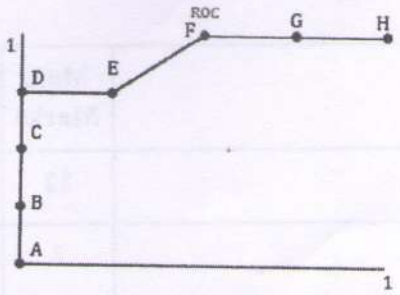
Date : 20-05-2022

K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)
End Semester Exam
May 2022 – June 2022
B.Tech Program: Electronics and Telecommunication Engineering
Examination: TY Semester: VI
Course Code: 1UEXC602 and Course Name: Machine Learning
Duration: 03 Hours Max. Marks: 60

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

		Max. Marks	CO	BT level
Q 1	**Solve any six questions out of eight:**	**12**		
i)	Explain types of Machine Learning.	**2**	CO1	U
ii)	Explain learning rate.	**2**	CO2	U
iii)	Explain logistic regression.	**2**	CO3	U
iv)	Explain soft margin classifier.	**2**	CO3	U
v)	Define F1-score.	**2**	CO4	U
vi)	Define ROC.	**2**	CO4	U
vii)	Explain clustering.	**2**	CO5	U
viii)	Explain anomaly detection.	**2**	CO6	U
Q.2	**Solve any four questions out of six.**	**16**		
i)	Explain Reinforcement Learning with example.	**4**	CO1	U
ii)	Compare advantages and disadvantages of Normal Equation over Gradient Descent.	**4**	CO2	An
iii)	Design expression for overall cost function in SVM. Explain the cost function with a graph.	**4**	CO3	C
iv)	Explain Confusion Matrix with an example.	**4**	CO4	U
v)	Write short note on dimensionality reduction.	**4**	CO5	U
vi)	Write short note on online learning.	**4**	CO6	U

Q.3	**Solve any two questions out of three.**	**16**		
i)	<table><tr><td>Age</td><td>20</td><td>32</td><td>18</td><td>29</td><td>47</td><td>45</td><td>46</td><td>48</td><td>45</td></tr><tr><td>Salary</td><td>86000</td><td>18000</td><td>82000</td><td>80000</td><td>25000</td><td>26000</td><td>28000</td><td>29000</td><td>22000</td></tr><tr><td>Bought</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td></tr></table> The above data shows the database of an automobile company whether a person with the given age and salary has bought a car or not. For $\theta_0 = 0.0002$, $\theta_2 = -10$, calculate predicted value, error and updated value of θ using logistic regression after an iteration.	**8**	CO3	A
ii)	Define ROC. Which of the following point on ROC gives the best threshold for the application that predicts if the patient can be discharged after Covid treatment? Why? 	**8**	CO4	An
iii)	Explain how large dataset is useful in machine learning? What is the problem with large dataset in gradient descent? How large datasets are dealt with in gradient descent.	**8**	CO6	An
Q.4	**Solve any two questions out of three.**	**16**		
i)	Illustrate process of learning with the gradient descent for a univariate linear regression, using a bell shaped error curve. Explain how a step size is modulated on every iteration.	**8**	CO2	An
ii)	Write expression for hypothesis, cost function and for parameter using gradient descent for multivariate linear regression. Explain each term in short.	**8**	CO1	U
iii)	Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into 3 clusters: 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**	CO5	A