Pre-Test for SIC Artificial Intelligence

Date:

Name:

Q1. Which of the following is NOT true about classification?

1. Classification is a data mining technique that predicts a group to which a data value belongs.
2. It is not required to understand the definitions of each group.
3. Classification that categorizes customer records into particular groups is similar to clustering.
4. Classification is used to categorize (0,1) or like (0,1,2,3,4).

Q2. Find the below function’s derivative using Python.



Q3. Below are the exploration techniques for univariate data. Which is NOT a continuous data exploration technique?

1. Histogram
2. Box plot
3. Kernel density curve
4. Pie chart

Q4. Suppose that the result of a dice roll is an even number. What is the probability of its value being a divisor of 4?

1. 2/3
2. 1/3
3. 1/4
4. 3/5

Q5. Which of the following is NOT used to select the classification variable in a decision tree?

1. Generalized variance
2. Gini index
3. Entropy index
4. P-value of chi-square statistics

Q6. Which of the following is NOT true about clustering?

1. Clustering is an analysis technique that divides the data into groups based on similarity or dissimilarity.
2. The most common clustering technique is k-means.
3. K-medoids is one of the hierarchical clustering algorithms.
4. In hierarchical clustering, the number of groups is determined only after clustering.

Q7. Which of the following is a process that extracts meaningful data from large-scale documents? It is also known as text analytics, knowledge discovery from text data, or document mining.

1. Text mining
2. Web mining
3. Opinion mining
4. Facebook mining

Q8. Which of the following is NOT true about the pros and cons of neural networks?

1. It is difficult to interpret the results.
2. It is possible to understand the complicated relationship between variables.
3. Neural networks show excellent performance even with noisy input data.
4. Neural networks require a small amount of training data to train.

Q9. The MNIST dataset consists of handwritten digits from 0 to 9. If you want to create a neural network that can categorize the dataset, which will be the activation function in the final fully-connected (= dense) layer?

1. Sigmoid
2. ReLU
3. Maxout
4. Softmax

Q10. Which of the following types adequately understands data analysis techniques and structures but fails to correctly recognize the analysis target within the organization?

1. Optimization
2. Solution
3. Insight
4. Discovery