# Random Forests:

1. Ensemble of Decision Trees:

* RF is an ensemble method that creates multiple decision trees during training. Each DT is a weak Learner, but when combined, they form a strong learner.

2. Randomness in Model Building:

* Bootstraping: For each tree in the forest, a random subset of the training data is selected. This is called bootstraping, which means sampling with replacement.

3. Training:

· Each DT is built using the subset of data and features, and it tries to learn the decision boundary in the feature space.

4. Prediction:

· Once the for DT's is built, predictions for new data are made by aggregating the predictions of all individual trees.

· For classification, this is done using majority voting

5.

· For regression, the final prediction is the average of all the predictions from the individual trees.

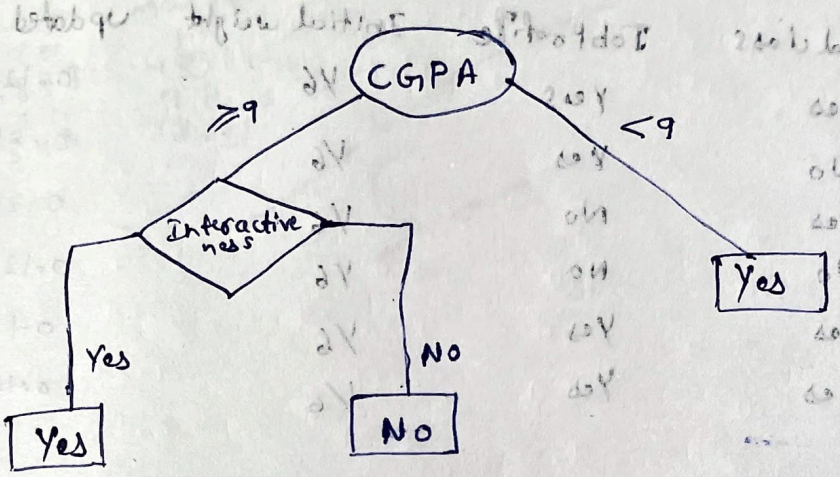
5. Model Evaluation.

Evaluate the performance of the RF model using metrics such as accuracy or MSE.

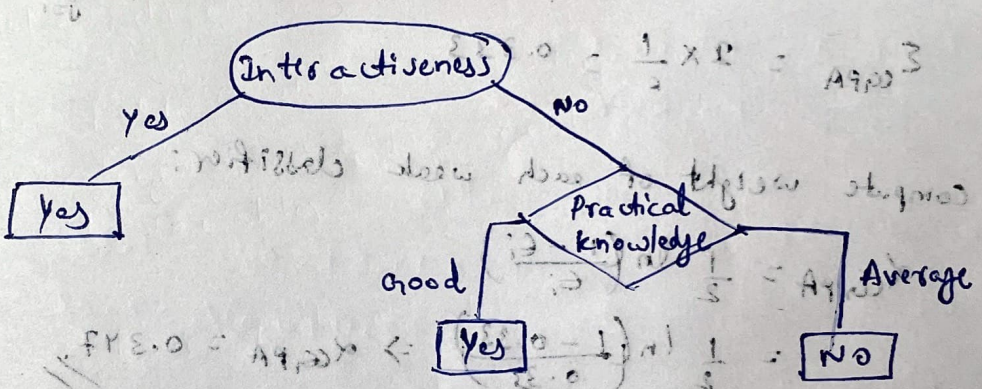
16.04.

# Random Forest - Lab 7

CGPA as root Node:




Interactiveness as root node:




i) For "9815.csv" dataset

Best - Accuracy :

Accuracy score = 0.973 (or) 97.3%.

Confusion Matrix:

$$\begin{bmatrix} 16 & 0 & 0 \\ 0 & 14 & 1 \\ 0 & 0 & 14 \end{bmatrix}$$

classified 44 out of 45 test samples correctly with 100 trees.