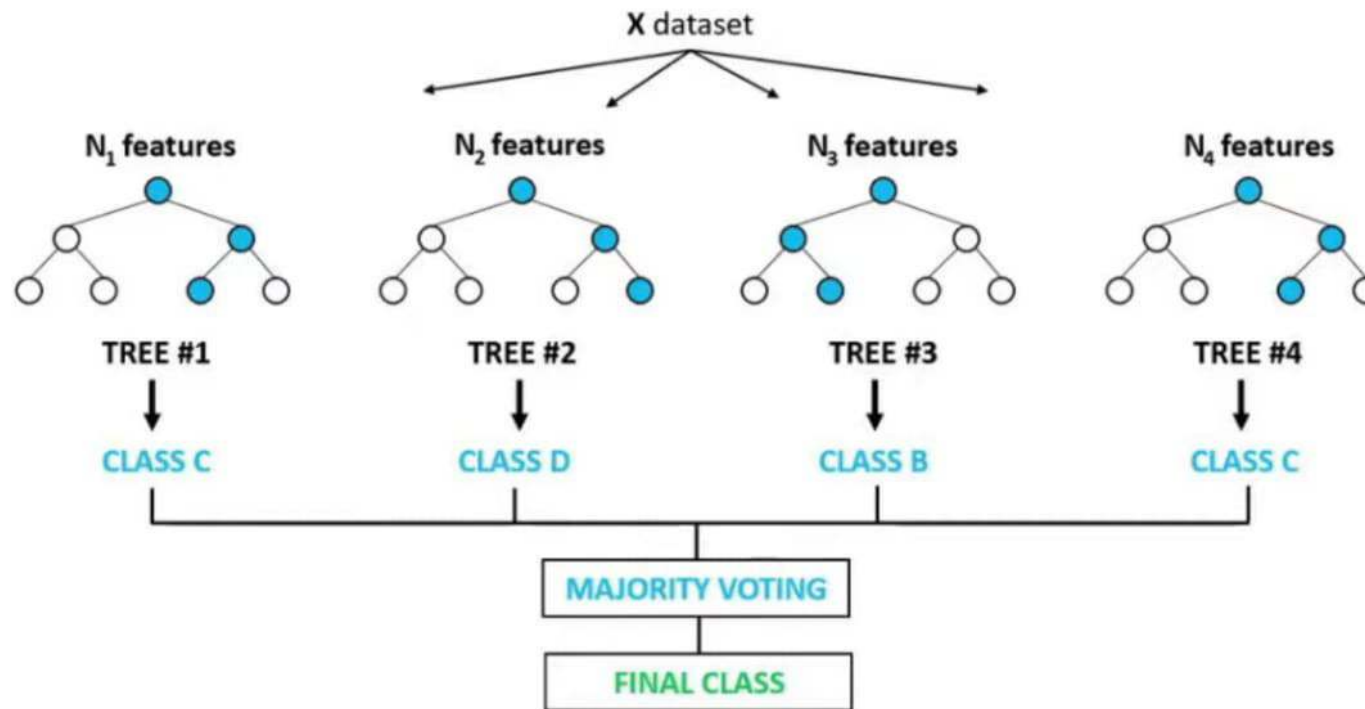
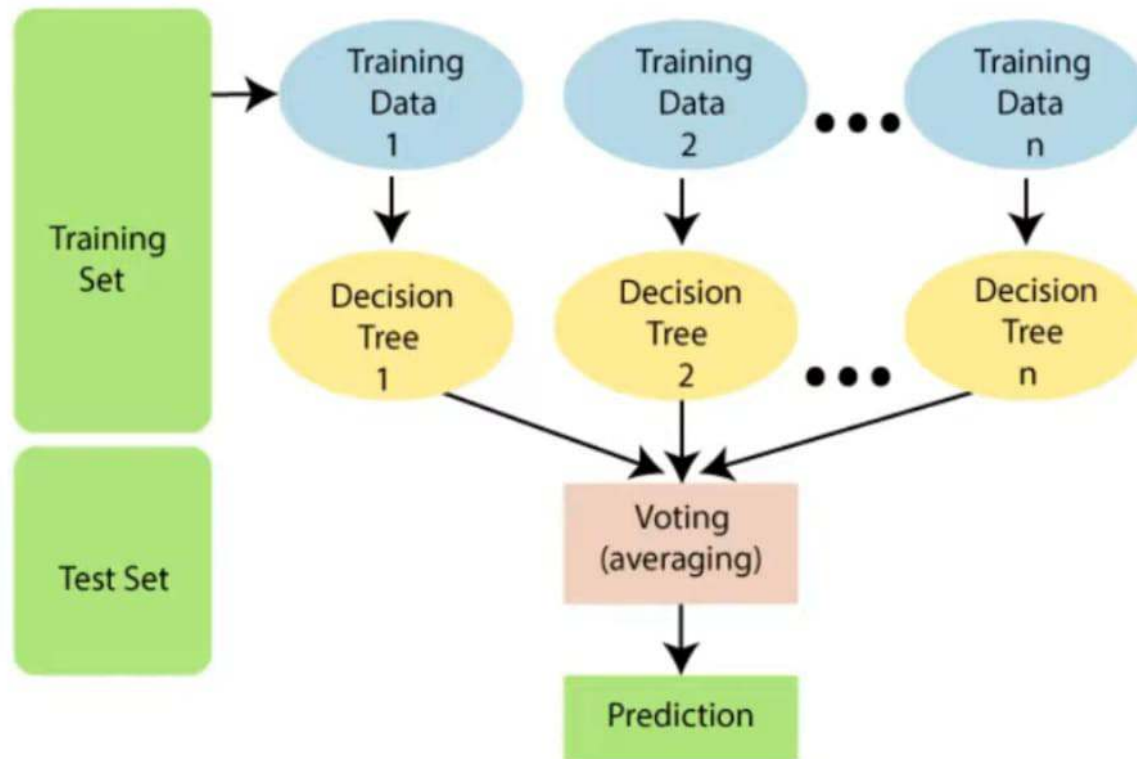


# Random Forest Algorithm



# Random Forest Algorithm



# Why use 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.

# Assumptions for Random Forest

- ⊠ There should be some actual values in the feature variable of the dataset so that the classifier can predict accurate results rather than a guessed result.
- ⊠ The predictions from each tree must have very low correlations.

# How does Random Forest algorithm work?

**Step-1:** Select random  $K$  data points from the training set.

**Step-2:** Build the decision trees associated with the selected data points (Subsets).

**Step-3:** Choose the number  $N$  for decision trees that you want to build.

**Step-4:** Repeat Step 1 & 2.

**Step-5:** For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

# Applications of Random Forest

1. **Banking:** Banking sector mostly uses this algorithm for the identification of loan risk.
2. **Medicine:** With the help of this algorithm, disease trends and risks of the disease can be identified.
3. **Land Use:** We can identify the areas of similar land use by this algorithm.
4. **Marketing:** Marketing trends can be identified using this algorithm.

# How does Random Forest algorithm work?

