

## Model Development Phase Template

Date	10 July 2024
Team ID	SWTID1720162737
Project Title	Predicting Compressive Strength Of Concrete Using Machine Learning
Maximum Marks	6 Marks

### Model Selection Report

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

### Model Selection Report:

Model	Description	Hyperparameters	Performance Metric ( R2 Score)
Gradient Boosting regression.	This estimator builds an additive model in a forward stage-wise fashion; it allows for the optimization of arbitrary differentiable loss functions. In each stage a regression tree is fit on the negative gradient of the given loss function.	RandomizedSearchCV	R2_Score= 94.59%
Linear Regression	Linear Regression fits a linear model with coefficients $w = (w_1, \dots, w_p)$ to minimize the residual sum	---	R2_Score=58.61 %

	of squares between the observed targets in the dataset, and the targets predicted by the linear approximation.		
Ridge Regression	This model solves a regression model where the loss function is the linear least squares function and regularization is given by the l2-norm. Also known as Ridge Regression or Tikhonov regularization. This estimator has built-in support for multi-variate regression.	---	R2_Score= 58.62%
Lasso regression	Lasso regression is a technique for building regression models that addresses overfitting. It works by adding a penalty term to the model that shrinks the coefficients of less important features towards zero. This encourages a simpler model that focuses on the most influential factors, potentially improving generalizability and reducing the risk of overfitting to the training data.	---	R2_Score= 58.78 %
Random Forest Regression	A random forest is a meta estimator that fits a number of decision tree regressors on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control overfitting.	RandomizedSearchCV, GridSearchCV	R2_Score= 93.23%

Decision Tree Regression	Decision Tree Regression predicts continuous values by splitting the data based on features. It builds a tree-like structure where each internal node represents a question about a feature, and the branches represent possible answers. The model predicts the target value based on the path a data point takes through the tree, reaching a leaf node with an average target value for similar data points.	--	R2_Score= 88.49%
XGB Regression	XGBoost (eXtreme Gradient Boosting) is a powerful regression algorithm that builds upon the concept of gradient boosting like GBRT. It excels at handling complex relationships in data and offers several advantages like Improved Performance, Scalability, Interpretability.	RandomizedSearchCV, GridSearchCV	R2_Score= 94.18 %