

JADAVPUR UNIVERSITY
DEPARTMENT OF INFORMATION TECHNOLOGY
M.E. INFORMATION TECHNOLOGY, FIRST YEAR, FIRST SEMESTER EXAM 2025
SUBJECT: MACHINE LEARNING

Full Marks :100

Time : 3Hrs

CO1 : (20)

Answer any two questions

2 X 10 = 20

i)

- a) Explain training versus testing.
- b) Apply Candidate Elimination algorithm on the data set given in Fig.1.

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

Fig.1

3+7=10

ii)

- a) Give the structure of a machine learning model.
- b) What do you mean by a well-posed learning problem? Explain the important features that are required to well define a learning problem.
- c) Explain Find -S algorithm with the data set given in Fig.1. Give its application.

2+3+5=10

iii)

- a) Consider the problem of sorting 'n' numbers. Is it wise to solve this problem using machine learning? Justify.
- b) Explain the various stages involved in designing a learning system.
- c) Explain the concept of Inductive Bias
- d) What is the significance of concept learning in designing machine learning models?

3+2+2+3=10

[Turn over

CO2 : (30)

Answer any two questions

2 X 15 = 30

i)

- a) What is the role of the maximum likelihood function in designing the logistic regression model?
- b) Why is the term "logistic regression" used to name this method?
- c) Marks obtained by 10 students in the class test and semester examination in machine learning are provided in the following table. Estimate the marks a student may obtain in the semester examination when she got 20 in class test, using linear regression.

Sl No	Class Test Marks	Semester Marks	Sl No	Class Marks Test	Semester Marks
1	28	63	6	28	51
2	27	49	7	26	66
3	23	43	8	21	36
4	17	36	9	22	31
5	24	39	10	19	37

- d) Explain soft margin and hard margin Support Vector Machine.

3+3+5+4 =15

ii)

- a) Explain the objective function of Soft SVM and clarify how it differs from Hard SVM?
- b) Justify the statement that "One can use infinite-dimensional spaces with the kernel trick" in the perspective of non-linear SVM classification.
- c) What is the use of regularization constant in SVM?
- d) A linearly separable dataset is given in the following Table. Predict the class of (0.6, 0.8) using a support vector machine classifier. Show all the relevant computations.

X1	X2	Y	Lagrange Multiplier
0.3858	0.4687	+1	65.5261
0.4871	0.611	-1	65.5261
0.9218	0.4103	-1	0
0.7382	0.8936	-1	0
0.1763	0.0579	+1	0
0.4057	0.3529	+1	0
0.9355	0.8132	-1	0
0.2146	0.0099	+1	0

4+3+3+5 =15

iii)

- a) Explain the purpose of linear regression to develop a data driven model.
- b) Suppose your model is $Y = \beta_0 + \beta_1 X$, where X is input (regress) variable and β 's are regression coefficients. Y is noted as predictor variable. Show that the regression coefficients are computed using the expression, $\beta = (X^T X)^{-1} X^T Y$
- c) Apply least square method to fit a model $Y = \beta_0 + \beta_1 X$, through the following data points.

X	1	2	3	4	5	6	7
Y	9	8	10	12	11	13	15

- d) What is the intuition of a large margin classifier?

3+5+4+3 =15

CO3 : (30)

Answer any two questions

2 X 15 = 30

i)

a) Consider the following set of training examples:

A1	A2	Classification
T	T	+
T	T	+
T	F	-
F	F	+
F	T	-
F	T	-

What is the entropy of this collection of training example with respect to the target function classification? What is the information gain of A2 relative to these training examples?

b) ID3 may suffer from overfitting problem – justify.

c) Define Information Gain and interpret each term in the equation.

(4+3)+3+5 =15

ii)

a) What is cohesion, and how can it be measured?

b) What are ordinal variables? Give example.

c) Construct a dendrogram showing each step using the given dataset and generate four clusters from it.

Points	A	B	C	D
A	0	2	6	10
B	2	0	5	9
C	6	5	0	4
D	10	9	4	0

5+3+7 =15

iii)

a) Briefly describe Complete Linkage algorithm.

b) How does K-Means handle outliers, and what techniques can be used to reduce the impact of outliers on clustering results?

c) In what scenarios is DBSCAN a better choice than K-Means, and what are the key advantages of DBSCAN in handling complex datasets?

d) How do ϵ and MinPts affect the clustering outcome in DBSCAN? What strategies or techniques can be used to determine optimal values for these parameters?

4+4+3+4=15

CO4 :(10)	Answer any two questions 2 X 5 =10 i) Explain different types of Gradient Descent in detail? ii) What is the difference between forward and backward propagation in neural network model? iii) How does the learning rate affect the training in neural network? iv) What is the role of activation function in Artificial Neural Network?
CO5 : (10)	Answer any two questions 2X5 =10 i) How does bootstrap resampling differ from traditional sampling techniques, and what are its advantages and limitations in statistical modeling? ii) Briefly describe Bagging Algorithm. iii) What are policies in reinforcement learning? iv) What do you mean by Long -term Reward and Immediate Reward?