Birla Institute of Technology & Science, Pilani
Work Integrated Learning Programmes Division
Second Semester 2018-19
M.Tech. (Data Science and Engineering)
Comprehensive Examination

SOLUTIONS

Course No.	: DSECL ZG565
Course Title	: MACHINE LEARNING
Nature of Exam	: Open Book
Weightage	: 40%
Duration	: 2 Hours 30 Minutes
Date of Exam:	November 10, 2019

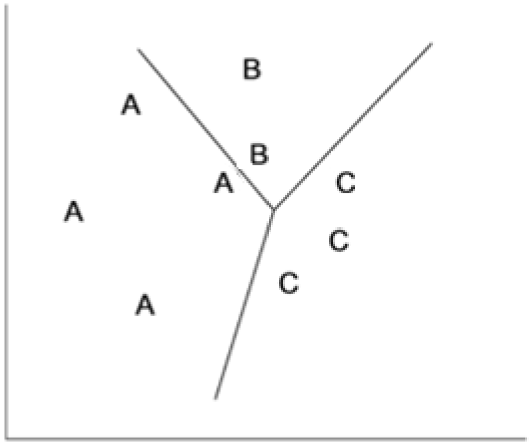
Time of Exam: 10:00 AM – 12:30 PM

No. of Pages = 3
No. of Questions = 6

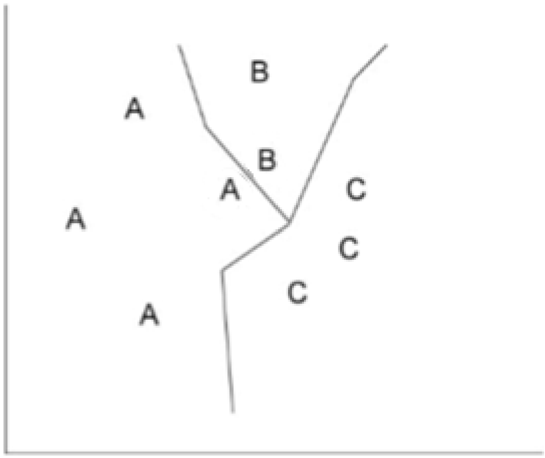
Note:
1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

Question 1. Draw the decision boundary generated by SVM and 1 nearest neighbour classifiers. [2+2=4]

[to be answered on page 3-4]



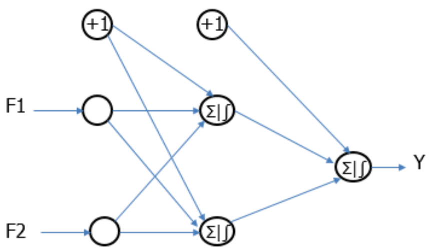
Support Vector Machine



1 Nearest Neighbour

SVM decision boundaries correspond to maximum margin lines between support vectors of each of the two classes. 1 Nearest Neighbour decision boundary corresponds to the set of perpendicular bisectors of nearest data points belonging to different classes.

Question 2. Consider the neural network model with exp( ) activation function to classify the training data given in the adjoining table. [6+1.5 = 7.5] **[to be answered on page 5-7]**



F1	F2	Y
1	2	0
1	1	+1
2	1	0
2	2	+1

(a) Obtain the node weights
(b) Bias weights to classify the training data correctly

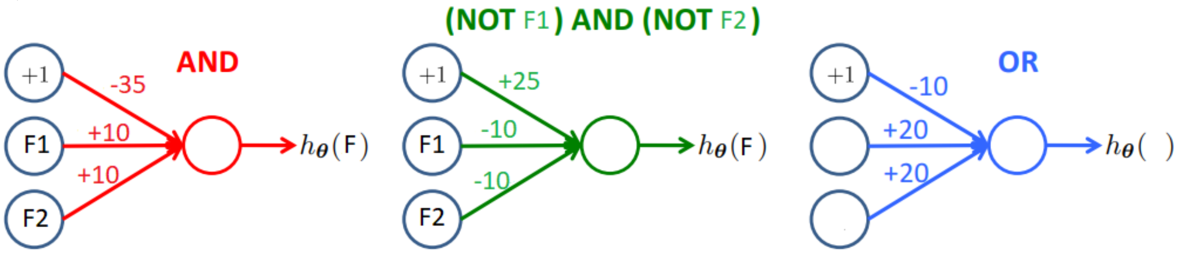
Solution with either sigmoid or exp activation problem is acceptable.
Note: There can be other combinations of weights that will give correct classification. If output >=0.5, class is denoted by +1, else it is class 0.

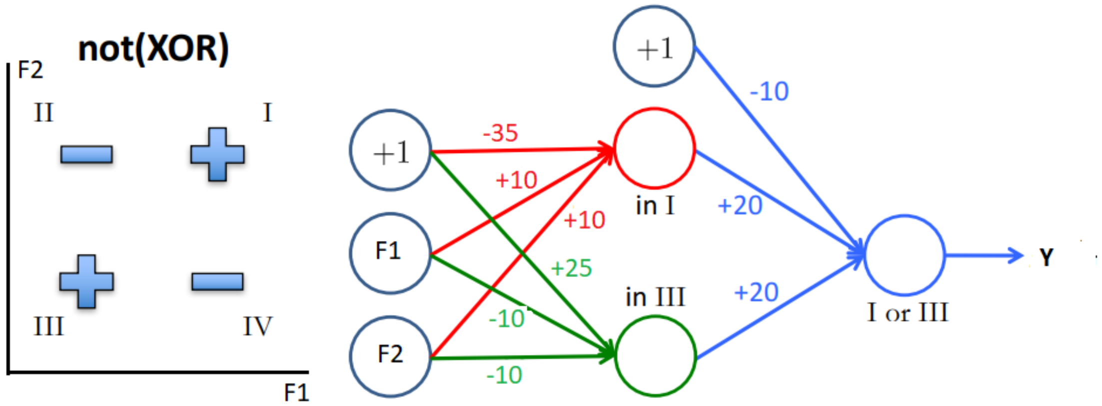
The problem qualitatively maps to not(XOR) problem, which can be expressed as OR of (f1 AND f2) & (NOT(f1) AND NOT(f2)). [f1 = F1 -1, f2 = F2-1]

For sigmoid activation function output = 1/(1+exp(-input)), and output --> +1 for input --> large + and output --> 0 for input --> large - .

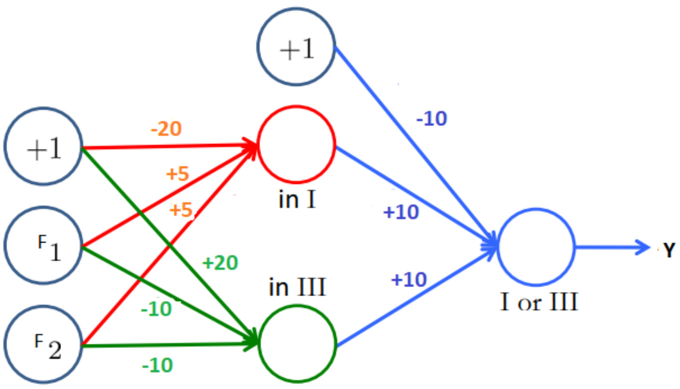
For F1 AND F2, assume weights from input F1/F2 to hidden nodes = w and bias weight = b. Then b+4w >> 0, b+2w <<0, b+3w << 0. So, -4w < b < -3w. Thus (w,b) can be chosen as (10, -35).

Similarly, for NOT(F1) AND NOT(F2), (w,b) can be chosen as (-10,+25), and X1 OR X2, (w,b) = (20, -10).





For exp activation function output = exp(input), output = +1 for input --> 0 and output = 0 for input << 0 .



Question 3. Answer the following questions. [3+3+2=8] **[to be answered on page 8-10]**

Feature probability distribution of two classes of data points are given by

$$P\ (x, y\ |class_0) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \frac{-(x-1)^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}$$

and

$$P\ (x, y\ |class_1) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \frac{-x^2}{2\sigma_y^2} - \frac{(y-2)^2}{2\sigma_x^2}$$

Both classes are equally likely.

a) Derive the equation of decision boundary.

$$\frac{(x-1)^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} = \frac{x^2}{\sigma_y^2} + \frac{(y-2)^2}{\sigma_x^2}$$

b) What is the equation and shape of the boundary if $\sigma_x = \sigma_y = \sigma$? Draw the distributions and mark decision boundary.

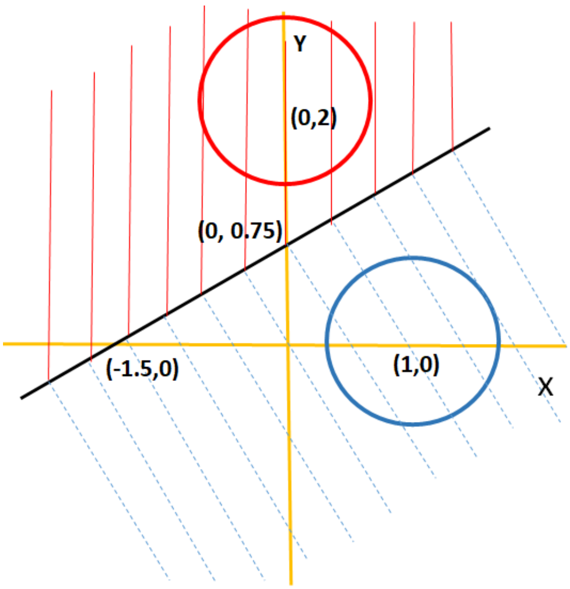
$$\frac{(x-1)^2}{\sigma^2} + \frac{y^2}{\sigma^2} = \frac{x^2}{\sigma^2} + \frac{(y-2)^2}{\sigma^2}$$

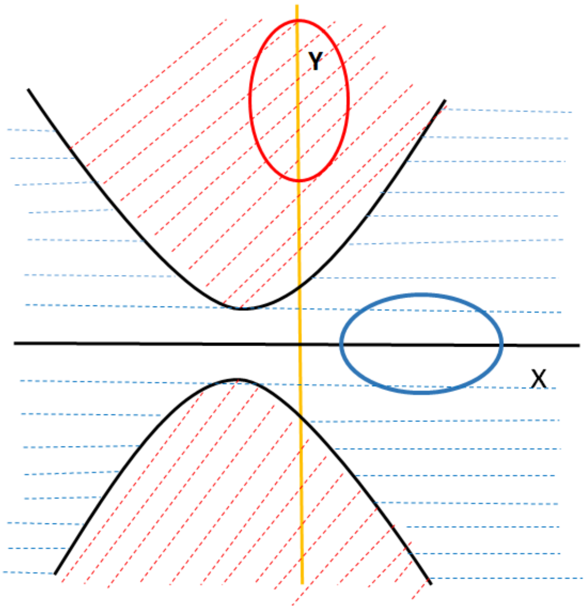
Or, -2x+1= -4y+4

Or, 2y – x = 1.5

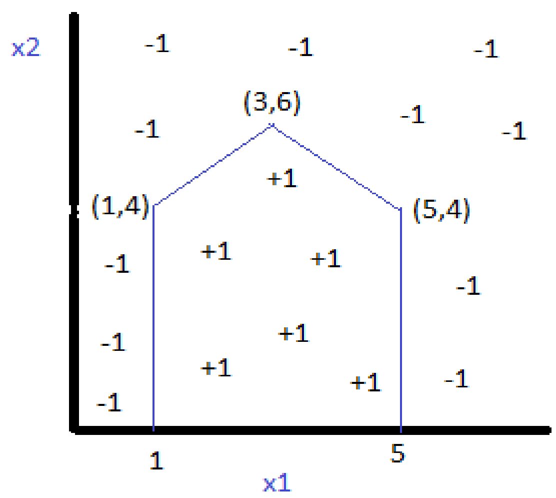
c) What is the shape of the boundary if $\sigma_x \neq \sigma_y$? Draw the distributions and mark decision boundary.

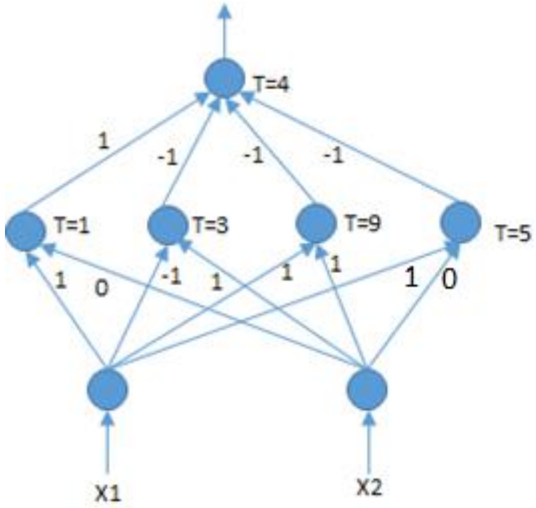
Shape of the decision boundary is quadratic, specifically hyperbolic, oriented along y dimension.





Question 4. The decision boundary for a two-class +1/-1 problem is given below in left. Specify the parameters weights and bias of the perception network on the right that can generate the given decision boundary. Note for a perceptron node, the output is +1 for input greater than or equal to 0, else output is -1. [6 + 2 = 8] **[to be answered on page 11-13]**





The decision boundary consists of the following four straight lines: x1=1, -x1+x2=3, x1+x2=9 and x1=5. The four hidden nodes weights and biases correspond to the coefficients of these four lines, respectively.

Now points labelled as positive class fall on the right side of x1=1, below the lines -x1+x2=3, x1+x2=9 and left side of x1=5. So corresponding weights from hidden nodes to output can be obtained as 1, -1, -1 and -1, respectively.

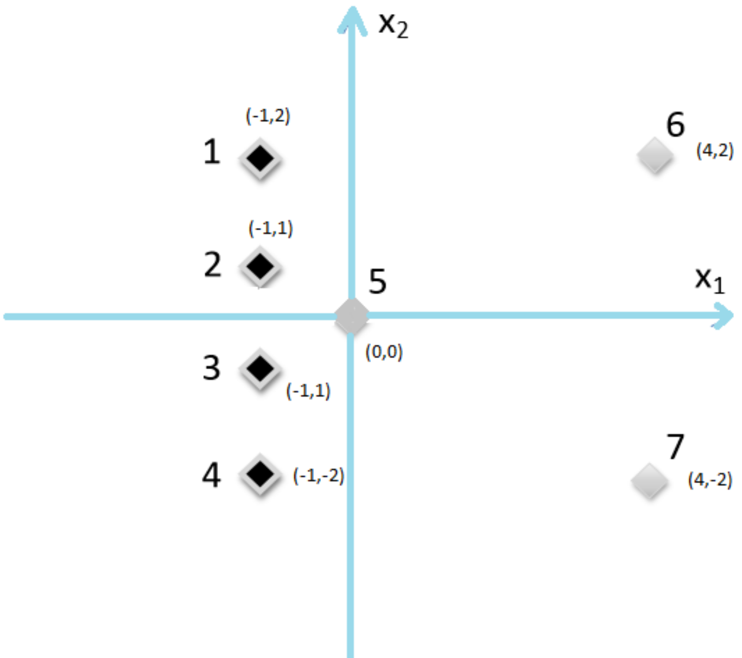
For points labelled as positive class +, total input to the output node is 4, whereas for points labelled as negative class -, total input to output node is less than 4. So bias at the output node is obtained as 4.

Note: Correct classification can be obtained for other combinations of weights and bias T also.

Question 5. Consider a 2-class classification problem in a 2 dimensional feature space **x**=[x1, x2] with target variable y=±1.The training data comprises 7 samples as shown in Figure on next page. 4 black diamonds denote the positive class and 3 gray diamonds for the negative class. [2+1+1+2.5 = 6.5]
[to be answered on page 14-16]

Note: Leave-one-out cross validation is performed by training a classifier with (n-1) training samples and testing on the remaining one sample. This process is repeated for every sample in the training set. Error rate is the ratio of number of wrongly classified samples and total number of samples.

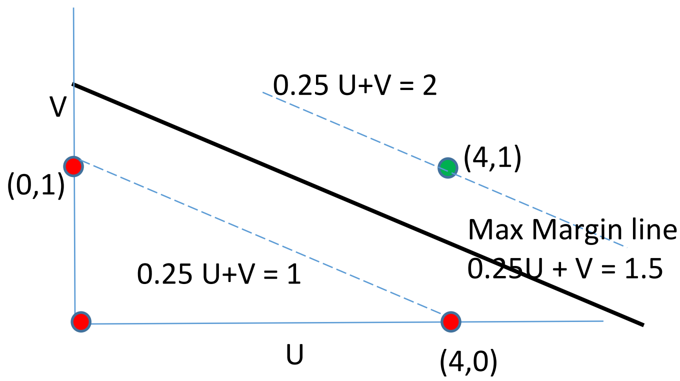
- **a)** What is the equation of the decision boundary if SVM is used? **X1= - 0.5**
- **b)** What is the training error rate? **0**
- c) The removal of which sample will change the decision boundary? Sample 5 - **(0,0).** If this sample is removed, decision boundary will be changed to X1 = 1.5
- **d)** What is the leave-one-out cross validation error rate? **1/7**



Note: There is a typo. Sample 2 and 3 coordinates are given to be same. But it does not have a bearing on the correct answer.

Question 6. Design a maximum margin SVM classifier for the following dataset. Draw the decision boundary along with the data points in (F1, F2) space. [5+1=6] **[to be answered on page 17-19]**

F1	F2	Class	$U=F1^2$	$V=(F2-1)^2$
0	1	0	0	0
2	1	0	4	0
2	2	1	4	1
0	2	0	0	1
-2	1	0	4	0
-2	2	1	4	1
-2	0	1	4	1
0	0	0	0	1
2	0	1	4	1



Maximum margin hyperplane in (F1,F2) space: $0.25\ F1^2 + (F2\text{-}1)^2 = 1.5$ which is an ellipse centered at (0,1) and major axis aligned to F1.