AIML Assessment BC-101

Points: 20/20

✓ **Correct** 1/1 Points

1. You have a dataset of exam scores for multiple subjects. You create a box plot for each subject's scores. In which part of the box plot would you find the first quartile (Q1)?

- ○ Upper Whisker
- ◉ Box (IQR)
- ○ Lower Whisker
- ○ Median

✓ **Correct** 1/1 Points

2. You are exploring a dataset and want to visualize the distribution of data points to identify potential outliers. Which graphical representation is commonly used for this purpose?

- ○ Line Plot
- ◉ Box Plot

- Pie Chart
- Bar Chart

✓ **Correct** 1/1 Points

3. What will be the output of following code?

```
import numpy as np
arr = np.array([100, 200, 300, 400, 500, 600, 700])
print(arr[::-2])
```

- [100 300 500 700]
- [200 400 600]
- [700 500 300 100]
- [600 400 200]

✓ **Correct** 1/1 Points

4. In regression analysis, the variable that is being predicted is:

- the source variable
- the target variable
- usually denoted by x
- usually denoted by r

✓ **Correct** 1/1 Points

5. Which of the following can be used to create various inputs using pandas DataFrame?

- ○ Lists, Dict
- ○ Series
- ○ Numpy Arrays
- ◉ All of the above mentioned

✓ **Correct** 1/1 Points

6. You have a CSV file named "employee_data.csv" containing the details of employees working in an organization. Which Pandas function would you use to read this CSV file into a DataFrame?

- ○ pd.load_csv("employee_data.csv")
- ○ pd.read_excel("employee_data.csv")
- ◉ pd.read_csv("employee_data.csv")
- ○ pd.load("employee_data.csv")

✓ **Correct** 1/1 Points

7. What evaluation metrics are used for Regression ML Algorithm?

- ○ Precision, Recall and F1-Score
- ◉ MSE, RMSE and MAE
- ○ Any one of the Above
- ○ None of the Above

✓ **Correct** 1/1 Points

8. Which of the following module is used to create n-dimensional arrays?

- ○ pandas
- ○ plotly
- ◉ numpy
- ○ seaborn

✓ **Correct** 1/1 Points

9. You are working on a project to predict the salary of an employee who is joining the organization based on certain parameters. Which type of regression algorithm would be most suitable for this task?

- ○ Logistic Regression
- ○ Decision Tree Regression
- ◉ Linear Regression
- ○ None of the Above

✓ **Correct** 1/1 Points

10. What will be the output of the below code?

```
import pandas as pd
srs_1 = pd.Series([10, 20, 30, 40, 50])
srs_2 = srs_1 * 3
print(srs_2.values)
```

- ○ [10, 20, 30, 40, 50, 10, 20, 30, 40, 50, 10, 20, 30, 40, 50]

- [ ] [10, 10, 10, 20, 20, 20, 30, 30, 30, 40, 40, 40, 50, 50, 50]
- [x] [30, 60, 90, 120, 150]
- [ ] Result an Error

✓ **Correct** 1/1 Points

11. Which of the following statement will display 3rd row to 7th row and first 3 columns?

- [x] df.iloc[2:7, :3]
- [ ] df.iloc[3:8, 1:3]
- [ ] df.iloc[3:7:2, :3]
- [ ] df.iloc[2:7:2, 1:3]

✓ **Correct** 1/1 Points

12. Which of the following is used to find the distance between two data points?

- [x] euclidean distance
- [ ] precision and recall
- [ ] mean squared distance
- [ ] All of the above

✓ **Correct** 1/1 Points

13. You are building a predictive model for housing prices. During data preprocessing, you observe some outliers in the target variable (prices). Which of the following

actions is generally recommended to handle these outliers?

- Replace outliers with the median of the target variable
- Replace outliers with the mean of the target variable
- No action is required; outliers do not affect regression models
- Exclude the outliers from the dataset

✓ **Correct** 1/1 Points

14. A residual (error) is defined as

- The difference between the actual Y values and the mean of Y
- The predicted value of Y for the average X value
- The square root of the slope
- The difference between the actual Y values and the predicted Y values

✓ **Correct** 1/1 Points

15. You are working on a dataset that has a continuous target variable. After analyzing the data, you find a strong linear relationship between the input features and the target variable. Which regression algorithm would be most appropriate for this scenario?

- K-Nearest Neighbour
- Logistic Regression
- Linear Regression
- K-Means Clustering

✓ **Correct** 1/1 Points

16. In a dataset, you want to use a standardized score to identify outliers. Which statistical measure involves calculating how many standard deviations a data point is from the mean?

- ○ Variance
- ○ IQR
- ○ Skewness
- ◉ Z-Score

✓ **Correct** 1/1 Points

17. In the regression equation, y = mx + c, c is the

- ○ slope of the line
- ◉ y-intercept
- ○ independent variable
- ○ coefficient of determination

✓ **Correct** 1/1 Points

18. You have a DataFrame named df with missing values in some rows. Which of the following technique will be used to fill the null records?

- ○ Replace with Mean, Meadian or Mode
- ○ Forward Filling
- ○ Backward Filling

- [x] All of the Above

✓ **Correct** 1/1 Points

19. After applying k-means clustering, you want to assign each data point to its respective cluster. Which step of the k-means algorithm is responsible for this assignment?

- [ ] Initialization of Centroids
- [ ] Updating Centroids
- [ ] Calculating distances
- [x] Assigning Data Points to Clusters

✓ **Correct** 1/1 Points

20. For two runs of K-Means Clustering, is it expected to get the same clustering results?

- [ ] Yes
- [x] No