Started on	Monday, November 1, 2021, 9:40 AM
State	Finished
Completed on	Monday, November 1, 2021, 9:50 AM
Time taken	9 mins 31 secs
Grade	**10.00** out of 10.00 (**100**%)

Question 1

Correct

1.00 points out of 1.00

In numpy, what is a correct syntax to print the numbers [2, 3, 4] from the array below:

arr = np.array([1,2,3,4,5,6,7])

- ⦿ a. arr[1:4] ✔
- ○ b. arr[2:4]
- ○ c. arr[2:5]
- ○ d. arr[1:3]

The correct answer is:
arr[1:4]

Question 2

Correct

1.00 points out of 1.00

What is the difference between classification and regression algorithms?

- ○ a. Classification is usually done for supervised learning, whereas regression is done for unsupervised learning.
- ⦿ b. Classification predicts a categorical variable, whereas regression predicts a number variable. ✔
- ○ c. Classification requires all data to be in categorical form, whereas regression does not have such a requirement.
- ○ d. There is not a big difference between them as they use similar data and predicts similar variables.

The correct answer is:
Classification predicts a categorical variable, whereas regression predicts a number variable.

Question 3

Correct

1.00 points out of 1.00

What kind of problem is predicting total precipitation (raining) for Ankara in 2022 based on historical data?

- ⦿ a. Regression ✔
- ○ b. Classification
- ○ c. Similarity
- ○ d. Clustering

The correct answer is:
Regression

Question 4

Correct

1.00 points out of 1.00

"40 students have attempted this quiz." is an example of what?

- ⦿ a. Information ✔
- ○ b. Data
- ○ c. Wisdom
- ○ d. Knowledge

The correct answer is:
Information

Question 5

Correct

1.00 points out of 1.00

"Based on our data, we are expecting 4% churn rate next year. So, we should do something to prevent this."

This situation is related with what kind of analytics?

- ◉ a. Prescriptive ✔
- ○ b. Predictive
- ○ c. Descriptive
- ○ d. Diagnostic

The correct answer is:
Prescriptive

Question 6

Correct

1.00 points out of 1.00

Which one is true about open source licenses?

- ○ a. Weak copyleft license is relaxed when compared to strong because now if covered code is modified, we can distribute it under any license we want.
- ○ b. If you used existing packages which are released under GPL (a strong copyleft license), your software does not have to have GPL license.
- ◉ c. BSD license can be preferred if we are not interested in whether users can modify it for their own use. ✔
- ○ d. With permissive licences, modified covered codes should be distributed under same license with the original software.

The correct answer is:
BSD license can be preferred if we are not interested in whether users can modify it for their own use.

Question 7

Correct

1.00 points out of 1.00

We are dealing with trillions of data, so we are prone to have erroneous data. This sentence reflects which side of the big data?

- ○ a. Variability
- ○ b. Velocity
- ○ c. Variety
- ◉ d. Veracity ✔

The correct answer is:
Veracity

Question 8

Correct

1.00 points out of 1.00

Which of these problems would you address with an unsupervised learning algorithm?

- ○ a. You have an Excel sheet with house prices based on size. You want to estimate a price for your house.
- ○ b. You have historical data on Real Madrid vs Barcelona. You want to predict who will win the next match.
- ◉ c. You have many papers written on Turkish History. You want to find a way to automatically group them into some groups based on their similarity. ✔
- ○ d. You have all of your family member's and your genetic data. You would like to know whether you are likely to develop any genetic disorders in your adulthood.

The correct answer is:
You have many papers written on Turkish History. You want to find a way to automatically group them into some groups based on their similarity.

Question 9

Correct

1.00 points out of 1.00

What kind of problem is deciding whether a loan applicant can repay his/her debt?

- ◉ a. Classification ✔
- ○ b. Data Reduction
- ○ c. Regression
- ○ d. Clustering

The correct answer is:
Classification

Which one is false about data science frameworks?

- ○ a. Although they are beneficial for companies, many organizations still use ad hoc processes.

- ○ b. OSEMN, KDD, SEMMA and CRISP-DM are example data science frameworks that companies can use.

- ○ c. They can be used along with agile methodologies.

- ● d. As SEMMA does not have a business understanding step like CRISP-DM, SEMMA should not be preferred. ✔

The correct answer is:
As SEMMA does not have a business understanding step like CRISP-DM, SEMMA should not be preferred.