DECISION TREES ASSUMPTIONS

Random Forest is an ensemble of decision trees. Each tree is constructed independently, and the final prediction is a combination (often an average or voting) of the predictions of individual trees.

1. No Assumption of Linearity:

Random Forest does not assume a linear relationship between features and the target variable. It can capture complex, non-linear relationships in the data.

2. No Assumption of Normality:

There is no assumption regarding the distribution of the features or the target variable. Random Forest can handle both categorical and continuous variables without assuming a specific distribution.

3. Robust to Overfitting:

Random Forest tends to be robust to overfitting, thanks to the combination of multiple trees and the use of randomness in feature selection and bootstrapping (sampling with replacement).

4. Feature Importance:

Random Forest provides a measure of feature importance, indicating the contribution of each feature to the model's predictive performance. This is based on how often a feature is used to split nodes and the improvement it brings to the model.

5. Insensitivity to Outliers:

Random Forest is less sensitive to outliers compared to some other models. The averaging or voting mechanism helps mitigate the impact of extreme values.

6. Parallelization:

Random Forest can be easily parallelized, making it suitable for parallel computing environments. This enables efficient training on large datasets.

7. Bias-Variance Tradeoff:

While Random Forest tends to have low variance (due to the ensemble nature), it may have some bias. However, this bias is often acceptable in practice.

8. Tree Depth and Number of Trees:

The hyperparameters related to individual decision trees (such as tree depth) and the number of trees in the ensemble are crucial for model performance. However, Random Forest is not very sensitive to the exact specification of these hyperparameters.

9. No Assumption of Independence:

Unlike assumptions in statistical models, Random Forest does not assume independence between observations. It can handle correlated or dependent observations.