1.3 a) ~~show~~ $y(t) * w^T(t) * x(t) < 0$

Hint: $x(t)$ is misclassified by $w(t)$

$\therefore$ since $x(t)$ is misclassified, that means.

$$sign(w^T(t) \cdot x(t)) = -1 \quad \& \quad y(t) = +1$$

and $sign(w^T(t) \cdot x(t)) = +1 \quad \& \quad y(t) = -1$

For $sign(w^T(t) \cdot x(t)$ to be -1,

we can refer $sign(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$

$\therefore sign(x) = -1 \therefore x < 0.$

$\therefore sign(w^T(t) \cdot x(t)) = -1$

which implies, $\boxed{w^T(t) \cdot x(t) < 0.}$

$\therefore$ As we already have $y(t) = +1$ for $sign(w^T(t) \cdot x(t)) = -1$

Then $y(t) * w^T(t) \cdot x(t) = (+1) * $ (sum number less than 0)

$\Rightarrow (+1) * (m < 0)$ where m = some number $< zero$.

$\therefore (+1) * (m < 0) ==$ output < 0.

$\boxed{\therefore \text{Hence proved, } y(t)\, w^T(t)\, x(t) < 0}$ same for the other case

(b) show that $y(t)\, w^T(t+1)\, x(t) > y(t) w^T(t)\, x(t)$

Hint : Use (1.3).

$\therefore$ It is said that with $w(t+1)$ our $x(t)$ will be correctly classified.

$\therefore$ $y(t) = +1$ then ~~sign~~ sign $(w^T(t+1)\, x(t)) = +1$
as it is $(x(t))$ correctly classified.

$\therefore$ for sign $(w^T(t+1)\, x(t))$ to be $= +1$

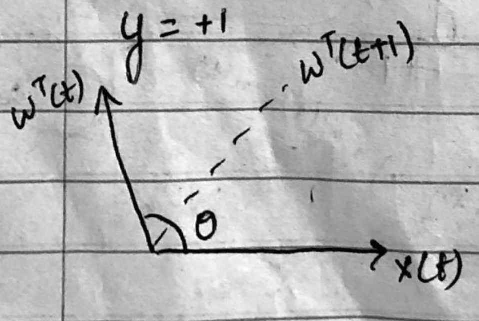
$w^T(t+1)\, x(t) > 0$. from $\text{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$

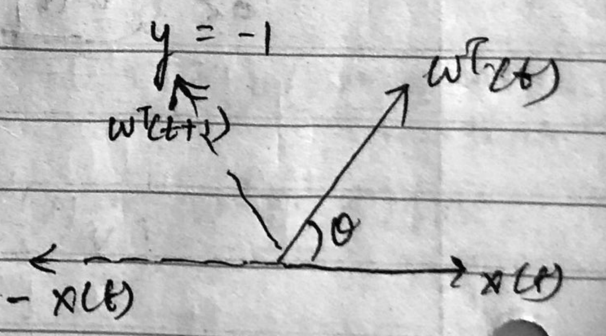
$\therefore$ As $y(t) = +1$ $\therefore$ $\boxed{y(t) * w^T(t+1) * x(t) > 0}$

As we have proved if 1.3(a) $y(t)\, w^T(t)\, x(t) < 0$

already which means,

$$\boxed{y(t) * w^T(t+1) * x(t) > y(t)\, w^T(t)\, x(t)}$$





Problem 1.6

a) Recommending a book to a user in an online bookstore :-

→ Supervised learning → As we have to have labelled purchase history of user to recommend the book.

→ we know what user has purchased in past & hence for every feature set 'x' we have 'y' as output whether "purchased" or not.

∴ Training data set :-

book author, no. of pages, type of book like its genre, ebook or hardcopy, graphical content or not, etc.

b) Playing a tic-tac-toe :- (Reinforcement learning)

- We do not have output of a player's win|loss given few steps.

→ The model has to learn whether the performed set of steps lead to a winning output or losing one, hence reinforcement learning.

Training data set :- positions of values for 'x' or 'o'
i.e first_box, second_box, third_box, fourth_box, five_box, sixth_box, seventh_box,

eigth_box, ninth_box, haveWon = {0, 1} = 0 for winning and '1' for losing.

each box will have {'x', 'o', winningStep} where winningStep = boolean i.e this step performed wining step or losing one.

c] Categorizing movies into different types:-

· It can be supervised or Unsupervised based on the data we have,

∴ Supervised → If we have attributes in dataset
↳ for which we have 'movie type' as output attribute for each input vector.

Unsupervised if,
↳ if do not have labelled output data for movie category then this becomes clustering problem which will be Unsupervised then.

dataset:-
→ movie_id, isAdult, isHorror, is Action, isRomantic, isDrama, isComedy, budget, actor, director, actress, length_of_movie, rating, language, runtime etc.

d) Learning to play music :-

Reinforcement Learning → As we have to learn by checking if output is proper music adhering to music notes or how.

Dataset :- Bandwidth, frequency, bass, contrast, audio attributes like Chroma, Tonnetz, roll off etc.

e) Credit limit : Deciding the maximum allowed debt for each bank customer :-

→ Supervised :- We have to have history of the customers the bank has & based on that only we can predict the future customer's debt.

dataset :- Age, Income, members in family, education, number of creditcards, creditscore, no. of loans, asset value etc.

Exercise 1.8 :-

Given, $\mu = 0.9 \Rightarrow$ probability of red marbles in bin.

$\therefore$ $p(red) = 0.9$

let's say $N \Rightarrow$ No of samples $= 10$

To find: $p(\nu \leq 0.1) \Rightarrow \nu = \frac{\text{No. of red}}{\text{Total sample}} = \frac{red}{N}$

$\therefore$ $p(\text{marble} \,!= \text{red}) = 1 - 0.9 = 0.1$

$p(\text{marble} = \text{red}) = 0.9$

$\therefore$ $p(\nu \leq 0.1) = p\left(\frac{red}{N} \leq 0.1\right) = p(red \leq 0.1 \times N)$

$= p(red \leq 1) = p(red = 0) + p(red = 1)$

$\therefore$ $p(red = 0) =$ each sample has probability of marble being not red.

$\therefore$ $P(\underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}$

$\Rightarrow (0.1)^{10} = \underline{1 \times 10^{-10}}$

Now, $P(red = 1) = \underline{0.9}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}$

plus, $\underline{0.1}\ \underline{0.9}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}\ \underline{0.1}$.

$\Rightarrow (0.9)(10)(0.1)^9$

$\Rightarrow 9 \times 10^{-9}$.

$\therefore$ $p(\nu \leq 0.1) = p(red = 0) + p(red = 1)$

$= 1 \times 10^{-10} + 9 \times 10^{-9}$

$= 0.1 \times 10^{-9} + 9 \times 10^{-9} = \underline{\underline{9.1 \times 10^{-9}}}$

Scanned with CamScanner

Exercise 1-9

$\mu = 0.9$. We have to bound $p(\nu \leq 0.1)$.

$\therefore$ By Hoeffding inequality,

$$p\left[|\nu - \mu| > \varepsilon\right] \leq 2e^{-2\varepsilon^2 N} \quad \text{For any } \epsilon > 0$$

$\therefore \quad |\nu - \mu| > \varepsilon$

$\therefore$ Squaring 2 sides, $(\nu - \mu)^2 > (\varepsilon)^2$

$$(0.1 - 0.9)^2 > (\varepsilon)^2$$

$$\Rightarrow (0.8)^2 > (\varepsilon)^2$$

So we can say for $\leq 2e^{-2\varepsilon^2 N}$, $\varepsilon = 0.8$.

$$\Rightarrow \therefore P\left[|\nu - \mu| > \varepsilon\right] \leq 2e^{-2(0.8)^2(10)}$$

$$\leq (2)\, e^{-2(6.4)}$$

$$\leq 5.52 \times 10^{-6}$$

$$\Rightarrow p(\nu \leq 0.1) < p\left[|\nu - \mu| > \varepsilon\right] \leq 5.52 \times 10^{-6}$$

$$\Rightarrow \cancel{\text{9...}}\ 9.1 \times 10^{-9} \leq p\left[|\nu - \mu| > \varepsilon\right]$$

$$\leq 5.52 \times 10^{-6}$$

Problem 1.2

$h(x) = \text{sign}(w^T x)$

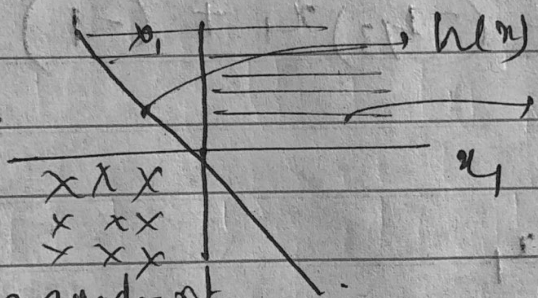
where $w = [w_0, w_1, w_2]^T$

$x = [1, x_1, x_2]^T$

As we know, $\text{sign}(x) = \begin{cases} 1 & \text{where } x > 0 \\ 0 & x = 0 \\ -1 & x < 1 \end{cases}$

$\therefore$ $h(x) = +1$ which means points are in +ve quadrant.

$h(x) = -1$ which means points are in -ve quadrants.

$\therefore$ 

$h(x)$ where +ve quadrant.

-ve quadrant.

Also, $h(x) = \text{sign}(w^T x) = \text{sign}(w_0 x_0 + w_1 x_1 + w_2 x_2)$

when $w^T x = 0$ we say all points are not line (1) & $\vec{w}$ is perpendicular to $\vec{x}$.

(a)

$\therefore$ $w_0 x_0$ = constant, we can express $w_0 x_0 = \cancel{k}$ ~~or~~ C some constant,

$\therefore$ $x_2 = (-w_1)/(w_2)^x + (-w_0 x_0)$

$= (-w_1/w_2)\, x + C$

This is a equation of line.

Now, we have given this, is, $x_2 = ax_1 + b$

where a = slope & b = intercept

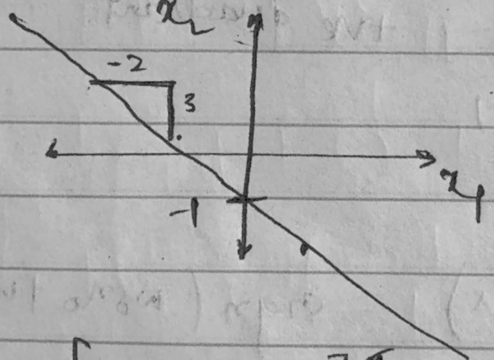
$$\therefore\ x_2 = -w_1/w_2\, x_1 + \left(-w_0 x_0 / w_2\right)$$

Ans, part a)

$$\therefore\ \boxed{\begin{aligned} a &= -w_1/w_2 \\ b &= -w_0/w_2 \text{ as } x_0 = 1 \end{aligned}}$$

part (b) ∴ 1) $w = [1, 2, 3]^T$

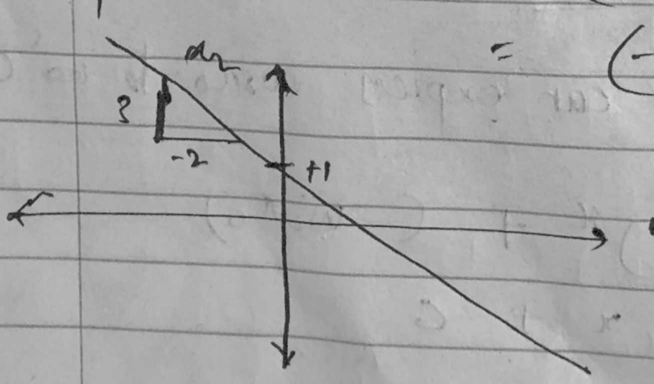
∴ equation → $x_2 = (-2/3)\, x_1 + (-1)$



2) $w = [-1, -2, -3]^T$

∴ equation → $x_2 = \left(-(-2)/-3\right) x_1 + \left(-(-1)\right)$

$= (-2/3)\, x_1 + (+1)$



problem 1.11:-

a] Supermarket: $\therefore$ matrix $\Rightarrow$

$$\begin{array}{c|cc} & & f \\ h & +1 & -1 \\ \hline +1 & 0 & 1 \\ -1 & 10 & 0 \end{array}$$

$\therefore \quad E_{in}(g) = \frac{1}{N} \sum_{n=1}^{N} e(h(x_n), f(x_n))$

$\therefore \quad E_{in}(g) = \frac{1}{N} \sum_{n=1}^{N} [y_n \neq g(x_n)]$

$$\boxed{E_{in}(g) = \frac{1}{N}\left[\left(\sum_{y_n=+1}(g(x_n) \neq +1)\right)\times 10 + \left(\sum_{y_n=-1}(g(x_n) \neq -1)\right)\times 1\right]}$$

b] for CIA: matrix $\Rightarrow$

$$\begin{array}{c|cc} & & f \\ h & +1 & -1 \\ \hline +1 & 0 & 1000 \\ -1 & 1 & 0 \end{array}$$

$\therefore \quad E_{in}(g) = \frac{1}{N} \sum_{n=1}^{N} [y_{(n)} \neq g(x_n)]$

$$\boxed{E_{in}(g) = \frac{1}{N}\left[\left(\sum_{y_{(n)}=+1}(g(x_n) \neq +1)\right)\times 1 + \left(\sum_{y_{(n)}=-1}(g(x_n) \neq -1)\right)\times 1000\right]}$$