below depicts a few experimental values of RAC beams consisting of its measurement that was compiled from multiple research papers. This database was then split into a training and testing set to apply different ML models and the ones providing highest accuracy were optimized to solve this supervised ML Regression problem. The database comprises of 91 experimental results of RAC beams without Shear Reinforcement (Dataset 1) and 96 experimental results with shear reinforcement (Dataset 2). [8]

1. **BASIC DATA EXPLORATION AND STATISTICAL ANALYSIS**

Data analysis and exploration are performed to gauge the spread of the data. Each data point is analysed for its type, and whether it affects the values of the target variable. Python library functions are used to determine the size of the dataset and the number of rows and columns. The exact statistical details of the data such as minimum, maximum, average and standard deviation is found. Each variable is identified as categorical or continuous and the missing values are flagged and fixed. Data Mining is conducted to deal with inadequate and missing data. They are calculated from given data or assumed based upon experience. Finally, columns that don’t affect the target variable are deleted.



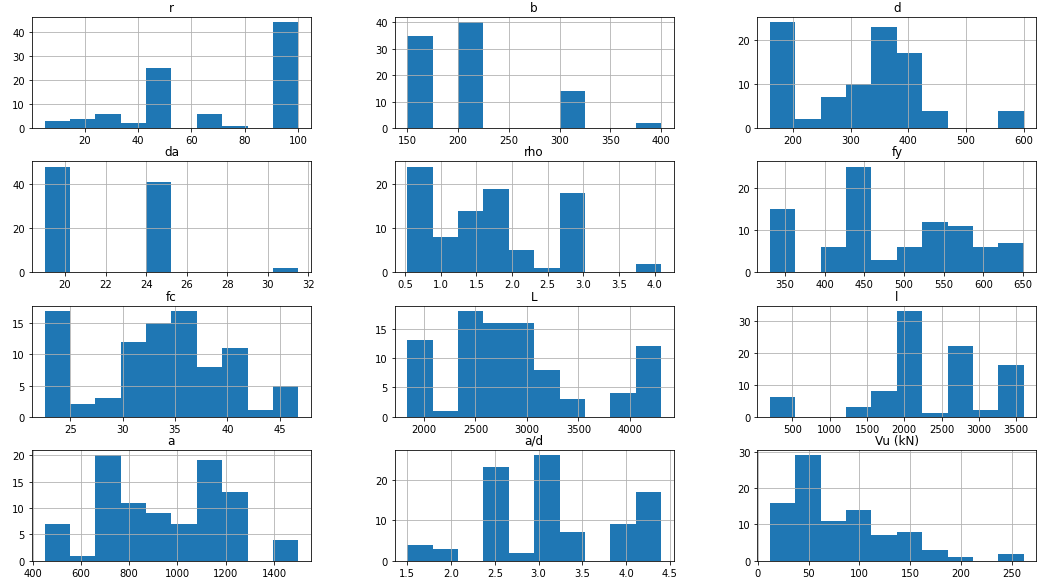
*Table 2 – Statistical range of parameters in Database 1*

**

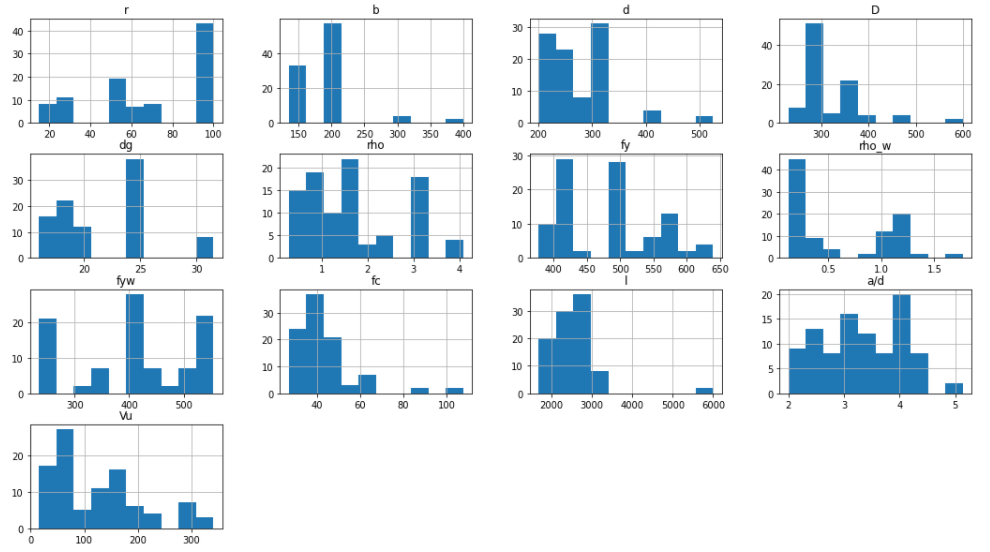
*Table 3 – Statistical range of parameters in Database 1*

1. VISUAL EXPLORATORY DATA ANALYSIS

Figure 3 and 4 depict the histogram plots for the Continuous Input Variables. The ideal outcome is a bell curve however, in reality, datasets are rarely perfect. Hence, the variables are accepted as long as there is good distribution across its range. [9]

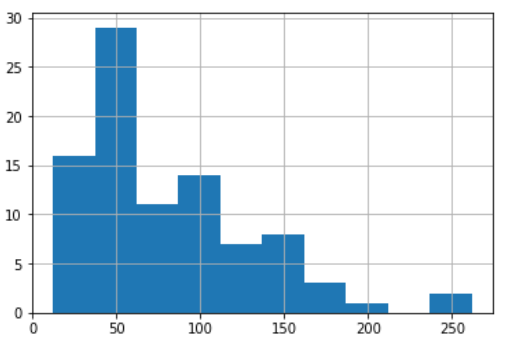


*Figure 3 - Distribution of the predictor variables for Database 1*

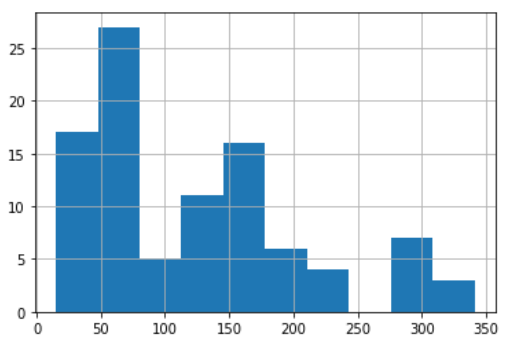
**

*Figure 4 - Distribution of the predictor variables for Database 1*

More importantly, the target variable is analysed and plotted on a histogram. The data distribution of the target variable depicted in Figure 5 and 6 for each dataset has sufficient range and frequency which makes our dataset suitable to build a model upon.



*Figure 5 – Distribution of the target variable for Database 1*

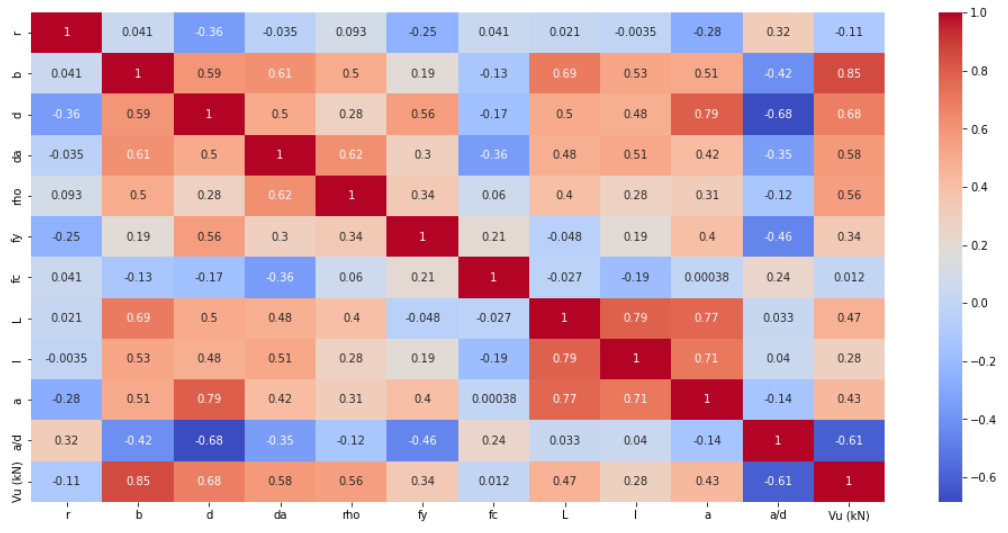


*Figure 6 – Distribution of the target variable for Database 2*

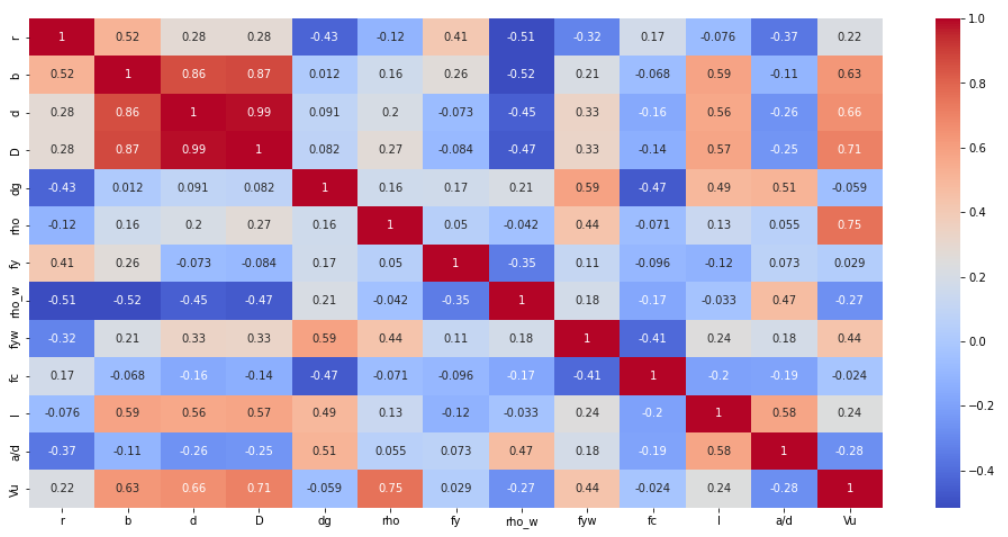
1. **ANALYSIS OF INPUT PARAMETERS AFFECTING SHEAR STRENGTH**

Since the target and predictor variable is continuous in nature, a correlation plot is used to understand the effect of each predictor variable on the target variable and the relationship between the two. The correlation value can be calculated for any two numeric variables. A negative correlation indicates an inversely proportional relation (downward trend) and a correlation between 0 and 1 indicates a directly proportional relation (upward trend). By studying the correlations between Target variable and all other predictor variables, the intensity of the effects of the predictors can be checked. Variables that are positively correlated and are directly proportional to each other have an increasing trend and variables that are negatively correlated and are inversely proportional to each other have a decreasing trend. An absolute correlation value of 0.5 and above signifies a strong relationship. For the datasets, beam size, maximum aggregate size, longitudinal reinforcement ratio, and shear span effect the target variable the most. Figure 7 and 8 depict the correlation plots for each dataset. A dark blue or red color indicates high correlation.

It is observed that for Dataset 1, beam size followed by the ratio a/d effects shear strength the most. This result was also found by Shahnewaz and Alam [10] , from the study of Meta-model of Optimal Prognosis to identify the important parameters. However, for Dataset 2, longitudinal reinforcement ratio followed by beam size effects shear strength the most. This result was also found by Rahman [7] while predicting shear strength of steel reinforced beams.



*Figure 7 – Heat map and correlation matrix of the variables in database 1*

**

*Figure 8 – Heat map and correlation matrix of the variables in database 2*

1. **SPLITTING DATASET AND MODEL SELECTION**

The existing dataset is shuffled, selected at random and divided 2 sets; Training and Testing with a 70-30% ratio. The training set is used to build the model and the testing set is used to validate it. Linear Regression, Decision Tree, L2 (Ridge) Regression, Support Vector Machine (Linear and RBF Kernel), ANN, Random Forest, Gradient Boosting and AdaBoost are the models applied on the dataset. The model providing with the highest accuracy is selected and optimized. Later, the model is saved and the predicted values are analysed against the original values of the training set. The efficiency produced by each model is depicted graphically below.

1. **PERFORMANCE MEASURES**

## The R2 values corresponding to each ML model for both datasets are analysed and compared with one another. For Dataset 1, Gradient Boosting Algorithm give maximum accuracy at 92.26% followed by Ridge regression at 90.65%. For Dataset 2, Linear Regression gives maximum accuracy at 94.27% followed by Ridge regression at 93.76%. Figure 9 and 10 compares the accuracy associated with each model for both datasets.

## Figure 9 – Performance of each ML model for Dataset 1

## Figure 10 – Performance of each ML model for Dataset 2

**CONCLUSION**

This study presents 9 machine learning-based approaches to predict the shear strength of beams with and without reinforcement. Based on the database, this paper developed models using Linear regression, Ridge Regression, SVM – (Linear and RBF kernel), Decision Tree, Neural Network, Random Forest, Gradient Boosting and AdaBoost. The datasets are tested on all 9 models and optimized to receive maximum efficiency for prediction. The findings of the present study can be concluded as:

* Gradient Boosting model gave the highest accuracy of **93.5%** to predict shear strength of RAC beams without shear reinforcement and Linear Regression gave the highest accuracy of **94.2%** to predict shear strength of RAC beams with shear reinforcement.The accuracy of these ML models is satisfactory with respect to our dataset and the comparison of the values predicted by the model with the original laboratory findings supports the applicability of this model for shear strength prediction. New input variables can be added to the algorithm as mentioned in the appendix to find new strength values.
* Beam size and the ratio a/d effects shear strength the most for Dataset 1. For Dataset 2, longitudinal reinforcement ratio followed by beam size effects shear strength the most. Hence, these values must be calculated cautiously as the greatly effect the target variable.
* The original vs Predicted values of shear strength are depicted in a scatter plot in Figure 11 and 12. The line fits almost at a 45 degree angle ensuring close values.

*Figure 11 – Graph of Original vs Predicted Values for Dataset 1*

*Figure 12 – Graph of Original vs Predicted Values for Dataset 1*

# References

|  |  |
| --- | --- |
| [1] | S. Sharma, "civilengineering.blog," [Online]. |
| [2] | C. J.-S. C. M.-T. Cheng M-Y, "Nature-inspired metaheuristic multivariate adaptive regression splines for predicting refrigeration system performance.," *Soft Computing,* p. 477–489, 2017. |
| [3] | A. A. S. D. S. M. Gandomi AH, "An empirical model for shear capacity of RC deep beams using genetic-simulated annealing," *Archives of Civil and Mechanical Engineering,* p. 354–369, 2013. |
| [4] | K. IF, "Prediction of shear strength of FRP-reinforced concrete beams without stirrups based on genetic programming," *Advances in Engineering Software,* p. 295–304, 2011. |
| [5] | C. A. R. J. M.-A. F. Pérez JL, "Optimization of existing equations using a new Genetic Programming algorithm: application to the shear strength of reinforced concrete beams.," *Advances in Engineering Software,* p. 82–96, 2012. |
| [6] | T. Zhang and P. V. D. J. Oehlers, "Shear strength of FRP RC beams and one-way slabs without stirrups," *Journal of Composites for Construction,* 2014. |
| [7] | K. S. A. ,. N. I. K. K. I. S. M. Jesika Rahman, "Data-driven shear strength prediction of steel fiber reinforced concrete beams using machine learning approach," *Engineering Structures ,* 2021. |
| [8] | S. K. S. V. B. Subhasis Pradhan, "Shear performance of recycled aggregate concrete beams: An insight for design aspects," *Construction and Building Materials,* pp. 593-611, 2018. |
| [9] | U. E. M. S. A. Tadesse Gemeda Wakjiraab, "Machine learning-based shear capacity prediction and reliability analysis of shear-critical RC beams strengthened with inorganic composites," *Case Studies in Construction Materials,* 2022. |
| [10] | A. M. Shahnewaz M, "Genetic algorithm for predicting shear strength of steel fiber reinforced concrete beam with parameter identifictaion and sensitivity analysis," *Journal of Building Engineering,* 2020. |