Code No: R22A6707

MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY
(Autonomous Institution – UGC, Govt. of India)
III B.Tech I Semester Regular Examinations, November 2024
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
***

		PART-A (10 Marks) **(Write all answers of this part at one place)**	**BCLL**	**CO(s)**	**Marks**
1	A	What is novelty of data science?	L1	CO-I	[1M]
	B	What is the significance of Sample Variance?	L2	CO-I	[1M]
	C	List the most common steps for data preparation	L1	CO-II	[1M]
	D	What is data preparation?	L1	CO-II	[1M]
	E	Why L2 weight regularization is used for?	L2	CO-III	[1M]
	F	What is soft margin in SVM?	L2	CO-III	[1M]
	G	What is the outcome of Silhouette Score?	L2	CO-IV	[1M]
	H	List the disadvantages of overfitting.	L2	CO-IV	[1M]
	I	What is the significance of PageRank?	L2	CO-V	[1M]
	J	What is Community toolbox?	L1	CO-V	[1M]

PART-B (50 Marks)

SECTION-I

			BCLL	**CO(s)**	**Marks**
2	A	Explain in detail about Jupyter Notebook Environment	L1	CO-I	[5M]
	B	Explain the various Fundamental Python Libraries for Data Scientists	L2	CO-I	[5M]

OR

3	A	List of most common aggregation functions	L1	CO-I	[5M]
	B	Explain Integrated Development Environments (IDE) using Python.	L3	CO-I	[5M]

SECTION-II

4	A	Write short notes on Exploratory Data Analysis	L1	CO-II	[5M]
	B	What is confidence interval? What is the impact of it?	L3	CO-II	[5M]

OR

5	A	Explain in detail about the process involved in data preparation	L1	CO-II	[5M]
	B	Write short notes on Descriptive statistics	L2	CO-II	[5M]

SECTION-III

6	A	Write a brief Note on Deriving Hard Margin Support Vector Machines.	**L1**	**CO-III**	**[5M]**
	B	Machine learning can be considered a subfield of artificial intelligence (AI) -justify the statement.	**L3**	**CO-III**	**[5M]**

<div align="center">OR</div>

7	A	Write a short notes on Learning Curves	**L1**	**CO-III**	**[5M]**
	B	Explain Training,Validation and Test with an example.	**L1**	**CO-III**	**[5M]**

SECTION-IV

8	A	How the Polynomial Regression is used in regression analysis ?Explain.	**L2**	**CO-IV**	**[5M]**
	B	Explain Hierarchical Clustering with real time examples.	**L2**	**CO-IV**	**[5M]**

<div align="center">OR</div>

9	A	Write short notes on Taxonomies of Clustering Techniques	**L2**	**CO-IV**	**[5M]**
	B	Compare different Hard Partition Clustering Algorithms.	**L2**	**CO-IV**	**[5M]**

SECTION-V

10	A	Write short notes on Ego-Networks	**L1**	**CO-V**	**[5M]**
	B	Explain Social Network Analysis with respect to data science?		**CO-V**	**[5M]**

<div align="center">OR   **L2**</div>

11	A	What is Community Detection? Explain with an example		**CO-V**	**[5M]**
	B	Explain network analysis in detail with Facebook Dataset	**L3**	**CO-V**	**[5M]**

<div align="center">*****</div>