

Nirma University

Institute of Technology

Semester End Examination, December 2020

BTech CSE Sem. V

2CS501 Machine Learning

Roll /
Exam
No.

Supervisor's
initial with
date

Time: 1.5 Hours

Max. Marks: 40

- Instructions:
1. Attempt all questions.
 2. Figures to right indicate full marks.
 3. Draw neat sketches wherever necessary.
 4. Assume suitable data wherever applicable and clearly mention them.
 5. CLO_ and BL_ have been mentioned against each question to map it as per Course Learning Objective and Bloom's taxonomy.

Q 1 Consider the data given in the following table. [10]

CLO2
BL3,5

Age in Years	Weight in KGs
12	40
19	60
40	80
55	85
23	70
13	70

Divide these data points in three clusters using k-means algorithm. Assume (12, 40), (19, 60) and (40, 80) as initial centroid of three clusters. Use Euclidean Distance as the distance measure. Perform only 1 iteration. What would be centroids of three clusters after 1 iteration?

Q 2 Assume a 3-2-1 feed-forward fully connected neural network [10]

CLO2
BL3,5

(3 neurons in the input layer denoted as neuron 1, 2 and 3, 2 neurons in the hidden layer denoted as neuron 4 and 5, and 1 neuron in the output layer denoted as neuron 6). Assume that the activation function of all the neurons in the hidden and output layers is log-sigmoid. Neurons in the hidden layer and output layer also have bias connection. Consider following table for the initial value of parameters.

W14	W24	W34	W15	W25	W35	W46	W56	b4	b5	b6
0.2	0.5	0.3	0.3	0.1	0.4	0.5	0.6	0.1	0.2	0.3

W14 is the weight of the connection between neuron 1 and 4. b4 is the weight of the bias connection of neuron 4. Other notations can be understood in the same way. Assume learning rate to be 0.1 and one half squared error as the loss function. Assume the input→output pair (training example) to be (2,5,3)→(1). Present this example and backpropagate error to calculate updated value of weight W14.

Q 3 What are support vectors in support vector machine? [2]

CLO1
BL1,2

Q 4 Write steps for performing polynomial regression. [2]
CLO1
BL1,2

Q 5 Differentiate between supervised, unsupervised, semi- [6]
CLO1
supervised and reinforcement learning.

BL4
Q 6 Assume data in the following table: [10]
CLO2
BL3,5

Weather	Temperature	Humidity	Wind	Play?
Sunny	Hot	High	Weak	No
Cloudy	Hot	High	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Cloudy	Mild	High	Strong	Yes
Rainy	Mild	High	Strong	No
Rainy	Cool	Normal	Strong	No
Rainy	Mild	High	Weak	Yes
Sunny	Hot	High	Strong	No
Cloudy	Hot	Normal	Weak	Yes
Rainy	Mild	High	Strong	No

Assuming ID3 decision tree algorithm and the data in the above table as training data, what is the entropy of the training set? What would be information gain, if we split on (i) Weather (ii) Temperature and (iii) Humidity? Assume “Play?” as the class label.

OR

Q 6 Assume the data given in the following table as the training [10]
CLO2
BL3,5 set. Assume “Gender” as the class attribute.

Gender	Weight	Foot Size
Male	180	12
Male	190	11
Male	170	12
Male	165	10
Female	100	6
Female	150	8
Female	130	7
Female	150	9

Use naïve Bayes classifier to classify a person with Weight = 155 and Foot Size = 9. Fit Gaussian distribution to the data. Units of the variables are ignored for the analysis.