

In order to make a prediction for a given observation, we typically use the mean or the mode of the training observations in the region to which it belongs. Since the set of splitting rules used to segment the predictor space can be summarized in a tree, these types of approaches are known as decision tree methods.

Decision trees and the building blocks for random forest model.

Random Forest consists of a large number of individual decision trees that operates as an ensemble.

What is ensemble?

Ensemble method uses multiple learning algorithm to obtain better predictive performance than could be obtained from any of the constituent learning algorithm alone.

Each individual tree in the random forest spits out a class prediction and the class with the most votes become our model prediction – the wisdom of crowds

The predictions made by the individual trees need to have low correlation with each other

Now how does the random forest ensure that the behaviour of each individual tree is not too correlated with the behaviour of any of the other trees in the model?

With help of these two methods:

- Bagging
- Feature Randomness

Decisions trees are very sensitive to the data they are trained on — small changes to the training set can result in significantly different tree structures.

Bagging – Random forest takes advantage of this by allowing each individual tree to randomly sample from the dataset with replacement, resulting in different trees . This is known as bagging.

Feature Randomness – Use different features to make decisions