Total No. of Printed Pages: 2

PRN No.	

PAPER CODE	U124-365

May 2024 (ENDSEM) EXAM

F.Y.B. TECH. (SEMESTER - II)

COURSE NAME: FUNDAMENTALS OF DATA SCIENCE **BRANCH: CSE(DATA SCIENCE)** **COURSE CODE: CD12235**

(PATTERN 2023)

Time: [1Hr. 30 Min] **[Max. Marks: 40]**

Instructions to candidates:

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

Q. No.	Question Description	Max. Marks	CO mapped	BT Level
Q.1	a) Give examples of real-world scenarios where Data Science techniques have been effectively employed to solve problems.	[5]	1	Understand
	b)If an organization wants to optimize its marketing strategies to target specific customer segments more effectively, how can the application of Data Science principles aid in achieving this goal?	[5]	1	Understand
Q2	a) Differentiate various data collection techniques such as surveys, interviews, and web scraping.	[5]	2	Understand
	b)Discuss sentiment analysis contribute to the understanding of public opinion and consumer behavior?	[5]	2	Understand
Q.3	a)Describe the various types of statistical models commonly used in predictive analytics. How do regression models differ from classification models in terms of their applications and underlying principles?	[5]	3	Understand
	b)Explain the role of exploratory data analysis in identifying outliers and anomalies. How can the detection of outliers impact the reliability of statistical models?	[5]	3	Understand
	c)Discuss the concept of feature engineering in the context of building statistical models. How does feature	[5]	3	Understand

	selection and transformation contribute to model accuracy and interpretability?			
	d)Define Exploratory Data Analysis (EDA) and its significance in data science. How does EDA help in understanding the underlying structure of data?	[5]	3	Understand
Q.4	a)Articulate the three main types of machine learning, supervised learning, unsupervised learning, and reinforcement learning. Provide examples of tasks suitable for each type and discuss the differences in their learning objectives.	[5]	3	Apply
	b)Discuss the role of labeled data in supervised learning. How does the availability and quality of labeled data impact the performance of supervised learning algorithms?	[5]	3	Apply
	c)Demonstrate the steps involved in using machine learning algorithms to automate tasks. From data preprocessing to model evaluation, analyze the key considerations at each stage of the machine learning pipeline using practical examples.	[5]	3	Apply
	d)Evaluate how techniques like model selection and hyperparameter tuning contribute to optimizing model performance in machine learning. Discuss how cross-validation can be utilized as a technique to enhance model performance, providing real-world examples to support your analysis.	[5]	3	Apply