Score: **73%**

No. of questions: **8**

Correct answer: **5**

Incorrect answer: **3**

Show incorrect attempt only

Question 1

1 Mark

What is the right sequence of steps while building a random forest:

1. Use feature sampling for each split in the decision tree.
2. Build a decision tree on every sample.
3. Create bootstrap Samples.
4. Aggregate all decision trees.

A 1, 2, 3, 4

B 4, 3, 2, 1

C 3, 2, 1, 4

D 4, 1, 2, 3

Correct Answer: C. 3, 2, 1, 4

To build a random forest model, we first have to create bootstrap samples, then build a decision tree on every sample, use feature sampling for each split in decision tree and finally, aggregate all the decision trees.

Question 2

2 Marks

Every tree in a Random forest is given a random set of features once.

A True

B False

Correct Answer: B. False

A feature can be considered multiple times for every tree of a random forest.

Question 3

3 marks

What do you understand by bagging?

A Bagging is a standalone predictive model by itself.

B It is the other name of bootstrapping.

C It is the process of aggregation of the results from diverse predictive models.

D It is the combined process of bootstrapping and aggregation of the predictions generated from the diverse models.

Correct Answer: D. It is the combined process of bootstrapping and aggregation of the predictions generated from the diverse models.

Bagging is the combined process of bootstrapping and aggregating predictions from diverse models.

Question 4

1 Mark

Suppose you are using a RandomForest model. Which of the following statements can be true?

i: Number of tree should be as large as possible
ii. You will have interpretability after using RandomForest

A i

B ii

C Both i and ii

D None of the above

Correct Answer: D. None of the above

After a certain number of estimators, there is no significant improvement in the model performance and the random forest may overfit. Secondly, as random forest build multiple decision trees on different subsets of data, the model is not very interpretable.

Question 5

2 Marks

How can we select best hyperparameters in tree based models?

A By measuring performance over training data

B By measuring performance over validation data

C Both A and B

D None of the above

Correct Answer: B. By measuring performance over validation data

Validation accuracy gives us how well the model will perform on validation data and hence it can be used to select the best hyperparameters.

Question 6

3 marks

Repetition of the samples are not allowed in the bootstrap sampling.

A True

B False ✓

Correct Answer: B. False

Repetition of samples is allowed in bootstrap sampling.

Question 7

1 Mark

Which of the following is a stopping criterion for random forest?

A n_jobs

B max_depth

C n_estimators

D All of the above ✗

Correct Answer: B. max_depth

max_depth is a stopping criterion which defines the growth of a random forest model.

Question 8

2 Marks

Which of the following is true about feature sampling in random forest?

A	Feature sampling takes place with replacement
B	Feature sampling takes place at tree level
C	Feature Sampling takes place at Node level
D	Feature Sampling does not consider replacement
E	Both C and D ✓

Correct Answer: E. Both C and D

Features are sampled at each node of a tree and it does not consider any replacement.