

Nirma University

Institute of Technology

Semester End Examination (IR), December - 2022

B. Tech. in CL / CH / ME / EE, Semester-V

2CSOE51 Machine Learning

Roll /
Exam
No.

Supervisor's
initial with
date

Time: 3 Hours

Max. Marks: 100

Instructions:

1. Attempt all questions.
2. Figures to right indicate full marks.
3. Use section-wise separate answer book.
4. Draw neat sketches wherever necessary.
5. Assume suitable data wherever necessary.
6. Use of scientific calculator is allowed.

SECTION - I

Q-1. Answer the following questions.

[16]

(A) Discuss supervised learning and unsupervised learning with respect to its' appropriate application as per your engineering branch.

(4)

CO3
BL2

(B) You are attempting to model the price of cars at auction and you find that your model has trained well, but subsequently does poorly on new data. Explain the reason why?

(6)

CO2
BL2

(C) In classification, what is a class imbalance problem and discuss how to handle this case?

(6)

CO2
BL2

Q-2. Do as directed.

[18]

(A) Discuss how Support Vector Machine can be used for binary classification.

(6)

CO2
BL2

(B) Generate box plot for the values of the attribute 'Test score' given in the table and discuss how can outliers be detected using box plot?

(6)

CO1
BL3

x, Hours of preparation	5	2	9	6	10
y, Test score	64	48	72	73	80

OR

(B) Consider the following data given in the table and do the following.

(6)

CO1
BL3

Year	Vehicle sales (in lacs)
2011	26
2012	27
2013	25
2014	26
2015	27
2016	30

- 1) Find the least square regression line $y=ax+b$.
- 2) Use the least square regression line as a model to estimate the sales of the company for the year 2022.

(C) What is feature scaling? Why is it required? Differentiate between min-max normalization and z-score normalization with suitable example. (6)
CO1
BL4

OR

(C) Discuss whether we can get multiple local optimum solutions or not if we solve a linear regression problem by minimizing the sum of squared errors using gradient descent? (6)
CO1
BL4

Q-3. Do as directed. [16]

(A) To build the decision tree for the data given in the following table, find the best attribute for the root node using Information Gain as the purity measure. Note: attribute "Loan-risk" is the class label. Don't create the full decision tree. Find only the root node. (8)
CO2
BL3

Age	Employment	Income-per-month	Pension-plan	Loan-risk
20-34	Salaried	50K-69K	No	High
35-50	Self	50K-69K	Yes	High
35-50	Salaried	70K-89K	Yes	Low
20-34	Self	50K-69K	No	High
51+	Salaried	70K-89K	Yes	Low
51+	Self	90K+	No	High
20-34	Self	90K+	No	Low
35-50	Salaried	70K-89K	No	Low

(B) Consider the training data shown in the table of Question-3(A). The data tuples are described by the attributes: *age*, *employment*, *Income-per-month* and *Pension-plan*. The class-label attribute: *Loan-risk* has two distinct values {High/Low}. Predict the class of the following tuple **X** using naïve-Bayes classification. Show all the computation steps. (8)
CO2
BL3

X = {age=51+, employment=self, Income-per-month=50K-69K, pension-plan=No}.

SECTION – II

Q-4. Do as directed. [16]

(A) Differentiate between Model accuracy and Error rate of the model. (4)
CO1
BL4

(B) What is Artificial neural network (ANN)? Explain the single-layer feed forward architecture of ANN with suitable diagram. (6)
CO2
BL2

(C) While predicting win-loss of teams in World Cup Football using a classification model. Following are the data recorded and consider class wins as positive class and losses as negative class. (6)
CO3
BL5

(a) Correct predictions – 73 wins, 13 losses

(b) Incorrect predictions – 9 wins, 5 losses

Show the Confusion matrix and calculate the precision, recall and F-measure of the model.

Q-5. Answer the following questions.

[18]

(A) Differentiate between Multilayer Feedforward neural network and Backpropagation neural network by taking suitable example and diagram.

(6)

(B) Justify the need to apply Normal Equation method in a regression problem with suitable diagram.

(6)

BL5

OR

(B) Describe the preprocessing steps to improve accuracy and efficiency of the classification/prediction process with proper justification.

(6)

CO3

BL5

(C) Explain the KNN classification method using suitable example and diagrams.

(6)

CO3

BL2

OR

(C) What is Clustering? Discuss the working of hierarchical agglomerative clustering techniques with a suitable example and diagram.

(6)

CO3

BL2

Q-6. Do as directed.

[16]

(A) Consider the following dataset of 5 objects to cluster the data using the k-means clustering technique. Take the value of $k = 2$ and select two initial seeds as M1 (4,6) and M2 (6,7). Show the steps for the two epochs. [Use Euclidean distance function].

(8)

CO2

BL3

Item no.	X	Y
1	5	6
2	4	5
3	4	6
4	6	7
5	7	8

(B) Assume 7 one dimensional input patterns {0.0, 0.17, 0.33, 0.51, 0.67, 0.83, 1.0}. Assume that first three patterns belong to class 0 (with desired output 0) and remaining patterns belong to class 1 (with desired output 1). Design a perceptron to classify these patterns. Use perceptron learning rule. Given learning rate = 0.1 and initial weight and bias is **(-0.4) and (-0.1)** respectively. Show computation for one epoch.

(8)

CO2

BL3