U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

June 2025 Semester End Main Examinations

Programme: B.E. **Semester: VII**

Branch: Electronics & Telecommunication Engineering **Duration: 3 hrs.**

Course Code: 22ET7PE3DS **Max Marks: 100**

Course: DATA SCIENCE

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

		UNIT - I	*CO*	*PO*	Marks
1	a)	Discuss the following in python, with an example code each. (i) Variables and types (ii) Operators on numbers and strings (iii) Assignments (iv) Typecasting	*CO2*	*PO1*	**10**
	b)	Explain functions with suitable code for the following cases in python. (a) function with args (b) function with *args (c) function with kwargs (d) function with **kwargs	*CO2*	*PO1*	**10**
		OR			
2	a)	Develop python functions for the following operations in linear algebra. (i) vector addition (ii) vector subtraction (iii) vector element by element addition (iv) scalar multiplication (v) dot product	*CO2*	*PO1*	**10**
	b)	List the types of data containers available in python and explain with an example code.	*CO2*	*PO1*	**10**
		UNIT - II			
3	a)	For the list given in x=[1,2,3,1,4,5,1,3,2,4,5,1,1], find : (i) mean (ii) median (iii) quantile (x,25%) (iv) mode	*CO3*	*PO2*	**10**

		b)	Define the following (i) Range (ii) Variance (iii) Standard deviation (iv) Covariance (v) Correlation coefficient	*CO2*	*PO1*	**10**
			OR			
	4	a)	Consider a family with two children. If we assume that each child is equally likely to be a boy or a girl and the gender of the second child is independent of the gender of the first child. Then find the probability of (i) No girls (ii) One girl, one boy (iii) Two girls (iv) Both children are girls conditional on the event "the older child is a girl". (v) Both children are girls conditional on the event "at least one of the children is a girl".	*CO3*	*PO2*	**10**
		b)	Answer the following with respect to normal distribution : (i) Write the Expression of normal distribution (ii) Python implementation of normal distribution (iii) Draw the following normal distributions on the same plot. for (mean=0, standard deviation=1) and (mean=-1, standard deviation=1)	*CO3*	*PO2*	**10**
			UNIT - III			
	5	a)	Describe PCA algorithm with relevant expressions and example.	*CO2*	*PO1*	**10**
		b)	With respect to Gradient Descent algorithm. (i) List the steps involved in minimizing a function. (ii) Develop a python code for Gradient Descent algorithm	*CO2*	*PO1*	**10**
			OR			
	6	a)	Implement PCA algorithm using python code	*CO3*	*PO2*	**10**
		b)	Develop the following functions in python with respect to the multi dimension data. (i) Shape( ) (ii) Dot_product( ) (iii) Sum_of_squares( ) (iv) Variance( ) (v) Standard deviation( )	*CO2*	*PO1*	**10**
			UNIT - IV			
	7	a)	Following are the statistics of a test conducted for a particular disease in a laboratory :	*CO3*	*PO2*	**10**

TP=70, FP=4930, FN=13930, TN=981070. Calculate the following.

(i) Accuracy
(ii) Precision
(iii) Recall
(iv) F1 Score

Also comment on the results optained.

b) For the dataset given below, Apply KNN algorithm to predict the possible sport for the query (Angelina, 5 years, female) using k=3. Assume Male=0 and Female =1

Name	Age (years)	Gender	Sport
Ajay	32	Male	Foot ball
Mark	40	Male	Neither
Sara	16	Female	Cricket
zaira	34	Female	Foot ball
Sachin	55	Male	Neither
Rahul	40	Male	Cricket

CO3 *PO2* **10**

OR

8 a) Discuss the working of Naive Bayes classifier with necessary equations. Also describe the mathematics behind a spam filter with an example. *CO2* *PO1* **10**

b) Apply multiple regression to estimate the value of 'y' given x1=3 and x2=2 as shown in the table below. *CO3* *PO2* **10**

Subject	y	x1	x2
1	-3.7	3	8
2	3.5	4	5
3	2.5	5	7
4	11.5	6	3
5	5.7	2	1
6	**?**	**3**	**2**

UNIT - V

9 a) With relevant expressions and graphs explain the working of
(i) Linear regression
(ii) Logistic regression *CO2* *PO1* **10**

b) The (x, y) coordinates of few points in 2 dimensional space belonging to 2 different class groups are as given below. Here 'x' is the point along 'x' axis and 'y' is the point along 'y' axis. *CO3* *PO2* **10**

Class 1	(1,1), (-1,1), (-1,-1), (1,-1)
Class 2	(2,0), (0,2), (-2,0), (0,-2)

Apply SVM algorithm to answer the following.

a. Plot the points on a 2-D plane with different representations for class1 & class2.
b. Find the transformed sample points.

			c. Estimate the hyper plane coefficients α with respect to each support vector d. Estimate the hyper plane equation along with weight vector and bias e. Draw the classifier line and hyper plane on given sample points in 2-D plane.			
			OR			
	10	a)	Describe the following algorithms with necessary equations and tree diagrams : (i) Decision tree algorithm (ii) Random forest algorithm	*CO2*	*PO1*	**10**
		b)	Describe the following algorithms with necessary equations and graphs : (i) K means clustering (ii) Agglomerative clustering	*CO2*	*PO1*	**10**

****