Date 5/5/25
Page

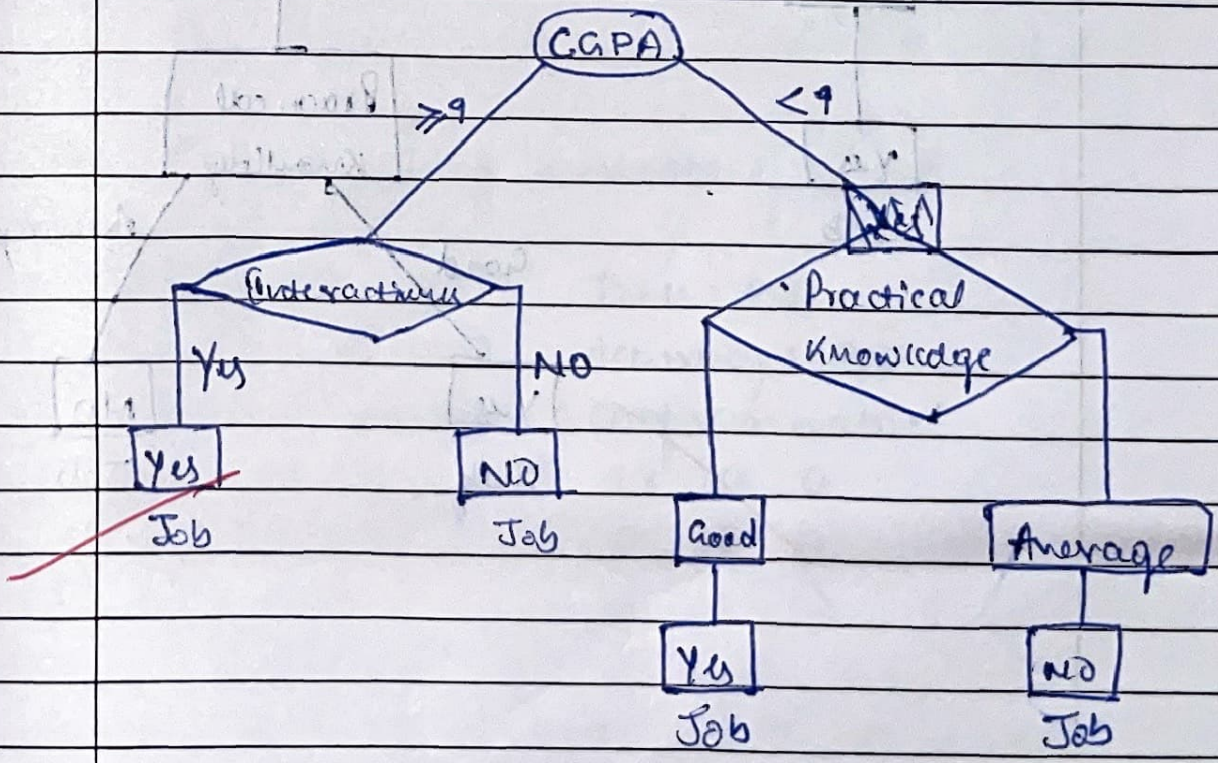
Lab-07

Implement Random forest ensemble method on a given dataset.

Sample 1:

To draw the decision tree considering CGPA as root node.

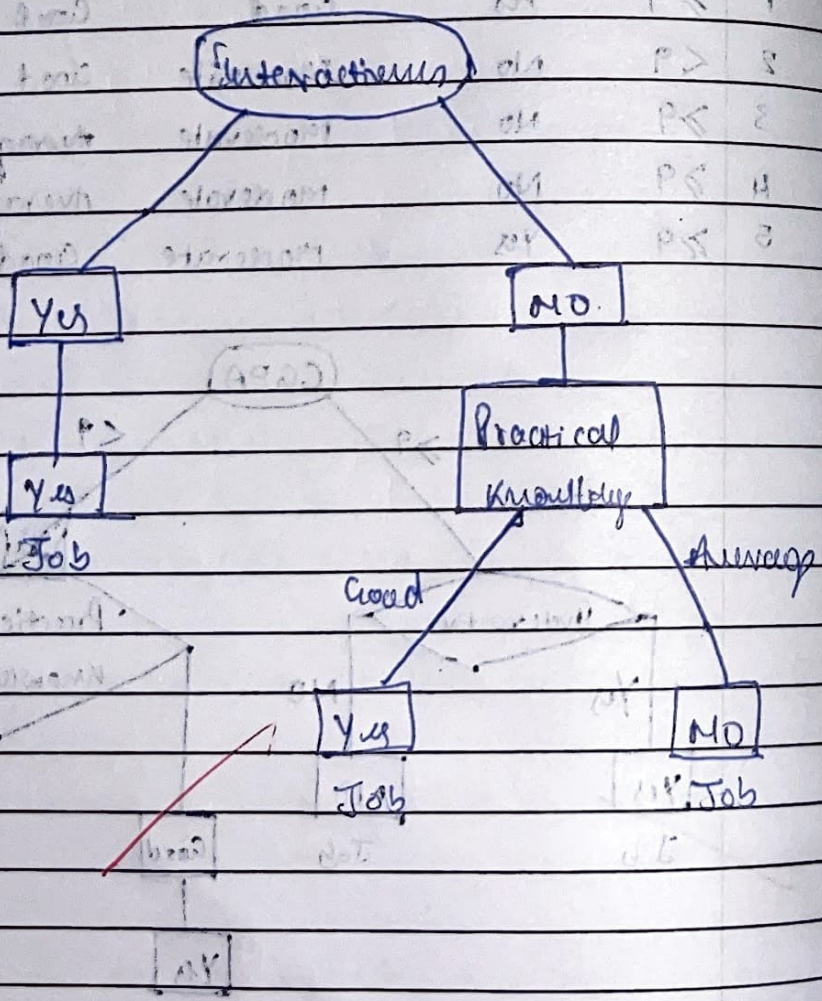
SN.NO	CGPA	Interactiveness	Communication Skills	Practical Knowledge	Job offer
1	≥9	Yes	Good	Good	Yes
2	<9	No	Moderate	Good	Yes
3	≥9	No	Moderate	Average	No
4	≥9	No	Moderate	Average	No
5	≥9	Yes	Moderate	Good	Yes



Date 5/5/25
Page

Sample 2:
Draw decision tree considering Interactiveness as root node.

SNO	CGPA	Interactiveness	Communication Skills	Practical Knowledge	Job offer
1					
2	<9	NO	Moderate	Good	Yes
3	>=9	NO	Moderate	Average	NO
3	>=9	NO	Moderate	Average	NO
5	>=9	Yes	Moderate	Average	Yes
5	>=9	Yes	Moderate	Good	Yes



Date 5/5/25
Page

→ Write python code to implement the following; consider "iris.csv" dataset.

Build a Randomforest (RF) classifier to classify IRIS flower dataset.

i) Measure prediction score using default n_estimators (10).

ii) Fine tune model by changing number of trees in your classifier & identify best score. Using, how many trees.

Sol: Output:

i) Trees = 10
Accuracy = 1.0000
Confusion matrix:

19	0	0
0	13	0
0	0	13

ii) By changing number of Trees

Trees = 20
Accuracy = 1.000
Confusion matrix:

19	0	0
0	13	0
0	0	13

Trees = 50
Accuracy = 1.000
Confusion matrix:

19	0	0
0	13	0
0	0	13

Best accuracy : 1.0000 with 10 trees.

Date 5/5/25
Page

What is the best accuracy score and confusion matrix of the classifier you observed and using how many trees?

Ans Best Accuracy score = 1.0000 (100%)
Number of trees used = 10
Confusion matrix: 19 0 0
0 13 0
0 0 13

* The random forest classifier achieved perfect classification on the "iris.csv" dataset with just 10 trees.
* Increasing the number of trees did not improve accuracy further.