

**K. J. Somaiya Institute of Technology, Sion, Mumbai-22**  
(Autonomous College Affiliated to University of Mumbai)

**April – May 2023**

(B.Tech / M.Tech.) Program: B.Tech(Scheme: I)

Examination: **TY**  
Course Code: **1UEXC602**  
Date of Exam: **27/05/23**

Semester: **VI**  
Course Name: **Machine Learning**  
Duration: **2.5 Hours**

Max. Marks: **60**

Instructions:

- (1) All questions are compulsory.
- (2) Draw neat diagrams wherever applicable.
- (3) Assume suitable data, if necessary.

	M	CO	BT
Q 1 Solve any six questions out of eight:	12		
i) Define Machine Learning.	2	1	U
ii) What is feature scaling?	2	2	U
iii) What is soft margin in SVM?	2	3	U
iv) Define precision.	2	4	U
v) Differentiate between Supervised Learning and Unsupervised Learning	2	5	U
vi) Differentiate between anomaly detection and supervised learning.	2	6	U
vii) Define accuracy.	2	4	U
viii)	2	3	A

Sepal length	Sepal width	Petal length	Petal Width	Class
3	2	3	4	Setosa
7	4	2	1	Versicolor

The above data represents the iris dataset. Find the distance between the datapoints using both Euclidean and Manhattan technique.

Q.2 Solve any four questions out of six.	16		
i) Identify the type of the machine learning problem. Suggest the	4	1	U
a) Predicting a house price			
b) Movie recommendation			
c) Autonomous vehicle			
d) Spam email detection			
ii) Write a short note on learning rate.	4	2	U
iii) Write expression for hypothesis, cost function and for parameter using gradient descent for univariate logistic regression. Explain each term in short.	4	3	U
iv) Write a short note on ROC.	4	4	A
v) Define dimensionality reduction. Write advantages of dimensionality reduction.	4	5	U

vi) Write short note on recommender system.

4 6 U

Q.3 Solve any two questions out of three.

16

i) What are the types of Machine Learning? Explain in brief with examples.

8 1 U

ii)

Height	175	188	182	177	170	175
Weight	74	96	100	93	69	83

8 2 A

Calculate the Squared Error function for the above dataset for  $\theta_0 = 4$  and  $\theta_1 = 0.45$ . Predict the weight of a person with height 174 cm.

iii) Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into 3 clusters and maximum 3 iterations. Consider A1, A4 and A7 as initial centroids. A1=(2,10), A2=(2,5), A3=(8,4), A4=(5,8), A5=(7,5), A6=(6,4), A7=(1,2), A8=(4,9).

8 5 A

Q.4 Solve any two questions out of three.

16

i)

Age	20	32	18	29	47	45	46	48	45
Salary	86000	18000	82000	80000	25000	26000	28000	29000	22000
Bought	0	0	0	0	1	1	1	1	1

8 3 A

Using KNN, predict whether a person with age 27 and salary 67000 will buy a car or not. Consider  $k=3$ .

ii) Draw a confusion matrix for a ML model that predicted 97 non spams and 17 spams mail correctly, while 4 non spam and 7 spam mail incorrectly. Calculate accuracy, precision, recall and F1 score.

8 4 A

iii) Write short note on

8 6 U

1. Anomaly detection
2. Online Learning

\*\*\*\*\*