

National Forensic Sciences University
School of Cyber Security and Digital Forensics

Course Name: M.Tech AI_DS (Batch: 2024-25)

Semester - 1 Exam: TA - I (SEP - 2024)

Subject Code: CTMTAIDS SI P4

Time: 2:30 PM to 3:15 PM

Subject Name: Fundamentals of Data Science and Machine Learning

Date: 10-sep-2024

Q1. Multiple Choice Question. (Attempt All)

[05 Marks]

1. Which statistical measure is used to quantify the spread of data around the mean?
 - A. Mode
 - B. Median
 - ☒ C. Standard Deviation
 - D. Range
2. In Pandas, what is the primary function of the dropna() method?
 - A. Fills missing values with a specified value.
 - B. Removes rows or columns containing missing values.
 - C. Converts missing values to a specific data type.
 - D. Replaces missing values with interpolated values.
3. Given a Pandas DataFrame df with a missing value in column 'A' at index 2, which interpolation method would be most suitable if you believe the missing value is likely to be close to the average of the values at indices 1 and 3?
 - A. linear
 - B. time
 - C. nearest
 - D. quadratic
4. In a confusion matrix for a binary classification problem, what does the diagonal represent?
 - A. The number of correct predictions.
 - B. The number of incorrect predictions.
 - C. The total number of instances.
 - D. The number of instances in the positive class.
5. A distribution with a long tail to the right is said to be:
 - ☒ A. Skewed left
 - B. Skewed right
 - C. Symmetric
 - D. Bimodal

Q2. Answer the following questions in short. (Attempt any 4)

[8 Marks]

1. How can you detect outliers using a box plot? What is lower and upper extreme formula using IQR?
2. Why preprocessing of data performed?
3. Write any 2 Data Reduction method.
4. What is use of kurtosis? explain its types with Diagram.
5. Provide step included in Data Science Process
6. Difference between Normalization and Standardization.

Q3. Answer the following questions in Detail. (Attempt any 3)

[12 Marks]

1. Write a python function to perform following transformation (final value can be float)

1 feet = 30.48 cm

1 inch = 2.54 cm

Name	Height
Ram	5' 11"
Shyam	6' 1"
Sita	5' 5"

=>

Name	Height
Ram	180
Shyam	185
Sita	165

2. Write a pandas code for creation of Dataframe as per following values.

Name	Height	Age	Weight	Salary
Ram	180	32	82	60K
Shyam	185	37	78	70K
Sita	165	30	58	50K

3. Provide 3 ways to handled null value with supportive python code.
4. What is Confusion Matrix, recall rate?
A model outputs 3 TP, 4 TN, 2 FP, and 1 FN.
Calculate the Accuracy, precision, FPR and Recall in Percentage.

$$\begin{array}{r} 5 \times 30 + 2 \times 11 \\ 150 + 22 \\ 172 \\ \hline 172 \end{array}$$