BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: B.Tech. **SEMESTER : IV**
BRANCH: Ftech **SESSION : SP/2025**

SUBJECT: FE223 STATISTICAL MACHINE LEARNING I

TIME: 3 Hours **FULL MARKS: 50**

INSTRUCTIONS:
1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
4. Before attempting the question paper, be sure that you have got the correct question paper.
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

			CO	BL
Q.1(a)	Consider the example of EnjoySport Task with Instances X: Possible days, each described by the attributes.	[7]	1	1,2

i. Sky (with possible values Sunny, Cloudy, and Rainy),
ii. AirTemp (with values Warm and Cold),
iii. Humidity (with values Normal and High),
iv. Wind (with values Strong and Weak),
v. Water (with values Warm and Cool), and
vi. Forecast (with values Same and Change).

Let us define a new hypothesis space H' that consists of all pairwise disjunctions of the hypotheses in H. For example, a typical hypothesis in H' is (? Cold, High, ?, ?, ?) v (Sunny, ?, High, ?, ?, Same). Trace the CANDIDATE-ELIMINATION for the hypothesis space H' given the sequence of training examples

Example	*Sky*	*AirTemp*	*Humidity*	*Wind*	*Water*	*Forecast*	*EnjoySport*
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

Q.1(b)	Describe Bias–Variance & Prediction Accuracy-Interpretability Tradeoff with suitable examples.	[3]	1	2,3
Q.2(a)	You have a friend who only does one of four things every Saturday afternoon: **go shopping, watch a movie, play tennis, or just stay in**. You have observed your friend's behaviour over 11 different weekends. On each of these weekends you have noted the weather (sunny, windy, or rainy), whether her parents visit (visit or no-visit), whether he/she has drawn cash from an ATM (rich or poor), and whether he/she had an exam during the coming week (exam or no-exam).	[5]	2	2,3

You have built the following data table:

# ex.	Weather	Parents	Cash	Exam	Decision
1	sunny	visit	rich	yes	cinema
2	sunny	no-visit	rich	no	tennis
3	windy	visit	rich	no	cinema
4	rainy	visit	poor	yes	cinema
5	rainy	no-visit	rich	no	stay-in
6	rainy	visit	poor	no	cinema
7	windy	no-visit	poor	yes	cinema
8	windy	no-visit	rich	yes	shopping
9	windy	visit	rich	no	cinema
10	sunny	no-visit	rich	no	tennis
11	sunny	no-visit	poor	yes	tennis

You now want to build a decision tree to predict the activity of your friend on any future Saturday afternoon from the observed values of Weather, Parents, Cash, and Exam.

PTO

i. Suppose that you want to build a very simple decision tree that allows you to predict the value of Decision from a single observable attribute (Weather, Parents, Cash, or Exam). What would be this attribute? If there is a tie among several attributes, give each of them. For each attribute that you have selected for each value of this attribute, answer what the decision tree would give using the majority rule, as well as the number of misclassified examples.

ii. Draw the full tree that correctly classifies all the examples.

Q.2(b) The table below provides a training data set containing six observations, three predictors, and one qualitative response variable. [5] 2 3

Obs.	X_1	X_2	X_3	Y
1	0	3	0	Red
2	2	0	0	Red
3	0	1	3	Red
4	0	1	2	Green
5	−1	0	1	Green
6	1	1	1	Red

Suppose we wish to use this data set to predict Y when X1 = X2 = X3 = 0 using K-nearest neighbours. Compute the Euclidean distance between each observation and the test point, X1 = X2 = X3 = 0. Evaluate your prediction with K = 3? Why?

Q.3(a) Consider the following entries in the dataset with Target attribute as Profit. Evaluate the information gain, gain & gain ratio of the attributes while creating the decision tree. [5] 3 4

Age	Competition	Type	Profit
Old	Yes	S/W	Down
Old	No	S/W	Down
Old	No	H/W	Down
Mid	Yes	S/W	Down
Mid	Yes	H/W	Down
Mid	No	H/W	Up
Mid	No	S/W	Up
New	Yes	S/W	Up
New	No	H/W	Up
New	No	S/W	Up

Q.3(b) Two astronomers, in different parts of the world, make measurements M1 and M2 of the number N of stars in some small regions of the sky, using their telescopes. Normally, there is a small probability of errors up to one star. Each telescope can also (with a slightly smaller probability) be badly out of focus (events F1 and F2), in which case the astronomer will underscore the count by three or more stars. Draw a Bayesian network structure that represents this story. [5] 3 3,4

Q.4(a) Consider a set of 2-dimensional training data points (x1, x2) belonging to two classes +1 and -1 respectively as shown below. We design a linear hard margin SVM to classify them. The equation of the optimal separating hyperplane is? [5] 4 4,4,6

+1: (3, 1), (3, −1), (6, 1), (6, −1)

-1: (1, 0), (0, 1), (0, −1), (−1, 0).

Q.4(b) Suppose that the task is to cluster the following eight points (with (x,y) representing location) into three clusters: [5] 4 4

A1(2,10), A2(2,5), A3(8,4), A4(5,8), A5(7,5), A6(6,4), A7(1,2), A8(4,9)

The distance function is Euclidean distance. Suppose initially we assign A1, A4 and A7 as the center of each cluster, respectively. Use the k-means algorithm to show only the three cluster centers after the first round of execution.

Q.5	Consider a neural network with two inputs, two hidden layers and two outputs. The input received by I_1 and I_2 are 0.1 and 0.5 respectively. The weights w_1 from I_1 to H_1 is 0.1, w_2 from I_1 to H_2 is 0.2, w_3 from I_2 to H_1 is 0.3 and w_4 from I_2 to H_2 is 0.4. The weights w_5 from H_1 to O_1 is 0.5, w_6 from H_2 to O_1 is 0.6, w_7 from H_1 to O_2 is 0.7 and w_8 from H_2 to O_2 is 0.8. The bias from Input to Hidden layer is 0.25 and from Hidden to Output layer is 0.35. The expected outputs from O_1 and O_2 are 0.05 and 0.95 respectively. Sigmoid function (Logistic function) is the activation function. Perform the backpropagation algorithms for any of the weights of your choice up to six decimal points precision. The learning rate is 0.6	[10]	5	4,5,6

:::::01/05/2025 M:::::