Reg. No. : | | | | | | | | | | | | |

Question Paper Code: REG71

B.E. DEGREE EXAMINATION, NOV 2025

Professional Elective

Artificial Intelligence & Data Science

R21ADVG71 – FUNDAMENTALS OF DATA SCIENCE

(Regulations R2021)

Duration: Three hours | Maximum: 100 Marks

Answer ALL Questions

PART - A (5 x 20 = 100 Marks)

1. (a) Define data reshaping in R and describe the operators available for manipulating data, including syntax and examples. CO1-U (20)

Or

(b) Describe the decision-making and loop constructs available in R, including their syntax and a sample program. CO1-U (20)

2. (a) Apply K-Means algorithm for the following data set and find the clusters for the data set where k=2. CO2-App (20)

Let's say we want to cluster a group of 20 individuals between the ages of 20 and 40. We have collected data on their ages, which are as follows:

25, 22, 28, 36, 32, 23, 27, 30, 31, 29, 33, 24, 26, 34, 37, 38, 21, 35, 39, 40

Or

(b) Apply naïve bayes theorem for the following: CO2-App (20)

Problem: If the weather is sunny & Rainy & Overcast, then the Player should play or not?

No.	Outlook	Play
1.	Rainy	Yes
2.	Sunny	Yes
3.	Overcast	Yes
4.	Overcast	Yes
5.	Sunny	No
6.	Rainy	Yes
7.	Sunny	Yes
8.	Overcast	Yes
9.	Rainy	No
10.	Sunny	No

11.	Sunny	Yes
12.	Rainy	No
13.	Overcast	Yes
14.	Overcast	Yes

3.	(a)	Explain in detail any three Big Data applications, supported by clear diagrams.	CO1-U	(20)
		Or		
	(b)	Define Hadoop and elaborate on application development within the Hadoop ecosystem.	CO1- U	(20)
4.	(a)	Explain clustering methods comprehensively, describing different approaches and techniques.	CO1- U	(20)
		Or		
	(b)	Provide a detailed explanation of decision trees, including their structure and use.	CO1- U	(20)
5.	(a)	Using your knowledge of Big Data applications and Hadoop components, propose a solution for monitoring tourist trends and preferences in the Travel and Tourism industry.	CO2- App	(20)
		Or		
	(b)	Given a scenario of a healthcare organization aiming to manage patient data, explain how Hadoop components can be used to handle and analyze Big Data efficiently.	CO2- App	(20)