**Concrete strength prediction model analysis Report And Coding Part**

**Contents**

[**Dataset & its Goal :** 3](#_Toc154149198)

[**Abstract:** 4](#_Toc154149199)

[**Data Characteristics:** 4](#_Toc154149200)

[**Summary Statistics:** 4](#_Toc154149201)

[**Variable Information:** 5](#_Toc154149202)

[**Project analysis report** 6](#_Toc154149203)

[Introduction: 6](#_Toc154149204)

[Data Exploration 6](#_Toc154149205)

[Data Cleaning 7](#_Toc154149206)

[Handling Missing Values 7](#_Toc154149207)

[Scaling the Data 7](#_Toc154149208)

[Data Splitting 7](#_Toc154149209)

[Model Choice 7](#_Toc154149210)

[Model Evaluation 7](#_Toc154149211)

[Conclusion 8](#_Toc154149212)

[**Project Code** 9](#_Toc154149213)

# **Dataset & its Goal :**

The given dataset contains information on the compressive strength of concrete. The data has nine input variables and one target variable.

The input variables are:

* cement: the amount of cement (in kg/m^3) in the concrete mixture.
* blast Furnace: the amount of blast furnace slag (in kg/m^3) in the concrete mixture.
* flyAsh: the amount of fly ash (in kg/m^3) in the concrete mixture.
* water: the amount of water (in kg/m^3) in the concrete mixture.
* superplasticizer: the amount of superplasticizer (in kg/m^3) in the concrete mixture.
* courseAggregate: the amount of coarse aggregate (in kg/m^3) in the concrete mixture.
* fineAggregate: the amount of fine aggregate (in kg/m^3) in the concrete mixture.
* age: the age (in days) of the concrete sample when the compressive strength was measured.

The target variable is:

* strength: the compressive strength of the concrete (in MPa).

The goal of this dataset is to develop a predictive model that can accurately estimate the compressive strength of concrete based on the given input variables. This model can be useful for engineers and construction companies to optimize the design of concrete mixtures and ensure that the resulting concrete structures meet the required strength specifications.

Data Type: multivariate

# **Abstract:**

Concrete is the most important material in civil engineering. The concrete compressive strength is a highly nonlinear function of age and ingredients. These ingredients include cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate.

# **Data Characteristics:**

The actual concrete compressive strength (MPa) for a given mixture under a specific age (days) was determined from laboratory. Data is in raw form (not scaled).

# **Summary Statistics:**

* Number of instances (observations): 1030
* Number of Attributes: 9
* Attribute breakdown: 8 quantitative input variables, and 1 quantitative output variable
* Missing Attribute Values: None

# **Variable Information:**

Given is the variable name, variable type, the measurement unit and a brief description.

The concrete compressive strength is the regression problem. The order of this listing

corresponds to the order of numerals along the rows of the database.

Name Data Type -- Measurement -- Description

* Cement (component 1) -- quantitative -- kg in a m3 mixture -- Input Variable
* Blast Furnace Slag (component 2) -- quantitative -- kg in a m3 mixture -- Input Variable
* Fly Ash (component 3) -- quantitative -- kg in a m3 mixture -- Input Variable
* Water (component 4) -- quantitative -- kg in a m3 mixture -- Input Variable
* Superplasticizer (component 5) -- quantitative -- kg in a m3 mixture -- Input Variable
* Coarse Aggregate (component 6) -- quantitative -- kg in a m3 mixture -- Input Variable
* Fine Aggregate (component 7) -- quantitative -- kg in a m3 mixture -- Input Variable
* Age -- quantitative -- Day (1~365) -- Input Variable
* Concrete compressive strength -- quantitative -- MPa -- Output Variable

**Project analysis report:**

**Project Analysis Report**: Concrete Compressive Strength Prediction

## Introduction:

This report presents an analysis of the Concrete Compressive Strength data set, which contains information about various components of concrete and their compressive strength. The goal of this analysis is to build a model that can predict the compressive strength of concrete based on its components.

## Data Exploration

The Concrete Compressive Strength data set consists of 1030 samples with 9 features:

* cement (component 1)(kg in a m3 mixture)
* blast furnace slag (component 2)(kg in a m3 mixture)
* fly ash (component 3)(kg in a m3 mixture)
* water (component 4)(kg in a m3 mixture)
* superplasticizer (component 5)(kg in a m3 mixture)
* coarse aggregate (component 6)(kg in a m3 mixture)
* fine aggregate (component 7)(kg in a m3 mixture)
* age (day)
* strength (MPa)

The first step in data exploration was to load the data set and examine its structure. The data set was loaded using the pandas library in Python, and its structure was examined using various pandas functions, such as head(), describe(), info(), and shape().

## Data Cleaning

The data set was cleaned by renaming some of the feature columns to make them more readable, and by checking for and dropping any duplicate rows.

## **Data** Preprocessing

The data set was preprocessed by performing several steps, such as handling missing values, scaling the data, and splitting the data into training and testing sets.

Handling Missing Values**:**

There were no missing values in the data set.

Scaling the Data**:**

To make sure that every feature was on the same scale and to keep any one feature from taking center stage in the model, the data was scaled using the StandardScaler function from the sklearn.preprocessing library.

Data Splitting**:**

The sklearn.model\_selection library's train\_test\_split() method was used to divide the data into training and testing groups.

## Model Choice

The data set served as the training and testing set for a number of models, including XGBoost, Lasso, Ridge, Decision Tree, and Linear regression. R-squared score and mean squared error (MSE), among other evaluation measures, were used to determine which model performed the best.

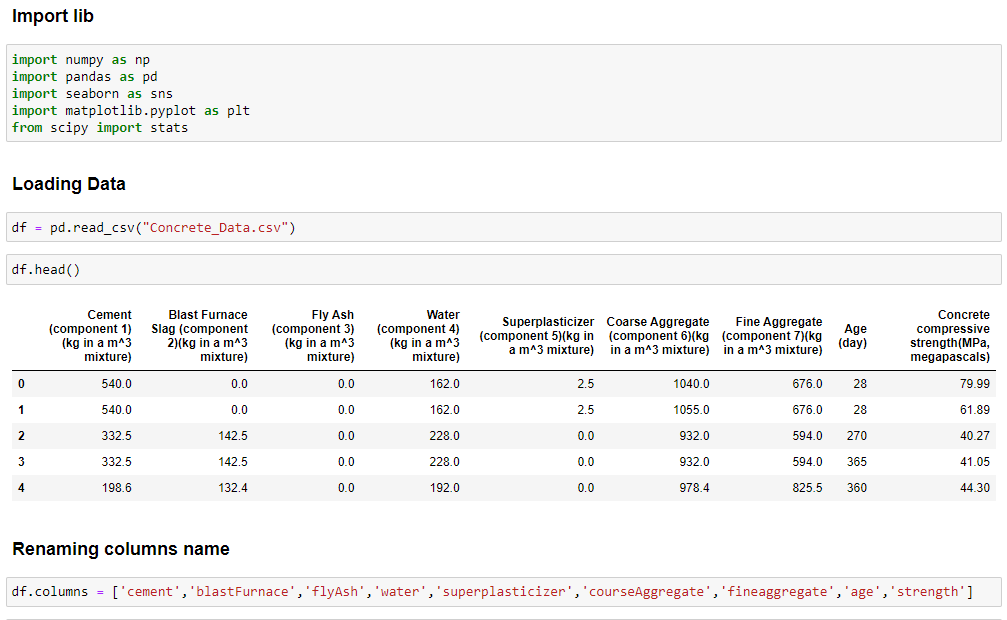
## Model Evaluation

The test data set was used to assess the chosen model, and a variety of assessment measures were used to analyze its performance. A variety of plots, including residual and scatter plots, were used to show the results.

Conclusion**:**

In summary In summary, the goal of this study was to develop a model that could predict concrete's compressive strength based on its constituent parts. In order to train and evaluate different machine learning models, the data set was examined, cleared, and preprocessed. Based on how well the model performed on the test data set and how many assessment criteria examined its performance, the best model was chosen. The completed model. can be useful in the building sector and used to predict the compressive strength of concrete.

**Project Code:**

****

