Random Forest Regression

Definition - Random Forest Regression is an **ensemble learning method** that builds multiple decision trees and combines their predictions to improve the accuracy and robustness of the model. Each tree is trained on a random subset of the data, and the final prediction is typically the average of the predictions from all individual trees.

Definition of Ensemble Learning - Machine learning technique where multiple models, often called "weak learners" or "base models," are combined to create a stronger overall model. The idea is that by aggregating the predictions of several models, the ensemble can achieve better accuracy, robustness, and generalization than any single model. Ensemble learning helps reduce errors like bias and variance by leveraging the strengths of different models.

Feature Scaling - **Feature scaling is not required** in Random Forest Regression for the same reason as Decision Trees. The splits within each tree are based on raw feature values, and the algorithm doesn't rely on distance-based calculations, making scaling unnecessary.

How It Works

- The algorithm begins by creating multiple subsets of the original data through a process called bootstrapping, where each subset is a random sample with replacement from the dataset.
- For each subset, a decision tree is built using a randomly selected subset of features. This random feature selection ensures that each tree in the forest learns different patterns and reduces the likelihood of overfitting.
- Once all the trees are built, the model predicts the target value by aggregating the predictions of all trees. For regression, this aggregation is usually done by averaging the predictions of each tree.
- The final output is a more accurate and generalized prediction due to the ensemble of diverse trees.

Ensemble Learning

Ensemble Learning is a machine learning technique where multiple models, often called "weak learners" or "base models," are combined to create a stronger overall model. The idea is that by **aggregating the predictions of several models**, the ensemble can achieve better accuracy, robustness, and generalization than any single model. Ensemble learning helps reduce errors like bias and variance by leveraging the strengths of different models.