R22

MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY
(Autonomous Institution – UGC, Govt. of India)
III B.Tech I Semester Supplementary Examinations, June 2025
Introduction to Data Science
(CSE-AIML & B.Tech-AIML)

Roll No										

Time: 3 hours   **Max. Marks: 60**

Note: This question paper contains two parts A and B
Part A is compulsory which carries 10 marks and Answer all questions.
Part B Consists of 5 SECTIONS (One SECTION for each UNIT). Answer **FIVE** Questions, Choosing ONE Question from each SECTION and each Question carries 10 marks.

***

		PART-A (10 Marks) (Write all answers of this part at one place)	**BCLL**	**CO(s)**	**Marks**
1	A	Define Data science.	L1	CO-I	[1M]
	B	What are the most popular Python toolboxes?	L2	CO-I	[1M]
	C	How to calculate Sample Variance?	L2	CO-II	[1M]
	D	What is Hypothesis Testing?	L1	CO-II	[1M]
	E	Define confusion matrix	L1	CO-III	[1M]
	F	Define kernel function	L1	CO-III	[1M]
	G	How to calculate MSE	L2	CO-IV	[1M]
	H	Draw the Logistic function for different lambda values.	L2	CO-IV	[1M]
	I	Which is the densest community (in terms of connections)?	L1	CO-V	[1M]
	J	How to represent degree of a node?	L2	CO-V	[1M]

PART-B (50 Marks)
SECTION-I

2	A	Data science allows us to adopt different strategies to explore the world using data – justify the statement	L2	CO-I	[5M]
	B	Discuss briefly about Web Integrated Development Environment (WIDE)	L1	CO-I	[5M]

OR

3	A	Explain the various fundamental Python Libraries for Data Scientists.	L2	CO-I	[5M]
	B	What are most common aggregation functions? Explain	L2	CO-I	[5M]

SECTION-II

4	A	Write short notes on Descriptive statistics	L1	CO-II	[5M]
	B	What is significance of Quantiles and Percentiles plotting ? explain	L2	CO-II	[5M]

OR

5	A	With the help of Statistical data we can estimates the approximate the values of unknown parameters of the	L3	CO-II	[5M]

dataset -justify the statement .

	B	Write short notes on Confidence Intervals.	L1	CO-II	[5M]

SECTION-III

6	A	Write short notes on L1,L2 weight regularization.	L1	CO-III	[5M]
	B	Random Forest is an ensemble technique -justify the it	L3	CO-III	[5M]

OR

7	A	Explain why validation is considered as best hyperparameters	L2	CO-III	[5M]
	B	Write short notes on Support Vector Machines	L1	CO-III	[5M]

SECTION-IV

8	A	What is the role of regression analysis in data science ? explain briefly	L2	CO-IV	[5M]
	B	Explain how Logistic Regression can be used as classifier?	L2	CO-IV	[5M]

OR

9	A	Explain Sparse Model with an example .	L2	CO-IV	[5M]
	B	Illustrate on the Metrics to Measure Clustering Quality.	L2	CO-IV	[5M]

SECTION-V

10	A	Explain in detail about Social network analysis Processes using social data structure.	L2	CO-V	[5M]
	B	Write short notes Ego-Networks	L1	CO-V	[5M]

OR

11	A	Write short notes on calculating the Centrality	L1	CO-V	[5M]
	B	Explain PageRank Algorithm.	L1	CO-V	[5M]

***