

Started on	Friday, April 14, 2023, 10:16 AM
State	Finished
Completed on	Friday, April 14, 2023, 10:24 AM
Time taken	8 mins 13 secs
Points	10.00/10.00
Grade	100.00 out of 100.00

Question **1**

Correct

1.00 points out of 1.00

Scikit-learn's SimpleImputer

Select one:

- ☐ requests inputs from the user.
- ☐ produces the simplest possible interpretation for a trained random forest model.
- ☐ initializes all hyperparameters of a machine learning algorithm to their default values.
- ☒ replaces missing values with mean, median, or most frequent values. ✓

Your answer is correct.

The correct answer is: replaces missing values with mean, median, or most frequent values.

Question **2**

Correct

1.00 points out of 1.00

In one-hot encoding

Select one:

- ☒ a categorical feature is transformed into a separate binary feature for each possible value. ✓
- ☐ targets must be decoded using zero-knowledge proofs.
- ☐ a "temperature" hyperparameter determines the learning rate.
- ☐ all feature values are normalized to values between (but not including) zero and one.

Your answer is correct.

The correct answer is: a categorical feature is transformed into a separate binary feature for each possible value.

Question 3

Correct

1.00 points out of 1.00

The k-means++ algorithm

Select one:

- ☐ uses inheritance to speed up the calculation of means.
- ☐ is an object-oriented version of the k-means algorithm.
- ☒ initializes centroids in a way that is intended to distribute them well, and then runs the classic k-means algorithm. ✓
- ☐ runs the k-means algorithm twice.

Your answer is correct.

The correct answer is: initializes centroids in a way that is intended to distribute them well, and then runs the classic k-means algorithm.

Question 4

Correct

1.00 points out of 1.00

A silhouette coefficient

Select one:

- ☐ is a clustering hyperparameter that must be set by the user.
- ☐ is the the number of examples in the foreground cluster divided by the number of examples in the background cluster.
- ☐ measures the size of the projection of a cluster onto a lower-dimensional space.
- ☒ characterizes the similarity of an example to others in its cluster vs. its similarity to examples in the closest cluster of which it is not a member. ✓

Your answer is correct.

The correct answer is: characterizes the similarity of an example to others in its cluster vs. its similarity to examples in the closest cluster of which it is not a member.

Question 5

Correct

1.00 points out of 1.00

The elbow method

Select one:

- ☐ is a more efficient implementation of the knee method.
- ☐ ensures that the lines connecting each centroid to its two closest neighbors form a right angle.
- ☐ nudges centroids away from one another during cluster formation.
- ☒ is a graphical method for determining the number of clusters in a dataset. ✓

Your answer is correct.

The correct answer is: is a graphical method for determining the number of clusters in a dataset.

Question 6

Correct

1.00 points out of 1.00

When using the k-means algorithm,

Select one:

- ☐ k starts at zero, and is updated for each data point, but will not always converge to the optimal value.
- ☒ k must be specified by the user. ✓
- ☐ k is the mean value of the targets.
- ☐ the algorithm determines the optimal value for k .

Your answer is correct.

The correct answer is: k must be specified by the user.

Question 7

Correct

1.00 points out of 1.00

Agglomerative clustering

Select one:

- ☐ is guaranteed to produce k clusters, regardless of the distribution of examples.
- ☐ considers all possible combinations of all examples for each cluster.
- ☐ begins with a cluster that contains all examples and then recursively splits the cluster into sub-clusters.
- ☒ builds up clusters by starting with clusters that each contain a single example, and repeatedly merging the two closest clusters. ✓

Your answer is correct.

The correct answer is: builds up clusters by starting with clusters that each contain a single example, and repeatedly merging the two closest clusters.

Question 8

Correct

1.00 points out of 1.00

SBS (Sequential Backward Selection) is

Select one:

- ☐ an iterative algorithm for error back-propagation.
- ☐ a parent selection algorithm for genetic programming.
- ☒ a greedy algorithm for feature selection. ✓
- ☐ a backtracking algorithm for the construction of optimal decision trees.
- ☐ a regression algorithm for predicting values in numerical sequences.

Your answer is correct.

The correct answer is: a greedy algorithm for feature selection.

Question 9

Correct

1.00 points out of 1.00

K-means is

Select one:

- ☐ a classification algorithm in which the class is determined from the k nearest neighbors.
- ☐ a hierarchical clustering algorithm.
- ☐ a cross-validation algorithm that compares predictions to the means of k random subsets of the target values.
- ☒ a prototype-based clustering algorithm. ✓

Your answer is correct.

The correct answer is: a prototype-based clustering algorithm.

Question **10**

Correct

1.00 points out of 1.00

For datasets in which the most natural clusters have unusual shapes that fit together like puzzle pieces, the most appropriate clustering method is likely to be

Select one:

- ☐ the elbow method.
- ☐ k-means++.
- ☒ density-based clustering. ✓
- ☐ soft clustering.
- ☐ k-means.

Your answer is correct.

The correct answer is: density-based clustering.

◀ Slides from class

Jump to...

Project Plan Assignment ▶