PRN No.

(AY:2024-25) May 2025 (ENDSEM) EXAM

SY (SEMESTER - II)

COURSE NAME: Data Centric AI   **Branch:** ALL   **COURSE CODE:** AIOEUA22239B

SY (Pattern 2023)

Time: [1Hr 30 Min]

[**Max. Marks: 40**]

(*) Instructions to candidates:
 1) **Figures to the right indicate full marks. Use of scientific calculator is allowed**
 2) **Use suitable data wherever required**
 3) **All questions are compulsory. Solve any two sub question each from Questions 1 , 2 ,3 and 4**

Q. No.	Question Description	Max. Marks	CO mapped	BT Level
Q.1	a) Why is it important to collect data from multiple sources, and what challenges might arise during this process?	[5]	1	Understand
	b) Explain the key steps involved in the process of acquiring data for analysis.	[5]	1	Understand
	c) What are the various techniques used to gather Primary data? Provide suitable examples for each method.	[5]	1	Understand
Q2	a) Given dataset: {28, 35, 1, 2, 4, 20, 20, 39, 37, 102, 101, 36} Find the outliers from the given dataset (Use IQR).	[5]	2	Apply
	b) What is data Annotation? Explain different types of Data Annotation methods.	[5]	2	Understand
	c) Describe common preprocessing steps such as handling missing values, removing duplicates, and standardizing data formats. Why are these steps necessary before training a model?	[5]	2	Understand
Q3	a) What are some common examples of ineffective visualizations and how can they be improved?	[5]	3	Apply
	b) A human resource manager wants to analyze the organizational structure of a company and visualize reporting hierarchies among employees. Which visualization category is most suitable? What type of chart or graph would best represent this data? Justify your answer.	[5]	3	Apply

	c) State and explain GESTALT principles for visualization.	[5]	3	Understand
Q4	a) Explain the difference between static visualizations and interactive dashboards. Why are interactive dashboards useful in data analysis?	[5]	4	Understand
	b) Give examples of common interactive features used in modern visualization tools and explain their purpose.	[5]	4	Understand
	c) Describe how AI techniques like K-Means clustering can help in data visualization. Why is clustering useful for understanding patterns in data?	[5]	4	Understand