Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – December 2025

Date: 08-12- 2025 **Time:** 01:00pm – 04:00pm

School: SOCSE	**Program**: B.Tech (CSD)	
Course Code : CSD1712	**Course Name:** Statistical Foundations of Data Science	
Semester: V	**Max Marks**:100	**Weightage**: 50%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**24**	**26**	**26**	**24**	

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **10Q x 2M=20M**

1.	Define descriptive statistics.	**2 Marks**	**L1**	**CO1**
2.	What does a box plot show?	**2 Marks**	**L2**	**CO1**
3.	Explain the main assumption of linear regression?	**2 Marks**	**L2**	**CO2**
4.	Define spline regression in one sentence.	**2 Marks**	**L1**	**CO2**
5.	What is penalized regression?	**2 Marks**	**L1**	**CO2**
6.	Define Model-Based Classifiers	**2 Marks**	**L1**	**CO3**
7.	Define Spectral Clustering?	**2 Marks**	**L1**	**CO3**
8.	What are the applications of K-Means Clustering	**2 Marks**	**L2**	**CO3**
9.	Define Feedforward Neural Network (FNN)?	**2 Marks**	**L1**	**CO4**
10.	Explain the purpose of regularization in deep learning?	**2 Marks**	**L2**	**CO4**

Part B

Answer the Questions. **Total Marks 80M**

11.	**a.**	**Explain the difference between descriptive and inferential statistics with examples.**	**10 Marks**	**L2**	**CO1**
		Or			
12.	**a.**	Explain measures of dispersion with examples.	**10 Marks**	**L2**	**CO1**
13.	**a.**	**Discuss graphical summaries: histograms and box plots, with their uses.**	**10 Marks**	**L2**	**CO1**
		Or			
14.	**a.**	Q1 Consider the data set: 17, 10, 9, 14, 13, 17, 12, 20, 14 Find Mean,Median,Mode and Varience	**10 Marks**	**L4**	**CO1**
15.	**a.**	**Explain Ridge regression. Discuss how it deals with multicollinearity.**	**10 Marks**	**L2**	**CO2**
		Or			
16.	**a.**	For the Given data X: 1,2,3,4,5 and Y: 2,4,5,4,5 in Linear regression model (a) Compute slope β1 and intercept β0. (b) Write the Linear regression equation.	**10 Marks**	**L3**	**CO2**
17.	**a.**	**What is penalized regression? Explain its need with examples of Lasso and Ridge.**	**10 Marks**	**L2**	**CO2**
		Or			
18.	**a.**	Compare Lasso regression (L1) and Ridge regression (L2)	**10 Marks**	**L4**	**CO2**
19.	**a.**	**Explain Naive Bayes Classifier in detail and discuss how it is used in real-world applications.**	**10 Marks**	**L2**	**CO3**
		Or			
20.	**a.**	Describe the Kernel Density Estimation (KDE) classifier and explain how bandwidth selection affects its performance.	**10 Marks**	**L2**	**CO3**
21.	**a.**	**Classify a** new fruit (F) **with:Weight =** 145 grams**Texture =** Smooth (1)**based on its** weight **and** texture **(smooth or rough). Using KNN Classifier**	**10 Marks**	**L4**	**CO3**

		Fruit	**Weight (grams)**	**Texture (1 = mooth, 0 = Rough)**	**Class**			
		A	150	0	Orange			
		B	170	0	Orange			
		C	140	1	Apple			
		D	130	1	Apple			
		E	160	0	Orange			
Or								
22.	**a.**	Describe Decision Trees in detail and discuss how they handle non-linear relationships in classification tasks.				**10 Marks**	**L2**	**CO3**
23.	**a.**	**Describe the structure and working of Convolutional Neural Networks (CNNs) for prediction tasks.**				**10 Marks**	**L2**	**CO4**
Or								
24.	**a.**	Explain the role of deep learning in recommender systems.				**10 Marks**	**L2**	**CO4**
25.	**a.**	**Describe how deep learning is applied in sentiment analysis.**				**10 Marks**	**L2**	**CO4**
Or								
26.	**a.**	Explain RNNs and how GRUs improve sequential modeling				**10 Marks**	**L2**	**CO4**