University of Asia Pacific

Department of CSE

Name: Rashik Rahman

Reg ID: 17201012

Year: 4th

Semester: 1st

Course Code: CSE 427

Course Title: Machine Learning

Date: 27.04.2021

"During Examination and upload time I will not take any help from anyone. I will give my exam all by myself."

University of Asia Pacific

Admit Card

Final-Term Examination of Fall, 2020

Financial Clearance	PAID

Registration No : 17201012

Student Name : Rashik Rahman

Program : Bachelor of Science in Computer Science and Engineering

SI.NO.	COURSE CODE	COURSE TITLE	CR.HR.	EXAM. SCHEDULE
1	CSE 400	Project / Thesis	3.00	
2	CSE 330	Industrial Training	1.50	
3	CSE 401	Mathematics for computer Science	3.00	
4	CSE 403	Artificial Intelligence and Expert Systems	3.00	
5	CSE 404	Artificial Intelligence and Expert Systems Lab	1.50	
6	CSE 405	Operating Systems	3.00	
7	CSE 406	Operating Systems Lab	1.50	
8	CSE 407	ICTLaw, Policy and Ethics	2.00	
9	CSE 410	Software Development	1.50	
10	CSE 427	Topics of Current Interest	3.00	

Total Credit: 23.00

1. Examinees are not allowed to enter the examination hall after 30 minutes of commencement of examination for mid semester examinations and 60 minutes for semester final examinations.

2. No examinees shall be allowed to submit their answer scripts before 50% of the allocated time of examination has elapsed.

3. No examinees would be allowed to go to washroom within the first 60 minutes of final examinations.

4. No student will be allowed to carry any books, bags, extra paper or cellular phone or objectionable items/incriminating paper in the examination hall.
Violators will be subjects to disciplinary action.

This is a system generated Admit Card. No signature is required.

Admit Card Generation Time: 25-Apr-2021 02:11 AM

17201012

Answer to the Q.No.1(a)

In an artificial neural network, the function which takes the incoming signals as input and produces the output signal is know as the activation function.

Some activation function with expression and diagram are given below:

i) Threshold activation function:

defined by, $f(x) = \begin{cases} 1; & \text{if } x > 0 \\ -1; & \text{if } x \leq 0 \end{cases}$

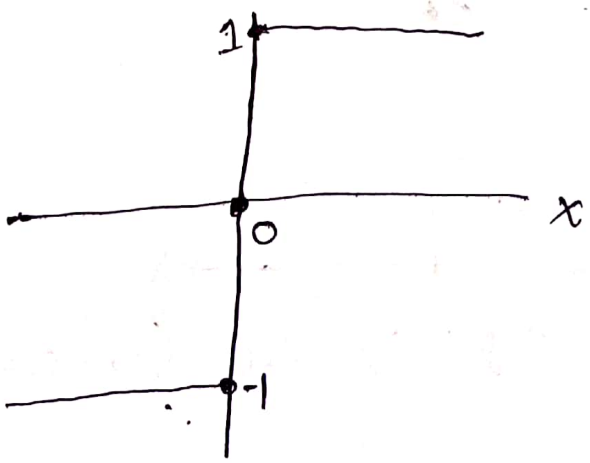


fig: Threshold function

②
17201012

ii) Unit step function:

~~definede~~
defined by, $f(x) = \begin{cases} 1 ; \text{ if } x \geq 0 \\ 0 ; \text{ if } x < 0 \end{cases}$

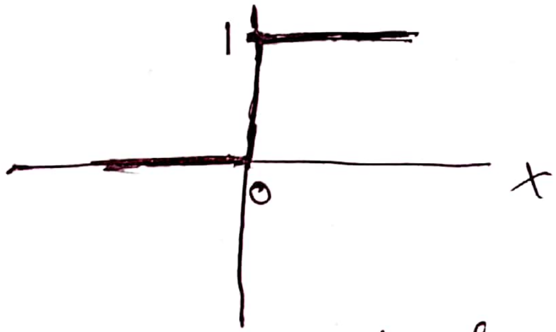


fig: unit step function

iii) Sigmoid activation function:

defined by, $f(x) = \dfrac{1}{1+e^{-x}}$

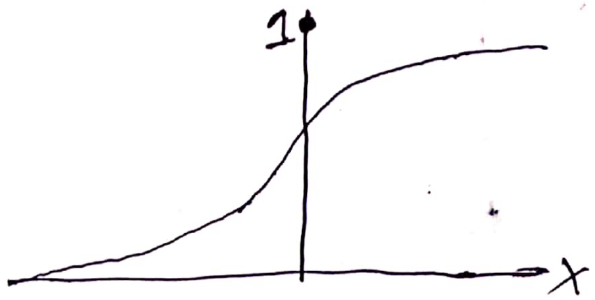


fig: sigmoid function

Scanned with CamScanner

(iv) Linear activation function:

defined by, $f(x) = mx + c$

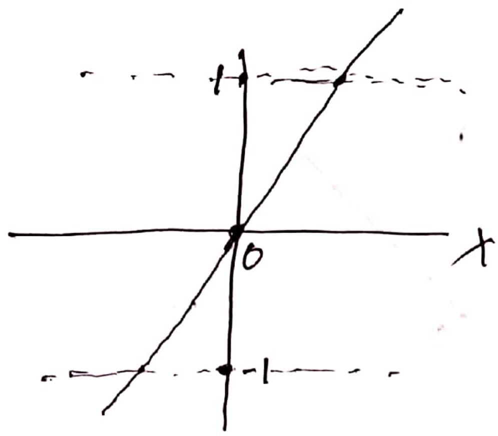


fig: Linear function

(v) Hyper-bolic tangential activation function:

defined by, $f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$

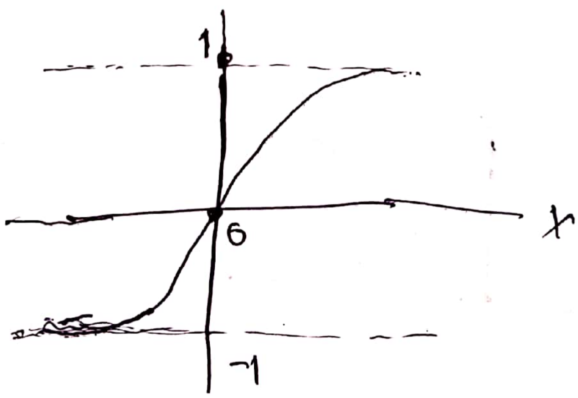


fig. hyperbolic tangential function

Scanned with CamScanner

(4)

17201012

(vi) ReLU activation function:

defined by, $f(x) = \begin{cases} x \, ; & \text{if } x \geq 0 \\ 0 \, ; & \text{if } x < 0 \end{cases}$

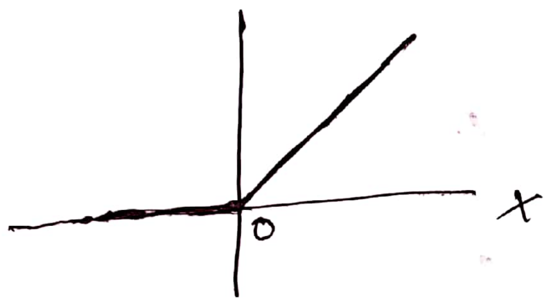


fig. ReLU function

(vii) Leaky ReLU activation function:

defined by, $f(x) = \begin{cases} x \, ; & \text{if } x \geq 0 \\ -0.000001 \, ; & \text{if } x < 0 \end{cases}$

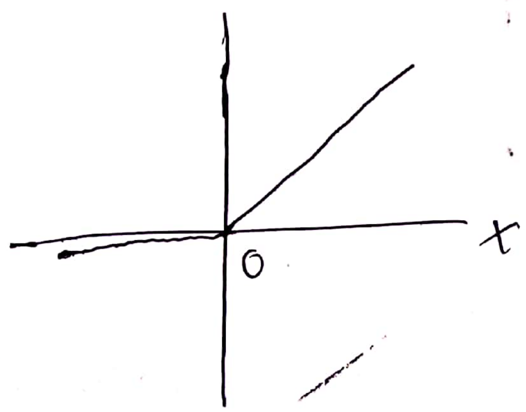


fig. Leaky ReLU function.

17201012

Answer to the Q. No. 1(b)

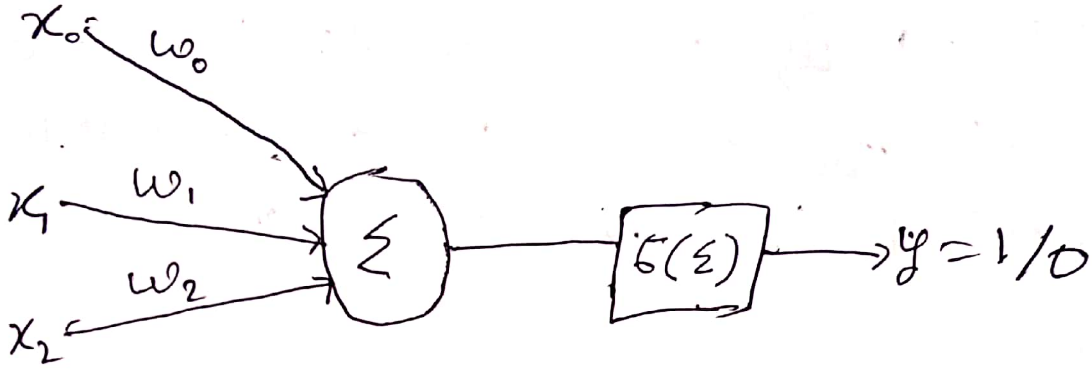


fig: Perceptron model.

AND gate:

A (x_1)	B (x_2)	y
0	0	0
0	1	0
1	0	0
1	1	1

Table: AND gate.

6

1720101R

~~So if we consider the w_i values as the following fig then we can build an AND gate.~~

So if we consider the value of $w_0 = -1.5$, $w_1 = 1$, $w_2 = 1$ and $x_0 = +1$ then we can get the same output as the AND gate output from perceptron model.

$\sigma(x) = \begin{cases} 1; & \text{if } x \geq 0 \\ 0; & \text{if } x < 0 \end{cases}$

for, $x_1 = 0; x_2 = 0$

$$y = \sigma(x_0 w_0 + x_1 w_1 + x_2 w_2)$$
$$= \sigma(1 \times (-1.5) + 0 \times 1 + 0 \times 1)$$
$$= \sigma(-1.5)$$
$$= 0$$

for, $x_1 = 0, x_2 = 1$

$$y = \sigma(x_0 w_0 + x_1 w_1 + x_2 w_2)$$
$$= \sigma(1 \times (-1.5) + 0 \times 1 + 1 \times 1)$$
$$= \sigma(-0.5)$$
$$= 0$$

Scanned with CamScanner

for, $x_1 = 1, x_2 = 0$

$$y = \sigma(x_0 w_0 + x_1 w_1 + x_2 w_2)$$
$$= \sigma(1 \times (-1.5) + 1 \times 1 + 0 \times 1)$$
$$= \sigma(-0.5)$$
$$= 0$$

for, $x_1 = 1, x_2 = 1$

$$y = \sigma(x_0 w_0 + x_1 w_1 + x_2 w_2)$$
$$= \sigma(1 \times (-1.5) + 1 \times 1 + 1 \times 1)$$
$$= \sigma(0.5)$$
$$= 1$$

∴

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

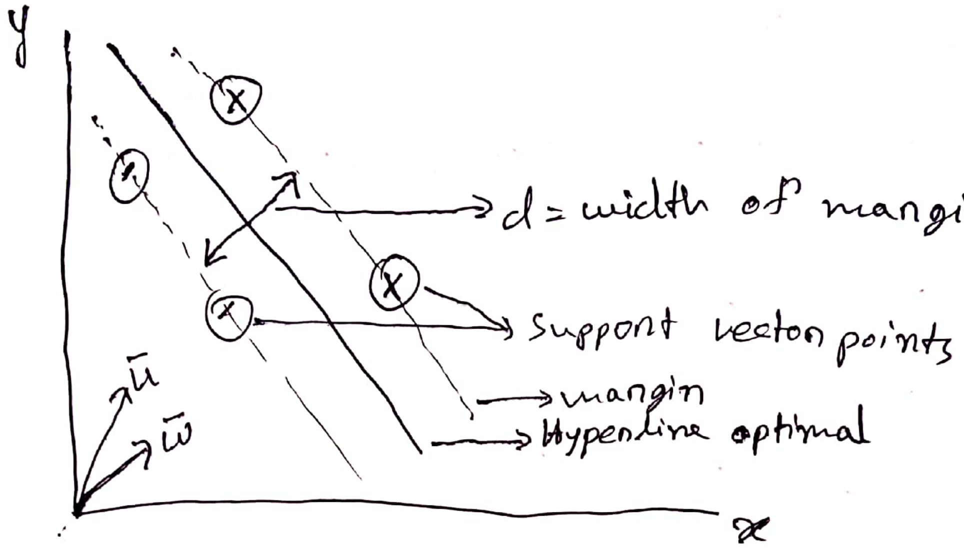
Table: Perceptron I/O.

~~Thus as~~ As both of the table matches we can say we have achieved AND gate through perceptron model.

8

17201012

Answer to the Q.No.3 (a)



$\bar{w}$ = weight vector perpendicular to hyperplane

$\bar{u}$ = unknown poin vector

Goal is to maximize d

$\bar{w}\bar{u} + b > 0$; for any (+) sample

$\bar{w}\bar{u} + b \leq 0$; for any (-) sample.

Equation for support vector points,

$\bar{w}\bar{x}_+ + b > 0$; for any (+) support vector points

$\bar{w}\bar{x}_- + b \leq 0$; for any (-) " " "

Now,

$$\left.\begin{aligned} y_i(\bar{w}.\bar{x}_+ + b) &\geq 1 \\ y_i(\bar{w}.\bar{x}_- + b) &\geq 0 \end{aligned}\right\}$$

∴ for all points (positive/negative)

$$y_i(\bar{w}.\bar{x} + b) - 1 \geq 0$$

$$\Rightarrow y_i(\bar{w}.\bar{x} + b) - 1 = 0$$

Scanned with CamScanner

9

17201012

$\therefore \bar{\omega}\cdot\bar{x}_+ + b \geq 1 \Rightarrow \bar{\omega}\cdot\bar{x}_+ + b = 1 \Rightarrow \boxed{\bar{\omega}\cdot\bar{x}_+ = 1-b}$

$\bar{\omega}\cdot\bar{x}_- + b \leq -1 \Rightarrow \bar{\omega}\cdot\bar{x}_- + b = -1 \Rightarrow \boxed{\bar{\omega}\cdot\bar{x}_- = -1-b}$

Now, maximizing d,

$$d = (\bar{x}_+ - \bar{x}_-)\frac{\bar{\omega}}{\|\bar{\omega}\|}$$

$$= \bar{x}_+ \frac{\bar{\omega}}{\|\bar{\omega}\|} - \bar{x}_- \frac{\bar{\omega}}{\|\bar{\omega}\|}$$

$$= (\bar{x}_+ \bar{\omega} - \bar{x}_- \bar{\omega})\frac{1}{\|\bar{\omega}\|}$$

$$= (1-b+1+b)\frac{1}{\|\bar{\omega}\|}$$

$$= \frac{2}{\|\bar{\omega}\|}$$

$$\therefore \max(d) = \max\left(\frac{2}{\|\bar{\omega}\|}\right)$$

$$\approx \max\left(\frac{1}{\|\bar{\omega}\|}\right)$$

$$\approx \min(\|\bar{\omega}\|)$$

$$\approx \min\left(\frac{1}{2}\|\bar{\omega}\|^2\right)$$

Here, $\|\bar{\omega}\| = \sum_{i=1}^{l} \lambda y_i \bar{x}_i$

1720102

$$\sum_{i=1}^{l} \lambda y_i = 0$$

$$\therefore \min(L) = \frac{1}{2}\|\bar{\omega}\|^2 - \sum_{i=1}^{l} \lambda_i \left(y_i (\bar{\omega} . \bar{x}_i + b) - 1 \right)$$

$$\therefore L = \sum_{i=1}^{l} \lambda_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \lambda_i \lambda_j y_i y_j x_j^T . x_i$$ → dual soln.

Answer to the Q. No. 3(b)

Given,

			class
x_1	2	1	+1
x_2	4	3	-1

Here, $l = 2$

$\bar{x}_1 = (2,1) \rightarrow y_1 = +1$

$\bar{x}_2 = (4,3) \rightarrow y_2 = -1$

$$L(w,b) = \sum_{i=1}^{l} \lambda_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \lambda_i \lambda_j y_i y_j \bar{x}_j^T . x_i$$

$$= (\lambda_1+\lambda_2) - \frac{1}{2}\{\lambda_1\lambda_1 y_1 y_1\, \bar{x}_1\cdot\bar{x}_1 + \lambda_1\lambda_2 y_1 y_2 \bar{x}_1\cdot\bar{x}_2 + \lambda_2\lambda_1 y_2 y_1 \bar{x}_2\cdot\bar{x}_1 + \lambda_2\lambda_2 y_2 y_2 \bar{x}_2\cdot\bar{x}_2)$$

$$= (\lambda_1+\lambda_2) - \frac{1}{2}\{\lambda_1\lambda_1 (+1)(+1)(2\times 2+1\times 1) + \lambda_1\lambda_2(+1)(-1)(2\times 4+1\times 3) + \lambda_2\lambda_1(-1)(+1)(4\times 2+3\times 1) + \lambda_2\lambda_2(-1)(-1)(4\times 4+3\times 3)\}$$

$$L = (\lambda_1+\lambda_2) - \frac{1}{2}\{5\lambda_1^2 - 22\lambda_1\lambda_2 + 25\lambda_2^2\} \longleftarrow (i)$$

We know,

$$\sum_{i=1}^{l} \lambda_i y_i = 0$$

$$\Rightarrow \lambda_1 y_1 + \lambda_2 y_2 = 0$$

$$\Rightarrow \lambda_1 \times 1 + \lambda_2(-1) = 0$$

$$\therefore \lambda_1 = \lambda_2$$

$\therefore$ From (i),

$$L = \lambda_1+\lambda_2 - \frac{1}{2}\{5\lambda_1^2 - 22\lambda_1\lambda_2 + 25\lambda_2^2\}$$

$$= 2\lambda_1 - \frac{1}{2}\{5\lambda_1^2 - 22\lambda_1^2 + 25\lambda_1^2\}$$

$$= 2\lambda_1 - 4\lambda_1^2$$

(12)

17201012

~~$f'(x)=0$~~

$$\frac{\partial}{\partial \lambda_1}(L)=0$$

$$\Rightarrow \frac{\partial}{\partial \lambda_1}(2\lambda_1 - 4\lambda_1^2)=0$$

$$\Rightarrow 2-2\times 4\lambda_1 = 0$$

$$\Rightarrow \lambda_1 = \frac{2}{8} = \frac{1}{4}$$

So, $\lambda_1 = \lambda_2 = \frac{1}{4}$

So,

$$w = \sum_{i=1}^{l} \lambda_i y_i \bar{x}_i$$

$$= \lambda_1 y_1 \bar{x}_1 + \lambda_2 y_2 \bar{x}_2$$

$$= \frac{1}{4}(+1)(2,1) + \frac{1}{4}(-1)(4,3)$$

$$\Rightarrow \frac{1}{4}(2,1) - \frac{1}{4}(4,3)$$

$$= \frac{1}{4}(2-4, 1-3)$$

$$= \frac{1}{4}(-2,-2) = \left(-\frac{1}{2}, -\frac{1}{2}\right)$$

Scanned with CamScanner

$$b = -\frac{1}{2}\left(\min_{y_i=+1} \bar{w}\cdot\bar{x} + \max_{y_i=-1} \bar{w}\cdot\bar{x}\right)$$

$$= -\frac{1}{2}\left[\left(-\frac{1}{2} * 2\right) + \left(-\frac{1}{2}\times 1\right) + \left(-\frac{1}{2} * 4\right) + \left(-\frac{1}{2}\times 3\right)\right]$$

$$= -\frac{1}{2}\left[-1 - \frac{1}{2} - 2 - \frac{3}{2}\right]$$

$$= \frac{5}{2}$$

$$f(\bar{x}) = \bar{w}\cdot\bar{x} + b = 0$$

$$= \left(-\frac{1}{2}, -\frac{1}{2}\right)\cdot(x_1, x_2) + \frac{5}{2}$$

$$= -\frac{x_1}{2} - \frac{x_2}{2} + \frac{5}{2} = 0$$

$$\Rightarrow \boxed{x_1 + x_2 - 5 = 0} \longleftrightarrow \text{Soln.}$$

(14)

17201012

Answer to the Q.No. 4(a)

The criterion for minimization of error in regression problem are as follows:

i) Power rule: $\frac{d}{dx} x^n = n x^{n-1}$.

ii) Constant rule: if $f(x) = c$ then $f'(c) = 0$

iii) A constant times a function: $\frac{d}{dx} m x^n = m \times n x^{n-1}$

iv) Differentiating a sum: $\frac{d\alpha}{dx} = \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i$

v) Product rule: $\frac{d}{dx}[f(x) g(x)] = f(x) g'(x) + f'(x) g(x)$

vi) Quotient rule: $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x) f'(x) - f(x) g'(x)}{(g(x))^2}$

vi) Chain rule: $F'(x) = f'(g(x)) g'(x)$

These methods can be used in minimizing errors using derivatives.

17201092

Answer to the Q. No. 4(b)

Let A and B any two events in a random experiment. If $P(A) \neq 0$, then.

$$\boxed{P(B|A) = \frac{P(A|B)P(B)}{P(A)}} \rightarrow \text{Bayes theorem}$$

here, $P(A|B)$ is called posterior probability of A given B.

$P(B|A)$ is called likelihood of B given A.

Freq Table

class	features										
	hieght			weight				foot size			
	5.00	5.50	6.00	130	150	170	180	6	8	10	Total
male	0	1	3	0	0	2	2	0	1	3	4
female	2	1	1	2	2	0	0	2	2	0	4
Total	2	2	4	2	2	2	2	2	3	3	8

16

17201012

Likelihood table:

Class	Features									
	height			weight				foot size		
	5.00	5.50	6.00	130	150	170	180	6	8	10
male	0	$\frac{1}{2}$	$3/4$	0	0	1	1	0	$1/3$	1
~~femal~~ female	1	$\frac{1}{2}$	$1/4$	1	1	0	0	1	$2/3$	0

Now for person ~~50~~ with

height = 6, weight = 130 and footsize = 8

class	feature									
	height			weight				footsize		
	5.00	5.50	6.00	130	150	170	180	6	8	10
male	0	$1/4$	$3/4$	0	0	$2/4$	$2/4$	0	$1/4$	$3/4$
female	$2/4$	$1/4$	$1/4$	$2/4$	$2/4$	0	0	$2/4$	$2/4$	0

$$P(male) = \frac{3}{4} \times 0 \times \frac{1}{4} = 0$$

$$P(female) = \frac{1}{4} \times \frac{2}{4} \times \frac{2}{4} = 0.0625$$

Overall probability,

$$P(male) = 0 \times 4/8 = 0$$

$$P(femal) = 0.0625 \times 4/8 = 0.03125$$

So we assign the class lable ~~for~~ female to the ~~test~~ given test ~~to~~ instance.

Example of Bayes theorem:

$P(A) = 10\% = 0.1$

$P(B) = 5\% = 0.05$

$P(B|A) = 7\% = 0.07$

$$\therefore P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$= \frac{0.07 \times 0.10}{0.05}$$

$$= 0.14$$

17201012

Answer to the Q. No. 2 (a)

Artificial Neural Network:

An artificial neural network (ANN) is a computing system inspired by the ~~b~~ biological neutral networks that consititute biological brain. An ANN is based on a collection ~~of~~ of connected units called artificial neurons. Each connection between artificial neurons can transmit a signal from one to another. The artificial neuron that receives the signal can process it and then signal ~~a~~ artificial neurons connected to it.

Characteristics of ANN:

An ANN can be defined and implemented in several different ways. The way the following characteristics are defined determines a particulate variant of an ANN.

→ The activation function:

This function defines how a neuron's combined input signal are transformed into a ~~sign~~ single output signal to be broadcast further in the network.

→ The network topology:

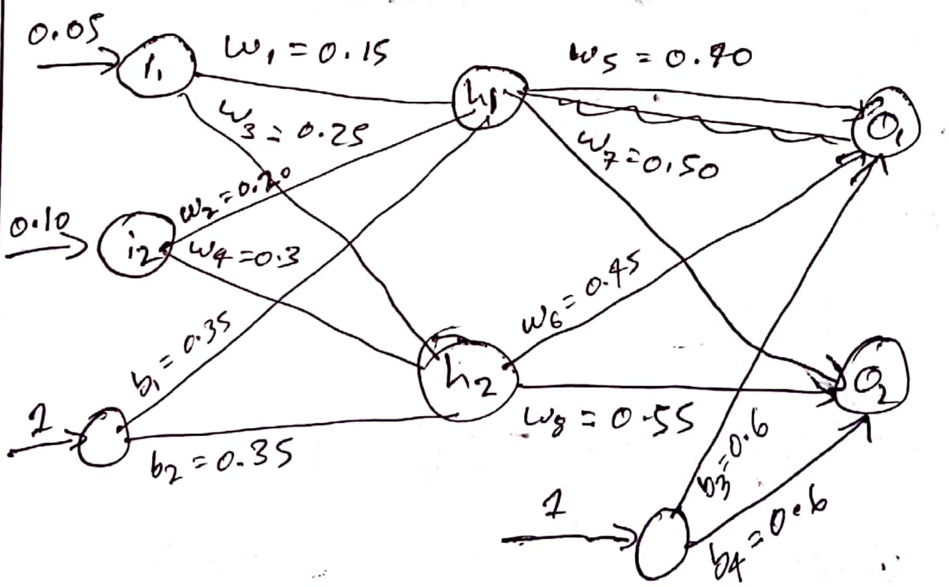
This describes the number of neurons in the model as well as the number of layers and manner in which they are connected.

→ The training algorithm:

This algorith specifies how connection weights are set in order to inhibit or excite neurons in proportion to the input signal.

Answer to the Q No. 2(b)

First observation



17201012

① Output calculation:

$Out_{h1} = \cancel{\phi}$ sigmoid $(w_1 i_1 \times w_2 i_2 \times b_1 1)$

sigmoid$(x) = \frac{1}{1+e^{-x}}$

$=$ sigmoid(0.3775)

$= 0.59$

$Out\cancel{\phi}_{h2} = \cancel{\phi}$ sigmoid $(w_3 i_1 + w_4 i_2 \cancel{\times} + b_2 \times 1)$

$=$ sigmoid(0.39)

$= \cancel{0.0} \; 0.59$

$Out_{o1} =$ sigmoid $(w_5 \times out_{h1} + w_6 \times out_{h2} + b_3 \times 1)$

$=$ sigmoid(1.11)

$= 0.75$

$Out_{o2} =$ sigmoid$(w_7 \times out_{h1} + w_8 out_{h2} + b_4 \times 1)$

$=$ sigmoid$(\cancel{0}\, 1.22)$

$= 0.77$

(ii) error calculation:

$E = \frac{1}{2}(T_1 - out_{o1})^2 + \frac{1}{2}(T_2 - out_{o2})^2$

$= (0.01 - 0.75)^2 + (0.99 - 0.77)2$

$= 0.298$

Scanned with CamScanner

(iii) weight adjust. of output layer

$\delta_{o1} = (T_1 - out_{o1}) \times out_{o1} \times (1 - out_{o1})$

$= -0.138$

$w_5^+ = w_5 + \eta \times \delta_{o1} \times out_{h1}$

$= 0.358$

$w_6^+ = w_6 + \eta\, \delta_{o1}\, out_{h2}$

$= 0.44$

$b_3^+ = b_3 + \eta\, \delta_{o1}\, 1$

$= 0.53$

$\delta_{o2} = (T_2 - out_{o2}) \times out_{o2} \times (1 - out_{o2})$

$= 0.03$

$w_7^+ = w_7 + \eta\, \delta_{o2}\, out_{h1}$

$= 0.511$

$b_4^+ = b_4 + \eta\, \delta_{o2}\, 1$

$= 0.62$

$w_8^+ = w_8 + \eta\, \delta_{o2}\, out_{h2} = 0.56$

22

17201012

(iv) weigh adjust for hidden layer:

$$\delta_{h1} = (\delta_{o1} w_5 + \delta_{o2} w_7) \; out_{h1} \times (1 - out_{h1})$$

$$= -0.008$$

$$w_1^+ = w_1 + \eta \, \delta_{h1} \times i_1$$

$$= 0.05$$

$$w_2^+ = w_2 + \eta \, \delta_{h1} \times i_2$$

$$= 0.19$$

$$b_1^+ = b_1 + \eta \times \delta_{h1} \times 1 = 0.34$$

$$\delta_{h2} = (\delta_{o1} w_6 + \delta_{o2} w_8) \times out_{h2} \times (1 - out_{h2})$$

$$= -0.009$$

$$w_3^+ = w_3 + \eta \, \delta_{h2} \times i_1 = 0.249$$

$$w_4^+ = w_4 + \eta \, \delta_{h2} \, i_2 = 0.299$$

$$b_2^+ = b_2 + \eta \times \delta_{h2} \times 1 = 0.345$$

For the second observation, values ~~of~~ will be

$i_1 = 0.25$

$i_2 = 0.18$

$T_1 = 0.23$

$T_2 = 0.79$.

For these values steps (i) - (iv) will be repeated for a single iteration.