Question 1

What is the subfield of computer science that gives "computers the ability to learn without being explicitly programmed"?

1 / 1 point

Computational science

Information management

**Machine learning**

Graphics and visual computing

Correct

Correct!

**2.**

Question 2

Which of the following is a Machine Learning technique?

1 / 1 point

Clustering

Classification

Regression/Estimation

**All of the above**

Correct

Correct! All of the above are considered machine learning techniques along with association, anomaly detection, sequence mining, and recommendation systems.

**3.**

Question 3

Which of the following is true for **Multiple Linear Regression?**

1 / 1 point

Observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables.

**Multiple independent variables are used to predict a dependent variable.**

The relationship between the independent variable x and the dependent variable y is modeled as an nth degree polynomial in x.

One independent variable is used to predict a dependent variable.

Correct

Correct! This contrasts simple linear regression, which only uses one independent variable.

**4.**

Question 4

Which of the below is an example of a classification problem?

1 / 1 point

To predict the category to which a customer belongs to.

To predict whether a customer switches to another provider/brand.

To predict whether a customer responds to a particular advertising campaign or not.

**All of the above.**

Correct

Correct! All of the above can be phrased as a classification problem.

**5.**

Question 5

Which of the following is an example of Logistic Regression?

1 / 1 point

The odds of a particular individual having a heart attack based on how much they exercise and how much they weigh.

The probability of a borrower defaulting on their mortgage based upon their credit score and age.

The probability of a person purchasing life insurance based on age and income.

**All of the above.**

Correct

Correct! All of these are examples of logistic regression as they try to predict the probability of a binary response.

**6.**

Question 6

Which statement is **FALSE**about k-means clustering?

1 / 1 point

k-means divides the data into non-overlapping clusters without any cluster-internal structure.

The objective of k-means, is to form clusters in such a way that similar samples go into a cluster, and dissimilar samples fall into different clusters.

**As k-means is an iterative algorithm, it guarantees that it will always converge to the global optimum.**

Correct

Correct! K-Means is a heuristic algorithm, so it is guaranteed to converge to a result that could be a local optimum.

**7.**

Question 7

Which one best describes the clustering process for k-means clustering?

1 / 1 point

k-means divides the data into clusters with minimal overlap such that there are low chances of dissimilar samples in the same cluster.

k-means creates clusters by grouping data points with similar labels.

k-means clustering creates a tree of clusters.

**The objective of k-means is to form clusters in such a way that similar samples go into a cluster, and dissimilar samples fall into different clusters.**

Correct

Correct! K-Means seeks to create non-overlapping clusters.

**8.**

Question 8

What is a statistical model that uses Logistic function to model the conditional probability?

1 / 1 point

Stepwise regression

Ridge regression

**Logistic regression**

Linear regression

Correct

Correct! Logistic regression uses the logistic cost function to return the probability of each class.

**9.**

Question 9

In comparison to mean absolute error, mean squared error:

1 / 1 point

Weighs small and large errors equally.

Is more interpretable by taking the same unit as the response.

**Focuses more on large errors.**

­Avoids cancellation of errors.

Correct

Correct! The squared term exponentially increases larger errors as compared to smaller ones.

**10.**

Question 10

When do we use regression trees instead of decision trees?

1 / 1 point

When the response is categorical instead of continuous

When all of the independent variables are continuous

**When the response is continuous instead of categorical**

When some of the independent variables are continuous

Correct

Correct! Regression trees split the data based on features like in decision trees, but the prediction is an average across the data points in that node.