IoT Fundamental: Big Data and Analysis

Started on	Saturday, 17 September 2022, 2:42 PM
State	Finished
Completed on	Saturday, 17 September 2022, 8:27 PM
Time taken	5 hours 45 mins
Marks	30.00/30.00
Grade	**100.00** out of 100.00

Question **1**

Correct

Mark 2.00 out of 2.00

What is the most commonly used statistical method for analyzing data?

Select one:

- mean analysis
- regression analysis ✔
- mean estimation
- sample proportion

Refer to curriculum topic: 4.1.2
Regression analysis is the most commonly used statistical method for analyzing data and there are many regression models available. Regression analysis can look for correlations between one predictor variable and one target variable or for correlations between more than one predictor variable and a target variable.

The correct answer is: regression analysis

Question **2**

Correct

Mark 2.00 out of 2.00

Which type of information can distort the results of an analysis and careful consideration should be given to their removal from a data set?

Select one:

- units of measurement
- outliers ✔
- azimuth
- z-axis

Refer to curriculum topic: 4.3.2
Outliers include corrupt or distorted data that deviates far from expected values and can distort the results of an analysis. After careful consideration has been given, these data points are frequently removed from the dataset.

The correct answer is: outliers

Question 3

Correct

Mark 2.00 out of 2.00

When a number of items are grouped together, which type of machine learning algorithm can determine which items in the group predict the presence of other items?

Select one:

- classification
- clustering
- regression
- association ✔

Refer to curriculum topic: 4.1.1
Two types of unsupervised machine learning algorithms are association and clustering. Association algorithms determine which items in the group predict the presence of other items when given a number of items that are grouped together. Clustering algorithms determine which items occur most often in clusters when given many items.

The correct answer is: association

Question 4

Correct

Mark 2.00 out of 2.00

If the results of a study do not align with previous studies, what question should an evaluator ask?

Select one:

- Can the study be replicated to verify the findings? ✔
- Who paid for the research study?
- Did the study have an appropriate sample size?
- Are there any experts that disagree with the findings?

Refer to curriculum topic: 4.2.3
When following the evaluation guidelines, if a study does not produce findings that confirm or align with the results of current studies in the field, the study should be replicated to verify the reliability of the findings.

The correct answer is: Can the study be replicated to verify the findings?

Question 5

Correct

Mark 2.00 out of 2.00

In a linear regression, which variable is also known as the target or response variable?

Select one:

- dependent ✔
- independent
- predictor
- first

Refer to curriculum topic: 4.1.2
The dependent variable is also known as the target or response variable. The independent variable is also known as the predictor or explanatory variable.

The correct answer is: dependent

Question 6

Correct

Mark 2.00 out of 2.00

What type of error has occurred when a data scientist records a measurement incorrectly after viewing the correct value on the measuring device?

Select one:

- instrumental
- gross ✔
- systematic
- random

Refer to curriculum topic: 4.2.2
The different types of errors in measurement include the following:

- Instrumental - Every device is limited in how precise it can be.
- Gross - An incorrect value is accidentally recorded after the correct value is viewed.
- Random - The measuring device is correctly measuring an item and providing a varying value.
- Systematic - The measuring tool is not correctly calibrated.

The correct answer is: gross

Question 7

Correct

Mark 2.00 out of 2.00

When you follow the scientific method, which step would occur after testing the hypotheses through experimentation?

Select one:

- Ask a question about an observation.
- Communicate the results of the process.
- Analyze data from an experiment to draw a conclusion. ✔
- Perform research.

Refer to curriculum topic: 4.2.1
The scientific method is commonly used in scientific discovery and contains the following steps:
Step 1. Ask a question about an observation such as what, when, how, or why.
Step 2. Perform research.
Step 3. Form a hypothesis from this research.
Step 4. Test the hypothesis through experimentation.
Step 5. Analyze the data from the experiments to draw a conclusion.
Step 6. Communicate the results of the process.

The correct answer is: Analyze data from an experiment to draw a conclusion.

Question 8

Correct

Mark 2.00 out of 2.00

What are two types of supervised machine learning algorithms? (Choose two.)

Select one or more:

- ✓ regression ✔
- ✓ classification ✔
- mode
- association
- clustering
- mean

Refer to curriculum topic: 4.1.1
Two algorithms used with supervised machine learning are classification and regression. Supervised machine learning algorithms are the most common algorithms used in big data analytics.

The correct answers are: classification, regression

Question 9

Correct

Mark 2.00 out of 2.00

What is the goal of linear regression?

Select one:

- to compute a line the interpolates the data, and which can be expressed as a weighted average of the predictor variables and any other function ✔
- to construct a flow chart
- to provide a summary of the data
- to provide a formula that does not require validation

Refer to curriculum topic: 4.1.2
Linear regression is used for predicting a value based on gathered data. Regression analysis has a trend line in a scatter plot that shows the target variable plotted on the y-axis and the independent variable plotted on the x-axis.

The correct answer is: to compute a line the interpolates the data, and which can be expressed as a weighted average of the predictor variables and any other function

Question 10

Correct

Mark 2.00 out of 2.00

Which type of regression analysis is often used to model variables that have an exponential relationship?

Select one:

- mean
- polynomial
- median
- nonlinear ✔

Refer to curriculum topic: 4.1.2
Nonlinear regression analysis is often used to model variables that have an exponential relationship. A nonlinear regression plot may appear as a set of points arranged to a curved path.

The correct answer is: nonlinear

Question **11**

Correct

Mark 2.00 out of 2.00

A researcher has measured the reliability of a test using the parallel-forms method. What is the expected result of this measurement?

Select one:

- What is the variation of scores for different items in the same test?
- How similarly do different people score on the same test?
- How much variation exists between scores for the same person taking a test multiple times?
- How similar are the scores of two different tests that are created from the same content domain? ✔

Refer to curriculum topic: 4.2.1
The four different types of reliability that a scientist could examine are as follows:

- **Inter-rater** - How similarly do different people score on the same test?
- **Test-retest** - How much variation exists between scores for the same person taking a test multiple times?
- **Parallel-forms** - How similar are the scores of two different tests that are created from the same content domain?
- **Internal consistency** - What is the variation of scores for different items in the same test?

The correct answer is: How similar are the scores of two different tests that are created from the same content domain?

Question **12**

Correct

Mark 2.00 out of 2.00

Which type of machine learning algorithm uses data sets verified by experts as its learning basis?

Select one:

- routing
- supervised ✔
- clustering
- association

Refer to curriculum topic: 4.1.1
Supervised machine learning algorithms can learn from a dataset that has already been processed by people. Two types of algorithms used with supervised machine learning are regression algorithms and classification algorithms.

The correct answer is: supervised

Question 13

Correct

Mark 2.00 out of 2.00

Which type of reliability would a scientist measure if the scientist wants to examine the variation between exam scores for a person taking a single test multiple times?

Select one:

- parallel-forms
- inter-rater
- test-retest ✔
- internal consistency

Refer to curriculum topic: 4.2.1
The four different types of reliability that a scientist could examine include the following:

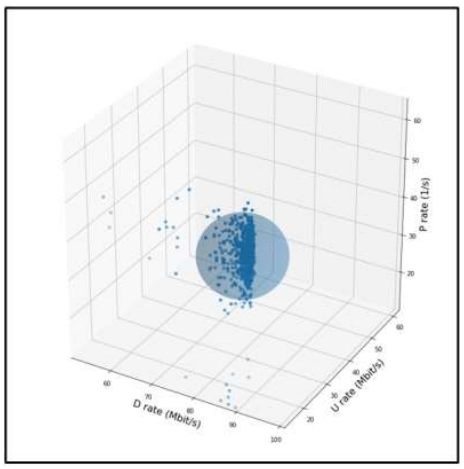
- **Inter-rater** - How similarly do different people score on the same test?
- **Test-retest** - How much variation exists between scores for the same person taking a test multiple times?
- **Parallel-forms** - How similar are the scores of two different tests that are created from the same content domain?
- **Internal consistency** - What is the variation of scores for different items in the same test?

The correct answer is: test-retest

Question 14

Correct

Mark 2.00 out of 2.00



Refer to the exhibit. What is the purpose of the blue sphere?

Select one:

- to display the mean
- to measure true error
- to categorize historical data
- to indicate data clusters ✔

Refer to curriculum topic: 4.3.2
A scientist must calculate a decision boundary to detect anomalies. Anomalous data points are points that lie beyond the decision boundary sphere.

The correct answer is: to indicate data clusters

Question **15**

Correct

Mark 2.00 out of 2.00

When is an experiment considered reliable?

Select one:

- if someone else can repeat the experiment and find different conclusions
- if someone else can modify the experiment and achieve the same conclusions
- if someone else can modify the experiment and achieve similar conclusions
- if someone else can repeat the experiment and find the same conclusion ✔

Refer to curriculum topic: 4.2.1
An experiment is considered reliable if someone else can repeat it and achieve the same results as the original scientist achieved.

The correct answer is: if someone else can repeat the experiment and find the same conclusion

◄ Chapter 4 Terms and Concepts Practice

Jump to...

Read Chapter 5: Storytelling with Data ►

NetAcad, a Cisco Corporate Social Responsibility program, is an IT skills and career building program available to learning institutions and individuals worldwide.

Terms and Conditions

Privacy Statement

Cookie Policy

Data Protection

Trademarks

Data Protection

Accessibility