Started on	Wednesday, January 31, 2018, 11:39 PM
State	Finished
Completed on	Wednesday, January 31, 2018, 11:59 PM
Time taken	20 mins
Grade	**56.00** out of 60.00 (**93**%)

Question 1

Complete

4.00 points out of 4.00

Which of the following are hyperparameters for Imputer in scikit learn preprocessing module?

Select one or more:

- [x] a. strategy
- [ ] b. fit_transform
- [ ] c. most_frequent
- [x] d. missing_values
- [x] e. axis

Question 2

Complete

4.00 points out of 4.00

What is the estimator we use to perform regression with extra algorithm called?

Select one:

- ( ) a. ExtraTreesClassifier
- ( ) b. ExtraTreesRegression
- (•) c. ExtraTreesRegressor
- ( ) d. ExtraTrees

Question 3

Complete

4.00 points out of 4.00

Which of the parameters can be adjusted to improve performance of Adaboost estimator?

Select one or more:

- [ ] a. n_classes
- [x] b. n_estimators
- [ ] c. random_state
- [x] d. learning_rate

Question 4

Complete

4.00 points out of 4.00

Which of the following python commands correctly imports gradient–boosted classification trees?

Select one:

Select one:

a.

```
from sklearn.ensemble import gradient_boosting_classifier
```

b.

```
from sklearn import Gradient-BoostedClassificationTrees
```

c.

```
from sklearn.ensemble import GradientBoostingClassifier
```

d.

```
from scikit_learn import Gradient_Boosting_Classifier
```

e.

```
from sklearn import GradientBoostingRegressor
```

Question 5

Complete

4.00 points out of 4.00

What does the StandardScaler method from sklearn.preprocessing do?

Select one:

a. Transform data to have zero mean and unit standard deviation.

b. Transform data to have unit mean.

c. Transforms data to lie between a zero and maximum value.

d. Transforms data to lie between a minimum and maximum value, often the range [0, 1].

Question 6

Complete

4.00 points out of 4.00

In the "Classification: Adult Model" example, why do we compute the zero model before creating the RandomForestClassifier e stimator?

Select one:

- a. It helps select features for further modeling
- b. It sets a useful baseline for how well an algorithm should perform in order to be useful in practice.
- c. The model provides insights into the data
- d. It is not subject to hyperparameter tuning to improve the performance

Question 7

Complete

4.00 points out of 4.00

Which of the following are hyperparameters for AdaboostClassifier or AdaboostRegressor?

Select one or more:

- [x] a. learning_rate
- [x] b. algorithm
- [ ] c. max_samples
- [x] d. n_estimators
- [ ] e. max_features
- [x] f. base_estimator

Question 8

Complete

4.00 points out of 4.00

Which boosting algorithm supports arbitrary cost or loss functions?

Select one:

- a. Random Boost
- b. Adaboost
- c. Gradient Tree Boosting
- d. Gradient Adaboost

Question 9

Complete

4.00 points out of 4.00

Which class or function do I use to create a Pipeline object in sklearn?

Select one:

- a. sklearn.Pipeline
- b. sklearn.pipeline.Pipeline
- c. Pipeline
- d. sklearn.pipeline

Question 10

Complete

4.00 points out of 4.00

Formally what is a Bootstrap?

Select one:

- a. Bootstrap refers to any statistical process that does not rely on the generation of random samples with replacement.
- b. Bootstrap refers to any statistical process that does not rely on the generation of random samples without replacement.
- c. Bootstrap refers to any statistical process that relies on the generation of random samples with replacement.
- d. Bootstrap refers to any statistical process that relies on the generation of random samples without replacement.

Question 11

Complete

4.00 points out of 4.00

One quick test to see if objects can be used in a pipeline is to check if transformation objects have a `transform` method, and learning algorithms have a `predict` method.

Select one:

- True
- False

Question 12

Complete

4.00 points out of 4.00

The extra trees algorithm does not select the best threshold on which to split automatically. Instead, the extra trees computes a set of random thresholds and selects the best split value from this random set.

Select one:

- True
- False

Question 13

Complete

0.00 points out of 4.00

Which best defines "soft" voting in Voting Classification?

Select one:

- a. majority voting on summed classification probabilities
- b. minority voting based on input classifications
- c. majority voting based on input classifications
- d. minority voting on summed classification probabilities

Question 14

Complete

4.00 points out of 4.00

Which of the following are the hyperparameters for GradientBoostingClassifier from the scikit-learn ensemble module?

Select one or more:

- [x] a. n_estimators
- [x] b. criterion
- [x] c. max_depth
- [x] d. learning_rate
- [x] e. min_samples_leaf
- [ ] f. n_jobs
- [x] g. loss

Question 15

Complete

4.00 points out of 4.00

According to the intro2bat notebook, which of the following are advantages of using bagging algorithms?

Select one or more:

- [x] a. Less prone to overfitting.
- [ ] b. Improved performance on unbalanced data
- [x] c. Improved prediction.
- [x] d. The out-of-bag data provides a useful metric for the performance of each individual tree used in the ensemble.