

III B.Tech II Semester Regular Examinations, May 2023

MACHINE LEARNING

(Artificial Intelligence and Data Science)

Time: 3 Hours**Max. Marks: 70****Note:** Answer one question from each unit.

All questions carry equal marks.

5 × 14 = 70M**UNIT-I**

1. a) How does the central nervous system helps in creating the intelligent machines? (6M)
- b) What is machine learning? Explain any three business applications of machine learning. (8M)

(OR)

2. a) What is a well posed learning problem? Specify task, performance measure and training experience for the following learning problems. (7M)
 - (i) A checkers learning problem.
 - (ii) A handwritten recognition learning problem.
 - (iii) A robot driving learning problem
- b) With a neat diagram explain the block diagram of machine learning. (7M)

UNIT-II

3. a) How Bayes' theorem supports the classification task? Explain with an example. (7M)
- b) Explain the concept of linear regression with least square error criterion. (7M)

(OR)

4. a) Consider the dataset D , calculate probability for unseen pattern $(x:\{M,1.85\})$ using Naïve Bayes Classifier. (7M)

	Gender x_1	Height x_2	Class	y
$s(1)$	F	1.6 m	Short	y_1
$s(2)$	M	2 m	Tall	y_3
$s(3)$	F	1.9 m	Medium	y_2
$s(4)$	F	1.88 m	Medium	y_2
$s(5)$	F	1.7 m	Short	y_1
$s(6)$	M	1.85 m	Medium	y_2
$s(7)$	F	1.6 m	Short	y_1
$s(8)$	M	1.7 m	Short	y_1
$s(9)$	M	2.2 m	Tall	y_3
$s(10)$	M	2.1 m	Tall	y_3
$s(11)$	F	1.8 m	Medium	y_2
$s(12)$	M	1.95 m	Medium	y_2
$s(13)$	F	1.9 m	Medium	y_2
$s(14)$	F	1.8 m	Medium	y_2
$s(15)$	F	1.75 m	Medium	y_2

- b) In a particular pain clinic, 10% of patients are prescribed narcotic pain killers. Overall, five percent of the clinic's patients are addicted to narcotics (including pain killers and illegal substances). Out of all the people prescribed pain pills, 8% are addicts. If a patient is an addict, what is the probability that they will be prescribed pain pills? (use Baye's theorem) (7M)

UNIT-III

5. Consider the following positively and negatively labeled data points.
Positively labeled data points: (3,1), (3,-1), (6,1), (6,-1)
Negatively labeled data points: (1,0), (0,1), (0,-1), (-1,0)
(i) Does these points linearly separable? (ii) Determine the hyper plane that classifies these points. (14M)

(OR)

6. a) Suppose that we have trained a classifier to classify medical data tuples, where the class label attribute is *cancer* and the possible class values are *true* and *false*. The confusion matrix of model is as given below. Compute the following. (8M)

	Predicted +ve	Predicted -ve
Actual +ve	4	7
Actual -ve	4	185

- (i) Success rate (ii) Miss classification rate (iii) Sensitivity (iv) Specificity
b) In support vector machines fixing the maximum margin sometimes leads to miss classification of some of the samples. How to overcome this problem? (6M)

UNIT-IV

7. a) Given the two objects represented by the tuples (22, 1, 42, 10) and (20, 0, 36, 8); Compute (i) Euclidean distance between the two objects. (ii) Manhattan distance between the two objects. (iii) Murkowski distance between the two objects, using $q=3$. (7M)
b) State the strengths and weaknesses of k-means clustering algorithm. (7M)

(OR)

8. Suppose that the data mining task is to cluster the following eight points (with (x, y) representing location) into three clusters: (14M)
A1(2, 10), A2(2, 5), A3(8, 4), B1(5, 8), B2(7, 5), B3(6, 4), C1(1, 2), C2(4, 9). The distance function is Manhattan distance. Suppose initially we assign A1, B1, and C1 as the center of each cluster, respectively. Use the k-means algorithm to show only
(i) The three cluster centers after the first round execution.
(ii) The final three clusters.

UNIT-V

9. a) Design a neural network that implements Boolean AND gate. (8M)
b) What is ANN? Briefly explain the problem characteristics for which ANN is most commonly used. (6M)

(OR)

10. a) Write the Back-propagation algorithm. How the error is minimized in the Back-propagation algorithm. (7M)
b) Draw the structure of a biological neuron and explain the functionality of axons and dendrites of the neuron. (7M)
