



SPRING END SEMESTER EXAMINATION-2024

6th Semester B.Tech

MACHINE LEARNING

CS 3035

(For 2021 & Previous Admitted Batches)

Time: 2 Hours 30 Minutes

Full Marks: 50

Answer any FIVE questions.

Question paper consists of two SECTIONS i.e. A and B.

Section A is compulsory.

Attempt any Four question from Sections B.

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable and all parts of a question should be answered at one place only.

SECTION-A

1. Answer the following questions: [1 × 10]

- (a) Why Multiple Linear Regression should not be applied as a prediction model to highly correlated data?
- (b) Calculate the MSE and MAE for the given test data set of a sales prediction

Year	Actual Sales	Predicted sales
2021	150	152
2022	162	165
2023	220	208
2024	272	269

- (c) What is the log odds coefficient and how does it help Logistic Regression for classification?
- (d) Analyze the need for Bias Variance tradeoff in classification.
- (e) What principle underlies the decision tree classifier?
- (f) Why Bayesian classification is called naive?

- (g) State the limitations of the K-Means Clustering algorithm and mention the remedies to overcome those situations.
- (h) How Divisive clustering is different from Agglomerative clustering. Also, mention the different types of Divisive clustering.
- (i) How to set threshold value in McCulloch-Pitt Neuron?
- (j) Discuss the importance of bias and the learning rate of neural networks.

SECTION-B

2. (a) In a sample observation, there are two independent predictors X_1 and X_2 , and a target Y as given in the table. [5]

X_1	X_2	Y
2	1	5.7
6	3	11.5
5	7	2.5
3	8	-3.7
4	5	3.5

Design a Multiple Linear Regression model that best fits the above observation. If a new sample has $X_1=8$, and $X_2=4$ and predict its target using your designed MLR model.

- (b) Derive the expression for computing the coefficients of Ridge Regression, formulate its pseudo code, and justify that it is a shrinkage method. [5]
3. (a) In a pathological laboratory, there were 10,000 blood samples collected in a day, out of which 300 are diabetic. The diabetic detection model detects 240 as diabetic, out of which 40 are diabetic. Formulate the Confusion Matrix and compute the Accuracy, Precision, Recall, and F1-Score of your detection system. [5]

- (b) What is the intuition of the maximum margin in the SVM classifier? Briefly explain the role of hyper-parameter in SVM and specify a situation where SVM is preferable over the Decision tree. [5]

4. (a) Apply the PCA algorithm to the following observations and compute the Covariance matrix and the Eigenvalues. [5]

Data points	X1	X2
S1	5	12
S2	14	6
S3	9	5
S4	8	15

- (b) Plot the hyper-plane for the following data points using a Support Vector Machine. [5]

Data points: (1,1), (2,1), (1, -1), (2,-1), (4,0), (5,1), (5, -1), (6, 0)

5. (a) Consider the following sample of observations. Divide the sample data set into two clusters using the K-means clustering algorithm by using Euclidean distance. [5]

Observation	Price of Battery (Rs.)	Battery Life (Years)
B1	299	4
B2	494	3
B3	499	2
B4	998	1
B5	463	2
B6	799	5

Use (799, 5) and (494, 3) as initial cluster centers.

- (b) For the given set of points identify clusters using complete link and average link using agglomerative clustering and draw the dendrogram. [5]

Observations	A	B
P1	1	1
P2	1.5	1.5
P3	5	5
P4	3	4
P5	4	4
P6	3	3.5

6. (a) Implement Logical AND function using perceptron networks for bipolar inputs and targets. (Note: initial weights and learning rate are set to zero and 1 respectively) [5]
- (b) Compute the error portions δ_k and δ_j as well as find the newly updated weights of the output layer and hidden layer respectively using the Back Propagation algorithm. [5]

The network is presented with input data $[-1, 1]$ and the target output is 1.

Use learning rate $\alpha=0.25$ and binary sigmoid activation function.


