R22 Regulation *Subject code: 4P5EC*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

(Autonomous, Accredited by NAAC with 'A+' Grade)

B.Tech V Semester Regular/Supplementary Examinations, November 2025

FUNDAMENTALS OF DATA SCIENCE

(CSE)

Maximum Marks: 60 Date:14.11.2025 Duration: 3 hours

Note: **1.This question paper contains two parts A and B.**
2. Part A is compulsory which carries 10 marks. Answer all questions in Part A.
3. Part B consists of 5 Units. Answer any one full question from each unit.
4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks (10X1M=10 Marks)		Marks	CO	Bloom Tx
1.a)	What is meant by the term "Data Science hype"?	1M	1	L1
b)	List the steps in the Data Science Process.	1M	1	L1
c)	Write the equation of a simple linear regression model.	1M	2	L1
d)	State Bayes theorem.	1M	2	L1
e)	Mention one application of feature generation in business analytics.	1M	3	L1
f)	Define about Entropy.	1M	3	L1
g)	List the key components of a recommendation algorithm.	1M	4	L1
h)	How does PCA reduce dimensionality?	1M	4	L1
i)	Give two examples of real-world social networks	1M	5	L1
j)	Write the purpose of visualization of a complex dataset.	1M	5	L1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	Bloom Tx
2	a) Discuss the current landscape of Data Science perspectives. b) Identify the essential skill sets required for a Data Scientist.	5M 5M	1	L2 L3
	OR			
3	a) Describe probability distributions and their significance in statistical modelling. b) Outline the basic tools of EDA such as plots, graphs, and summary statistics.	5M 5M	1	L2 L2
4	a) Explain the concept of Linear Regression with an example. b) Discuss how the choice of 'k' and distance metric affects the accuracy of the kNN algorithm.	5M 5M	2	L3 L3
	OR			
5	a) What is Naive Bayes classification algorithm? Explain different steps in Naive Bayes classification algorithm with an example b) Define Data Wrangling. Explain the key steps involved in data wrangling with suitable examples.	5M 5M	2	L2 L2

6	a) Describe how feature extraction helps in understanding and predicting user (customer) retention. b) How can brainstorming help in identifying useful features? Give suitable examples.	5M 5M	3	L4 L3
	OR			
7	a) Compare and contrast filter and wrapper methods of feature selection. b) Explain how Random Forests perform feature selection and the advantages they offer over single Decision Trees.	5M 5M	3	L4 L2
8	a) Describe how a user-facing data product like Netflix or Amazon recommendation engine is designed. b) Identify the limitations and assumptions of Principal Component Analysis	5M 5M	4	L4 L3
	OR			
9	a) Describe the need for dimensionality reduction in recommendation systems. b) Discuss how SVD can be used to improve recommendation systems with examples.	5M 5M	4	L4 L3
10	a) Make use of graph-based representation in understanding social interactions. b) Apply the concept of partitioning of graphs to explain the applications in social network analysis.	5M 5M	5	L3 L3
	OR			
11	a) Describe various tools used for creating data visualizations. b) Model some inspiring industry projects that effectively use data visualization	5M 5M	5	L4 L3