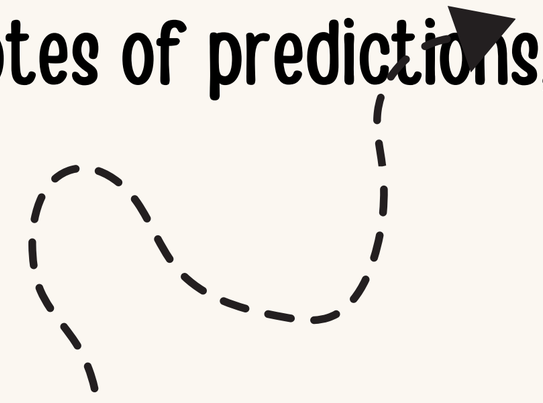
RANDOM FOREST

Introduction

Random forest is an ensemble of decision trees

It combine multiple classifiers to solve a complex problem and to improve the performance of the model.

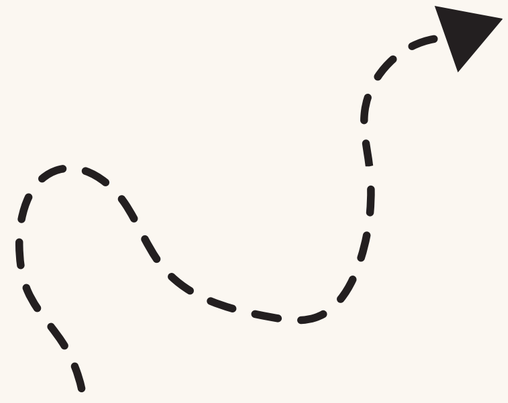
Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.



A depth unlimited decision tree or deep decision tree is a strong learner.

The random forest's main goal is to combine multiple overfitted decision trees to reduce variance or overfitting.

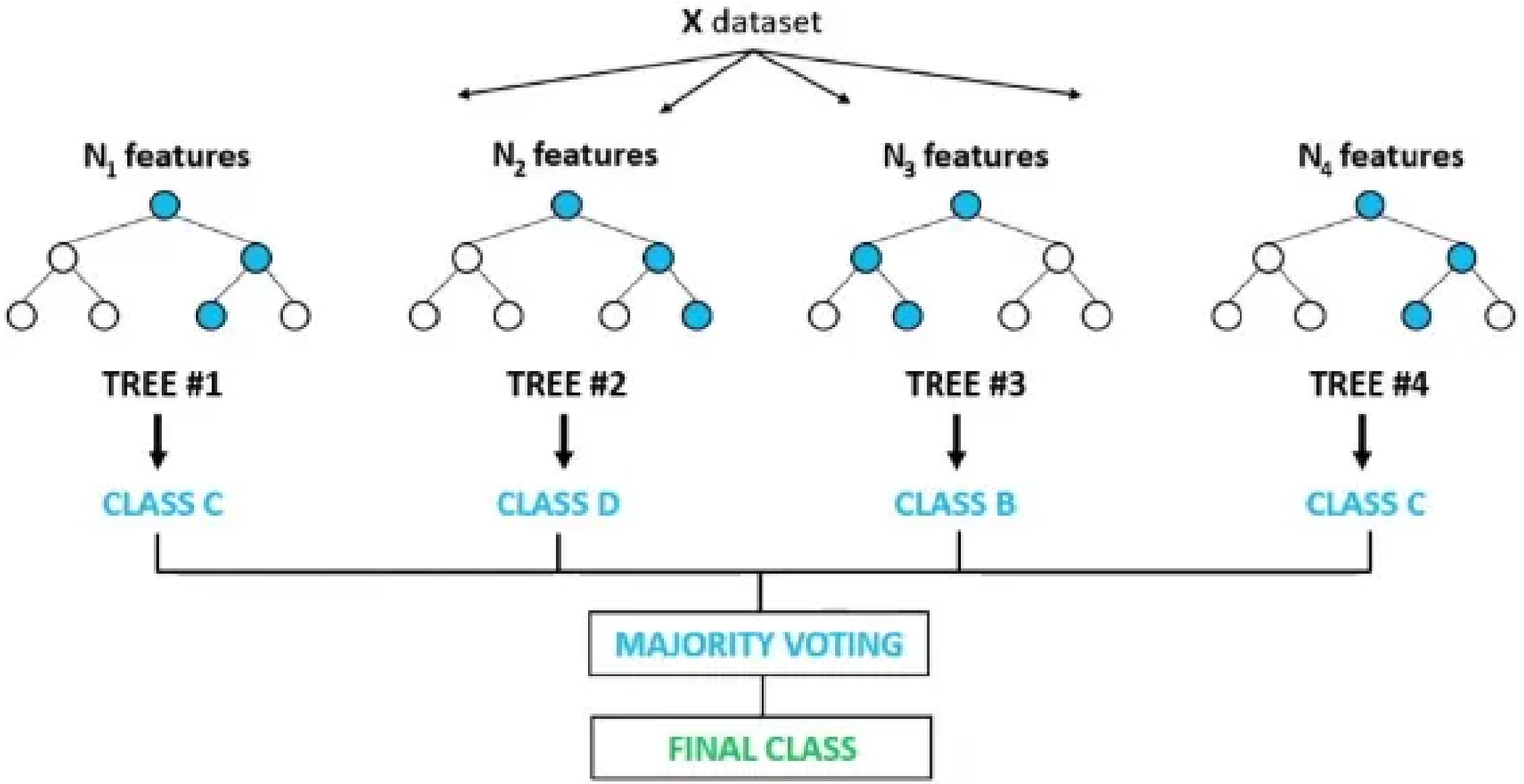
It adds randomness by using random feature subsampling to create decorrelated decision trees.



Advantages

1. Handle large datasets with high dimensionality
2. handle both numerical and categorical features
3. Enhances the accuracy of the model and prevents the overfitting issue

Random Forest Classifier



WHY RANDOM FOREST?

- It takes less training time as compared to other algorithms.
- It predicts output with high accuracy, even for the large dataset it runs efficiently.
- It can also maintain accuracy when a large proportion of data is missing.