INDIAN INSTITUTE OF TECHNOLOGY (ISM) DHANBAD

Mid-Term Examination

Machine Learning: MSD527

(Academic Year 2021-22)

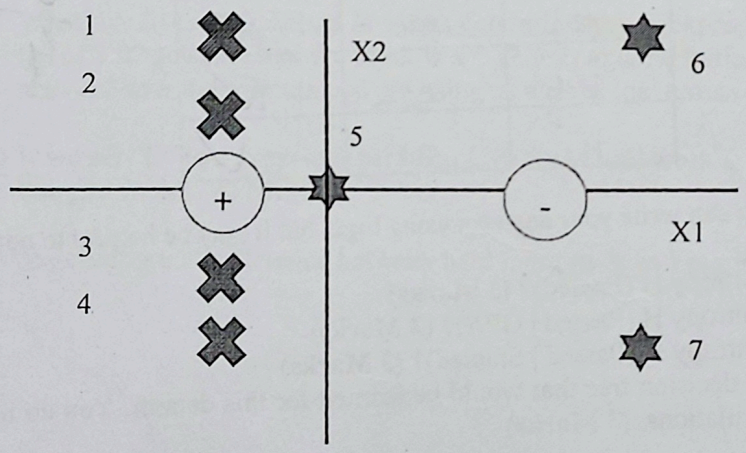
Course: Machine Learning **Max Marks: 40**

Date: 24/3/22 **Duration: 1 hours**

Instructions: Scientific Calculators are allowed

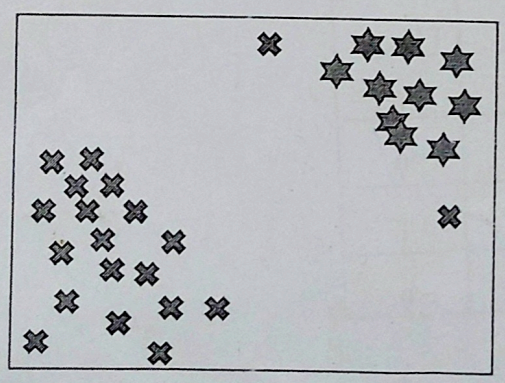
Q1. Consider below given training data:

a.) Draw the decision boundary obtained by the linear hard margin SVM method with a thick solid line. Draw the margins on either side with thinner dashed lines. Circle the support vectors. **(3 Marks)**



b.) The removal of which sample will change the decision boundary? **(Marks: 1)**

c.) Assume that we are training an SVM with a quadratic kernel and kernel function is a polynomial kernel of degree 2. This means the resulting decision boundary in the original feature space may be parabolic in nature. The dataset on which we are training is given below:



Where would the decision boundary be for very large values of C? (Remember that we are using a quadratic kernel). Justify your answer in one sentence and then draw the decision boundary in the above figure. **(3 Marks)**

d.) Where would the decision boundary be for C nearly equal to 0? Justify your answer in one sentence and then draw the decision boundary in the above figure. **(3 Marks)**

Q2. Consider the dataset given below to learn a decision tree which predicts if people pass machine learning (Yes or No), based on their previous GPA (High; H, Medium; M, or Low; L) and whether or not they studied.

GPA	Studied	Passed
L	F	F
L	T	T
M	F	F
M	T	T
H	F	T
H	T	T

For this problem, you can write your answers using log2, but it may be helpful to note that log2 $3 \approx 1.6$.

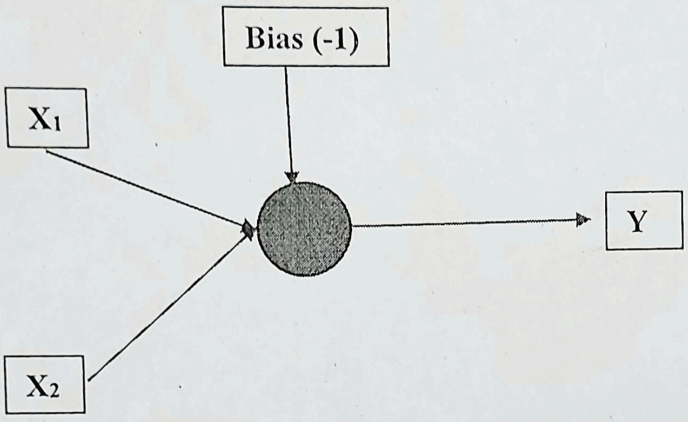
a.) What is the entropy H (Passed)? **(3 Marks)**
b.) What is the entropy H (Passed | GPA)? **(3 Marks)**
c.) What is the entropy H (Passed | Studied)? **(3 Marks)**
d.) Draw the full decision tree that would be learned for this dataset. You do not need to show any calculations. **(3 Marks)**

Q3. Answer the following questions:

a.) Can you represent the following Boolean function with a single logistic threshold unit (i.e., a single unit from a neural network)? If yes, show the weights. If not, explain why not in 1-2 sentences. **(3 Marks)**

A	B	F(A, B)
1	1	0
0	0	0
1	0	1
0	1	0

b.) Assume we have a single sigmoid unit and the weights are w0 = 0, w1 = 1, w2 = 1. What is the computed 'y' value for each of the points on the diagram below? Data points are: Negative: (1, 0) (2, 2) Positive: (1, 0) Recall that for neural nets, the negative class is represented by a desired output of 0 and the positive class by a desired output of 1. **(3 Marks)**



c.) What would be the change in w2 as determined by backpropagation using a step size (η) of 1.0? Assume that the input is x = (2, −2) and the initial weights are as specified above. Show the formula you are using as well as the numerical result. **(8 Marks)**

d.) If we set "C" to a large number, then how SVM will work & whether it will affect the margins or not? **(2 Marks)**

e.) State important difference between hard margins & soft margins? **(2 marks)**