You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam   0:34:50

Final Exam Instructions

1. Time allowed: **1 hour**

2. Attempts per question:

- One attempt - For True/False questions

- Two attempts - For any question other than True/False

3. Clicking the "**Final Check**" button when it appears, means your submission is **FINAL**. You will **NOT** be able to resubmit your answer for that question ever again

IMPORTANT: Do not let the time run out and expect the system to grade you automatically. You must explicitly submit your answers, otherwise they would be marked as incomplete.

QUESTION 1 (1/1 point)

Randomness is important in Random Forests because it allows us to have distinct, different trees that are based off of different data.

- ⦿ True ✔

- ◯ False

You have used 1 of 1 submissions

QUESTION 2 (1/1 point)

When building a decision tree, we want to split the nodes in a way that increases entropy and decreases information gain.

- ◯ True

- ⦿ False ✔

You have used 1 of 1 submissions

QUESTION 3 (1/1 point)

Which of the following is true?

Cookie Preferences

You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam   0:34:50

○ A high value of K in KNN creates a model with low bias and high variance

⦿ An observation must contain values for all features ✔

○ A categorical value cannot be numeric

○ None of the above

You have used 2 of 2 submissions

QUESTION 4 (1/1 point)

In terms of Bias and Variance, Variance is the inconsistency of a model due to small changes in the dataset.

⦿ True ✔

○ False

You have used 1 of 1 submissions

QUESTION 5 (1/1 point)

Which is the definition of entropy?

○ The purity of each node in a random forest.

○ Information collected that can increase the level of certainty in a particular prediction.

○ The information that is used to randomly select a subset of data.

⦿ The amount of information disorder in the data. ✔

You have used 2 of 2 submissions

QUESTION 6 (1/1 point)

Which of the following is true about hierarchical linkages?

Cookie Preferences

You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam   0:34:50

⦿ Average linkage is the average distance of each point in one cluster to every point in another cluster   ✔

○ Complete linkage is the shortest distance between a point in two clusters

○ Centroid linkage is the distance between two randomly generated centroids in two clusters

○ Single linkage is the distance between any points in two clusters

You have used 2 of 2 submissions

QUESTION 7 (1/1 point)

In terms of Bias and Variance, Variance is the inconsistency of a model due to small changes in the dataset.

⦿ True   ✔

○ False

You have used 1 of 1 submissions

QUESTION 8 (1/1 point)

Which is true about bootstrapping?

○ All data points must be used when bootstrapping is applied

⦿ The data points are randomly selected with replacement   ✔

○ The data points are randomly selected without replacement

○ It is the same as bagging

You have used 2 of 2 submissions

QUESTION 9 (1/1 point)

Machine Learning is still in early development and does not have much the current society.

Cookie Preferences

You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam

0:34:50

○ True

◉ False ✔

You have used 1 of 1 submissions

QUESTION 10 (1/1 point)

In comparison to supervised learning, unsupervised learning has:

◉ Less tests ✔

○ More models

○ A better controlled environment

○ More tests, but less models

You have used 2 of 2 submissions

QUESTION 11 (1/1 point)

Outliers are points that are classified by Density-Based Clustering that do not belong to any cluster.

◉ True ✔

○ False

You have used 1 of 1 submissions

QUESTION 12 (1/1 point)

Which of the following is false about Linear Regression?

Cookie Preferences

You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam

0:34:50

○ It does not require tuning parameters

○ It is highly interpretable

○ It is fast

◉ It has a low variability on predictive accuracy ✔

You have used 2 of 2 submissions

QUESTION 13 (1/1 point)

Machine Learning uses algorithms that can learn from data without relying on standard programming practices.

◉ True ✔

○ False

You have used 1 of 1 submissions

QUESTION 14 (1/1 point)

Which of the following are types of supervised learning?

○ Clustering

○ Regression

○ Classification

◉ Both A and B ✔

You have used 2 of 2 submissions

QUESTION 15 (1/1 point)

A Bottom Up version of hierarchical clustering is known as Divisive clustering. It is a more popular method than the agglomerative method.

Cookie Preferences

You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam   0:34:50

○ True

☑ False ✔

You have used 1 of 1 submissions

QUESTION 16 (1 point possible)

Which is NOT a specific outcome of how Dimensionality Performance improves production?

☑ Highlights the main linear technique called Principle Components Analysis. ✖

○ Creates step-wise regression.

○ Reduces number of features to be considered.

○ Highlights relevant variables only and omits irrelevant ones.

You have used 2 of 2 submissions

QUESTION 17 (1/1 point)

Feature Selection is the process of selecting the variables that will be projected from a high-order dimension to a lower one.

○ True

☑ False ✔

You have used 1 of 1 submissions

QUESTION 18 (1/1 point)

Hierarchical Clustering is one of the three main algorithms for clustering along with K-Means and Density Based Clustering.

☑ True ✔

○ False

Cookie Preferences

You are taking "Final Exam" as a timed exam. The timer on the right shows the time remaining in the exam.

End My Exam

0:34:50

You have used 1 of 1 submissions

QUESTION 19 (1/1 point)

Which one is NOT a feature of Dimensionality Reduction?

○ It can be divided into two subcategories called Feature Selection and Feature Extraction

◉ Removal of an "outsider" from the least cohesive cluster. ✔

○ Feature Selection includes Wrappers, Filters, and Embedded.

○ Feature Extraction includes Principle Components Analysis.

○ It reduces the number of variables/features in review.

You have used 2 of 2 submissions

QUESTION 20 (1 point possible)

Low bias tends to create overly generalized models, which can cause a loss of relevant relations between the features and target output. When a model has low bias, we say that it "under fits" the data.

◉ True ✖

○ False

You have used 1 of 1 submissions

Cookie Preferences