←Go Back to Advanced Machine Learning

Course Content

Weekly Quiz 1- Bagging and Random Forest

Type	:	Graded Quiz
Attempts	:	1/1
Questions	:	10
Time	:	45m
Due Date	:	Jul 29, 3:30 AM EDT
Your Marks	:	10/10

Instructions

Attempt History

Attempt #1
Jul 28, 5:11 PM

Marks: 10

Q No: 1 Correct Answer

Marks: 1/1

Which of the following statement(s) is/are true for ensemble learning?

A) Individual base learners in an ensemble model need to be dependent on each other in order to get a better prediction.

B) Ensembles are more complex than base models but they are not sensitive to slight variations in the data, hence, robust.

C) Ensembles are machine learning methods for combining predictions from multiple separate models

D) Ensemble models are only used in a classification problem.

- ○ A and B
- ◉ B and C — You Selected
- ○ C and D
- ○ A and D

Ensembles are more complex than base models but they are not sensitive to slight variations in the data, hence, robust. They are machine learning methods for combining predictions from multiple separate models. The individual base learners in an ensemble model need to be built on different data points so that they are independent of each other and result in better predictions. Ensemble models can be used for both classification and regression problems.

Q No: 2 Correct Answer

Marks: 1/1

Which of the following statement(s) is/are true?

1. Random forest is not necessarily built using decision trees. Other supervised learning algorithms can be employed to create a random forest instead of decision trees as long as the ensemble contains one particular type of task i.e. either classification or regression.
2. In Bagging, we can employ n number of models as long as the models are of the same type. For example, a bagging classifier can have 10 Decision trees, etc.

- ○ Only 1
- ◉ Only 2 — You Selected
- ○ 1 and 2
- ○ None

In bagging, we can use n number of estimators of any algorithm, however, for the random forest, we can use n number of estimators of decision trees only.

Q No: 3 Correct Answer

Marks: 1/1

Which of the following statement(s) is/are true for bagging classifier?

"In bagging, if n is the number of rows sampled and N is the total number of rows, then "

A) n can never be equal to N

B) n can be equal to N

C) n can be less than N

D) n can never be less than N

- ○ Only A
- ○ Only B
- ○ A and C
- ◉ B and C — You Selected

In bagging, if n is the number of rows sampled and N is the total number of rows, then n<=N since n is the subset of N.

Q No: 4 Correct Answer

Marks: 1/1

Select whether the following statement is True or False for Random Forest:

"A subset of rows is selected for each individual tree and a subset of columns is used to choose the best split at each level. "

- (●) True — You Selected
- ( ) False

In ensemble models, we draw a subset of data, build multiple classifiers, and then combine the output of all the models to make predictions. The basic idea of a random forest is that we draw multiple samples (rows) and for an individual tree/classifier we use a subset of features (columns).

For e.g., There are 100 rows and 10 columns.

Multiple subsets are selected using sampling with replacement. Let's say we draw 10 samples from the rows. We are to build 10 classifiers and each time we will take a subset of columns to build a tree.

This way we use a subset of both rows and columns to build an individual tree.

Q No: 5 Correct Answer

Marks: 1/1

In a given dataset, there are M columns. Out of these M, m columns are chosen each time for creating training samples for the individual trees in a random forest. What will happen if

A - m is almost equal to M

B - m is very small

- ( ) A will result in a weak tree but B will result in a very robust tree
- ( ) A will result in a robust tree but B will result in a weak tree
- (●) A will result in high correlation among individual trees resulting in lack of diversity, and B will result in very weak individual trees — You Selected
- ( ) A will have weak individual trees but B will result in high correlation among individual trees

A will result in a high correlation among the individual trees resulting in a lack of diversity and the final output won't be reliable, on the other hand, B will result in very weak individual trees.

Q No: 6 Correct Answer

Marks: 1/1

Select the statement(s) that highlight the key difference(s) between Bagging and Random Forest

A) Bagging considers all the features to decide the best split while Random Forest selects only a subset of features.

B) Bagging can have any number of estimators while Random Forest can not have any number of estimators.

C) Bagging selects only a subset of features to decide the best split while Random Forest considers all the features to decide the best split.

- (●) Only A — You Selected
- ( ) Only B
- ( ) Only C
- ( ) Both A and C

Bagging considers all the features to decide the best split while Random Forest selects only a subset of features.

Q No: 7 Correct Answer

Marks: 1/1

In a random forest, what does the out-of-bag (OOB) error rate indicate?

- ( ) Mean prediction error on each training sample x_i, using only the trees that did have x_i in their bootstrap sample
- (●) Mean prediction error on each training sample x_i, using only the trees that did NOT have x_i in their bootstrap sample — You Selected
- ( ) Mean prediction error on each testing sample x_i
- ( ) Total prediction error on each testing sample x_i

OOB error rate is the mean prediction error of each training sample x_i by using only the trees that did not have x_i in their bootstrap sample.

Q No: 8 Correct Answer

Marks: 1/1

Which of the following are true for 'class_weight' in the random forest for binary classification?

A) It is used when classes are balanced

B) It is used when classes are imbalanced

C) Random forest gives less importance to the class with more weight

D) Random forest gives more importance to the class with more weight

- ( ) A and B
- ( ) A and C
- ( ) B and C
- (●) B and D — You Selected

In random forest for binary classification, the 'class_weight' parameter is used when classes are imbalanced and the random forest gives more importance to the class with more weight while building the model.

Q No: 9 Correct Answer

Marks: 1/1

The bootstrap samples created during Bootstrap aggregation (Bagging) are created by sampling the data _____ and _____

- orderly, without replacement
- orderly, with replacement
- randomly, without replacement
- randomly, with replacement — You Selected

The bootstrap samples created during Bootstrap aggregation (Bagging) are created by sampling the data randomly and with a replacement which may lead to the repetition of some of the samples.

Q No: 10 Correct Answer

Marks: 1/1

Which of the following parameters can you vary to tune a random forest model?

1. The number of features to consider when looking for the best split

2. Number of estimators in the forest

- 1 only
- 2 only
- 1 and 2 — You Selected

Both the above two parameters can be varied to tune a random forest model

Comments:

+ Add comments

< Previous Next >