JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
TEST -3 EXAMINATION- 2025
M.Tech-I Semester (CSE-Data Science)

COURSE CODE (CREDITS): 22M1WCI133(3) MAX. MARKS: 35
COURSE NAME: INTRODUCTION TO STATISTICAL LEARNING
COURSE INSTRUCTORS: PROF. PARDEEP KUMAR MAX. TIME: 2 Hours

Note: (a) All questions are compulsory.
(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems
(c) Calculator is allowed.

Q. No	Question	CO	Marks
Q1	(a) Consider 2-dimensional dataset with label=+1 and label=-1: Label +1: (2,2), (4,4), (3,3) Label -1: (4,0), (0,0), (0,2) Find the hyperplane, support vectors and maximal marginal hyperplane (MMH) margin using linear support vector machine. (b) Consider the 2-dimensional dataset label=+1 and label=-1: Label +1: (2,0), (0,2), (-2,0), (0,-2) Label -1: (0,0), (0.5,0), (0,0.5) Consider the above points are not linearly separable in (x_1, x_2) space. The feature map function is given as : $\Phi(x_1, x_2, x_1^2 + x_2^2)$. Find the hyperplane, support vectors and maximal marginal hyperplane (MMH) margin using nonlinear support vector machine.	4	5+8
Q2	Consider 8 persons showing interest in personal loan from a reputed public sector bank. Let point(x,y,z) represents their age(in years), salary(in Indian rupees) and income from other sources(in Indian rupees). So the data is represented as Ram(20,35K,10K), Sita(30,25K,5K), Laxman(50,20K,12K), Anil(40,20K,2K),Prem(45,50K,15K),Sunil(21,46K,15K),Sourabh(23,32K,11K)andMeenu(32,27K,13K). The bank manager's task is to check whether the person is rich, medium earning or poor based on the given data. Use k-means clustering algorithm technique based on Euclidean distance to ease manager's task. Show the cluster formation using graphs at each iteration of K-means algorithm.	3	10
Q3	Suppose we have height, weight and T-shirt size of some customers and we need to predict the T-shirt size of a new customer given only height and weight information we have. Data including height, weight and T-shirt size information is shown below –	3	5

Height (in cms)	Weight (in kgs)	T Shirt Size
158	58	M

158	59	M	
158	63	M	
160	59	M	
160	60	M	
163	60	M	
163	61	M	
160	64	L	
163	64	L	
165	61	L	

New customer named 'Monica' has height 161cm and weight 61kg. Predict the T shirt size of Monica using lazy learner algorithm. Your initialized parameter should be 3.

Q4 Consider a shopping mall basket as shown below: 3 7

Customer	Items
C1	Milk, Egg, Bread, Chip
C2	Egg, Popcorn, Chip, Beer
C3	Egg, Bread, Chip
C4	Milk, Egg, Bread, Popcorn, Chip, Beer
C5	Milk, Bread, Beer
C6	Egg, Bread, Beer
C7	Milk, Bread, Chip
C8	Milk, Egg, Bread, Butter, Chip
C9	Milk, Egg, Butter, Chip

Let the minimum support is 40% and minimum confidence is 80%. Generate the information using Apriori algorithm to describe the behavior of customers.