Nirma University
Institute of Technology
Semester End Examination (IR), December - 2021
B. Tech. in Computer Science and Engineering, Semester-V
2CS501 Machine Learning

Q.1 **Answer the following questions:** **[10]**

(a) CLO:1 What is the need of pre-processing of the data? A data scientist [05] uses Naive Bayes, Decision tree, Logistic Regression and ANN for predicting if it will rain tomorrow or not. Initially he forgets to scale the data and applies all the mentioned algorithm and then he scales and apply the same algorithms again. Which of the four will show improvement in terms of time and in terms of output? Explain with reason.

(b) CLO:2 Consider one neuron Artificial Neural Network. Suppose that the [05] weights corresponding to the three inputs have the following values: $W_1 = 2$, $W_2 = -4$, $W_3 = 1$, and the activation of the unit is given by the step-function:

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Calculate what will be the output value y of the unit for each of the following input patterns:

Pattern	P_1	P_2	P_3	P_4
x_1	1	0	1	1
x_2	0	1	0	1
x_3	0	1	1	1

Q:2 **Answer the following questions:** **[15]**

(a) CLO:1 Identify and describe briefly (2 lines) an application for the [05] following machine learning techniques:
1) Linear Discriminant Analysis
2) Logistic Regression
3) Reinforcement Learning
4) Hierarchical Clustering
5) Convolution Neural Networks

(b) CLO:3 Suppose 10000 patients get tested for flu; out of them, 9000 are [05] actually healthy and 1000 are actually sick. For the sick people, a test was positive for 620 and negative for 380. For the healthy people, the same test was positive for 180 and negative for 8820. Construct a confusion matrix for the data and compute the precision and recall for the data.

(c) CLO:3 Assume text collection shown in the following table. Last column [05] shows the class of the text document. Build naïve Bayes text classifier considering the text collection shown in Table below as the training set. Construct bag-of-words index (In bag-of-words index, each cell in an index (i.e. table) corresponds to a row (i.e. text document) and a column (i.e. word in the text document)) and fit Multinomial distribution to the data. Classify the document "Chinese Chinese Chinese Tokyo Japan" using the Bayes classifier. Build the classifier considering Laplacian correction.

Document ID	Text of Document	Class
1	Chinese Beijing Chinese	c
2	Chinese Chinese Shanghai	c
3	Chinese Macao	c
4	Tokyo Japan Chinese	j

Q.3

(a) CLO:2 **Answer the following questions:** [10]
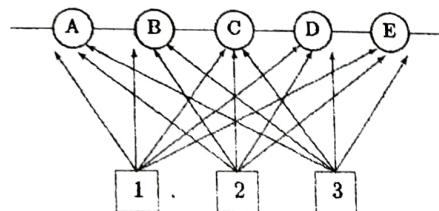
Differentiate between [04]
    a. self organizing maps and artificial neural network.
    b. K-means clustering and K-medoid clustering.

OR

(a) Answer the following questions based on a competitive learning [04] neural network as shown in the figure below which is used to classify the airplane based on the size, speed and passenger capacity. [04]

Initial weights of the output nodes are:

Node	A	B	C	D	E
W1	3	5	1	2	5
W2	2	1	5	3	2
W3	5	1	1	2	5

    a. How many type of planes can this network classify?
    b. Which node will be the winner if the input vector is (1,5,1)?

(b) CLO:2 Take 10 points in two dimensions having coordinate values as: A: [06] (2,10), B:(3,5), C:(2,2), D:(4,5), E:(2,6), F:(10,15), G:(12,14),

H:(14,11), I:(13,13), J:(15,12). Use hierarchical clustering to cluster them into two clusters. Use *min* distance to update the distance between the clusters.

OR

(b) Suppose we are given the dataset listed in Table below for a [06] classification problem. There are two binary input attributes $X1$ and $X2$, and three classes $Y \in \{1, 2, 3\}$. We learn a decision tree classifier.
a. Calculate the information gain for both $X1$ and $X2$.
b. Which attribute is used for the first split at the root node of the tree? Draw the decision tree resulting from using this split alone. Make sure to label the split attribute, the corresponding branches and the predicted label of each leaf. (decision tree of depth 1). (
c. How would this tree classify an example with $X1 = 1$ and $X2 = 0$?

X1	X2	Y
0	0	1
0	0	1
0	0	2
0	1	3
1	1	2
1	1	3

Q.4 **Answer the following questions:** [15]

(a) CLO:2 In an Artificial Neural Network what is the role of hidden layers? [05] What is the relation in the classification capability with number of hidden layer 0,1 2 and 3?

(b) CLO: 2 There are three features for all samples in a dataset. And there is [05] a requirement to construct the polynomial hypothesis of for non-linear regression. Write the equation for hypothesis and list out all the learnable parameters.

(c) CLO:2 Given the hyperplane defined by the line for a SVM [05]
$y = x_1 - 2x_2$
$y = (1, -2)^T x = w^T x$
are these point correctly predicted?
1) $y = 1$, $x = (1,0)$?
2) $y = 1$, $x = (1,1)$?

OR

(c) CLO:2 The problem is related to the Civil engineering Domain. The user [05] wants to estimate the nuber of days(WD) required to complete the task. The user has historical data. Identify 5 atributes that can help in estimating the WD. Explain the attributes, data types and range of values in every attribute. Write Normal equation for regression, also clearly specify the dimension of the vectors X, theta, and Y. Consider we have total 10 instances in the dataset.

****