BITS PILANI, DUBAI CAMPUS
DUBAI INTERNATIONAL ACADEMIC CITY, DUBAI

FIRST SEMESTER 2023 – 2024
COURSE: BITS F464 (Machine Learning)
COMPONENT: Practice Tutorial 4 **DATE:** 20 December 2023

Q1.

Consider a following plotted data points in Figure 1,

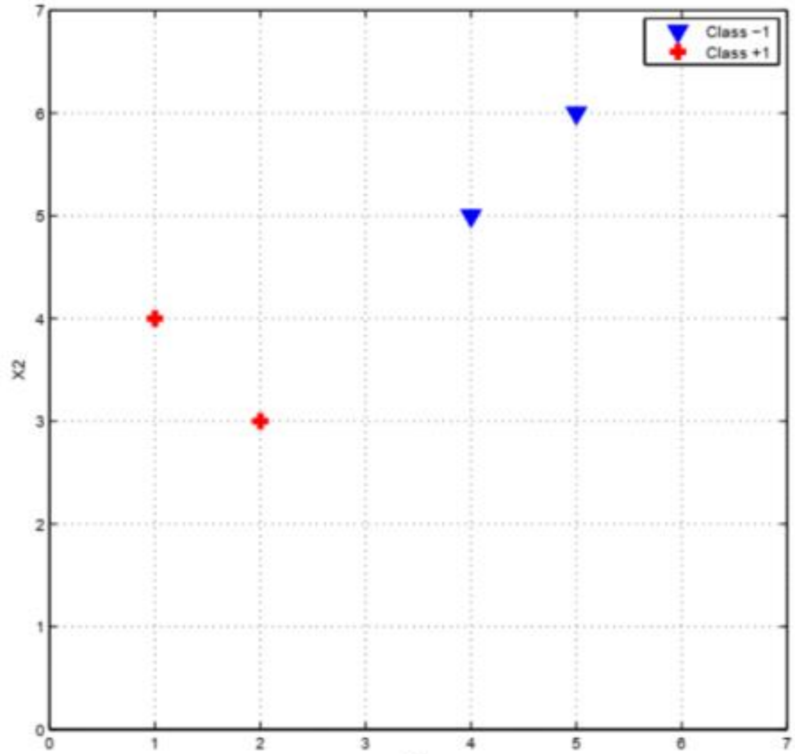


Figure 1 Q1

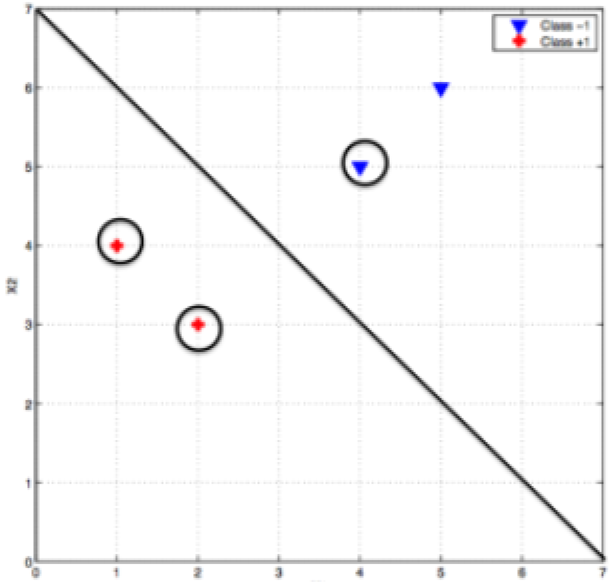
Find the support vectors and draw the decision boundry which will discriminate the given points into two classes. (Show step by step process of calculating a hyperplane)

Answer:

Support vector => (4,5), (2,3)

w1 = -1/2, w2 = -1/2, b = 7/2

(Refer class notes to know procedure to find it)

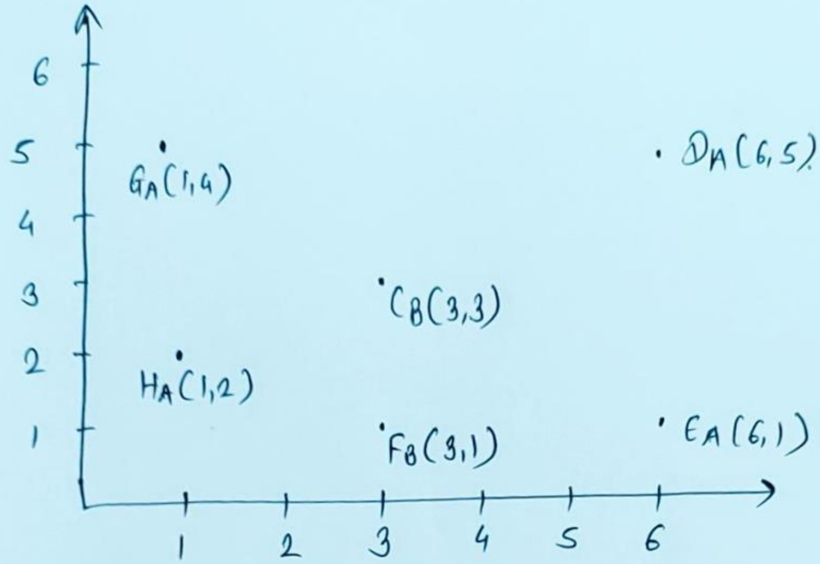


Q2.

Given a following datapoints and decision stumps. Find out the strong classifier (after two iterations) which will classify the data points into two classes.

Class A data points (X, Y)	*G (1,4), H (1,2), D (6,5), E (6,1)*
Class B data points (X, Y)	*C (3,3), F (3,1)*
Decision stumps	*X<2, X<4, X>7*

Q. 2 ⇒



Iteration 1 ⇒

Initial Weights ⇒ $W_g = W_h = W_c = W_d = W_e = W_f = 1/6$

Decision Stump	Wrong	Error
$X < 2$	D, E	$2/6 = 1/3$
$X < 4$	D, E, C, F	$4/6 = 2/3$
$X > 7$	G, H, D, E	$4/6 = 2/3$

$X < 2$ is the best performing classifier, now decide voting power of the classifier.

$$\epsilon = 1/3 \Rightarrow \alpha = 1/2 \log \frac{1-\epsilon}{\epsilon} = 1/2 \log \frac{(1-1/3)}{1/3}$$

$$= 1/2 \log 2$$

$$\boxed{h(x) = 1/2 \log 2 * F(X<2)}$$

Decide new weights ⇒

G, H, C, F are correctly classified points.

$$\therefore w_{new} = \frac{w_{old}}{2(1-\epsilon)} = \frac{1/6}{2(1-1/3)} = 1/8 \quad \therefore \quad \boxed{W_g = W_h = W_c = W_f = 1/8}$$

D, E are wrongly classified.

$\therefore w_{new} = \frac{w_{old}}{2\epsilon} = \frac{1/6}{2 \times 1/3} = \frac{1}{4}$

$\boxed{w_d = w_e = 1/4}$

Iteration 2 $\Rightarrow$

Decision Stump	Wrong	Error
$x < 2$	D, E	$2/4 = 1/2$
$x < 4$	D, E, C, F	$3/4$
$x > 7$	G, H, D, E	$3/4$

$x < 2$ is the best performing classifier, now decide voting power of the classifier.

$\epsilon = 1/2 \Rightarrow \alpha = 1/2 \log \frac{1 - 1/2}{1/2} = 1/2 \log 1$

$$\boxed{h(x) = 1/2 \log 2 * F(x<2) + 1/2 \log 1 \, F(x<2)}$$

Q3.

Consider a neural network with linear activation as shown in Figure 2. The output of each unit is a constant *C* multiplied by the weighted sum of inputs.

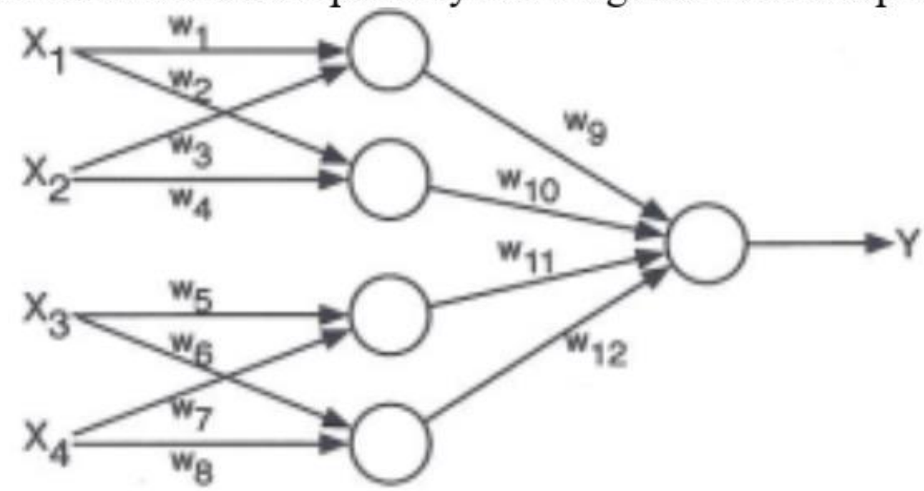
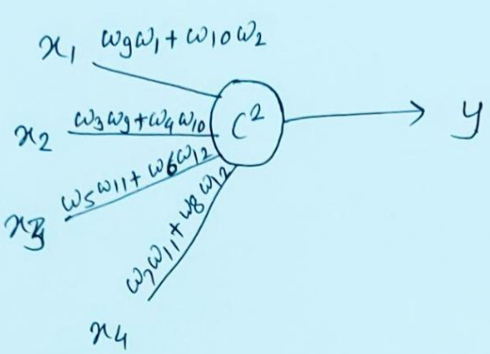


Figure 2 Q. A5

i. Represent the above network using single unit network, including the weights and activation function.

ii. Can it be possible to represent the above network using linear regression? If yes, write the equation of Y. If no, explain why?

(i)



(ii) Yes,

One of the indicative solution is,

$$Y = C^2(\omega_9\omega_1 + \omega_{10}\omega_2)x_1 + C^2(\omega_3\omega_9 + \omega_4\omega_{10})x_2 + C^2(\omega_5\omega_{11} + \omega_6\omega_{12})x_3 + C^2(\omega_7\omega_{11} + \omega_8\omega_{12})x_4$$

Q4.

Suppose we have the following training set of positive (+) and negative (-) instances and a single test instance (*o*). All instances are projected onto a vector space of two real-valued features (X and Y) as shown in Figure 2. Answer the following questions.

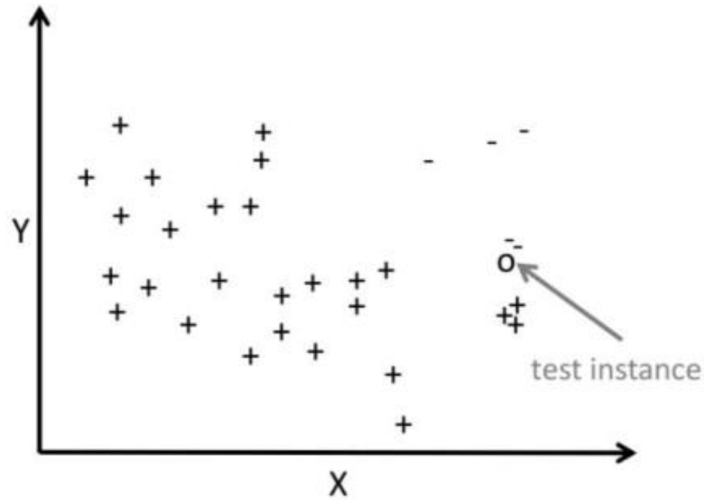


Figure 2 Q4

a. What would be the class assigned to this test instance for $K=3$?
b. Can setting $K=11$ is good for a given dataset? Give correct justification for your answer.

a. For K=3, this test instance would be predicted negative. Out of its three nearest neighbors, two are negative and one is positive.

b. There are only 5 negative instances in the training set. Therefore, any value of K > 10 would have a majority of positive instances

Q5. Calculate the principal components for following dataset. (Show all steps of calculations)

X1 = (x1, x2) = {(1,4), (3,7), (2,4)}

X2 = (x1, x2) = {(11,10), (9,11), (8,9)}

- ⇨ **Use class room notes ML_PCA (Same example as notes with different values)**