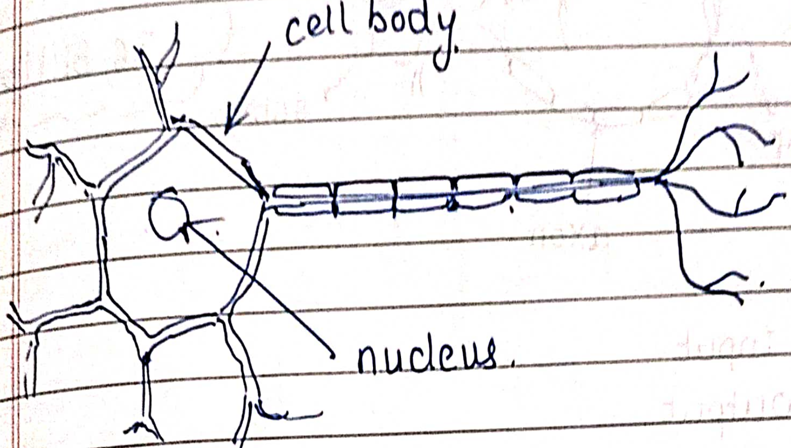
Deep learning.

1. Introduction to deep learning.

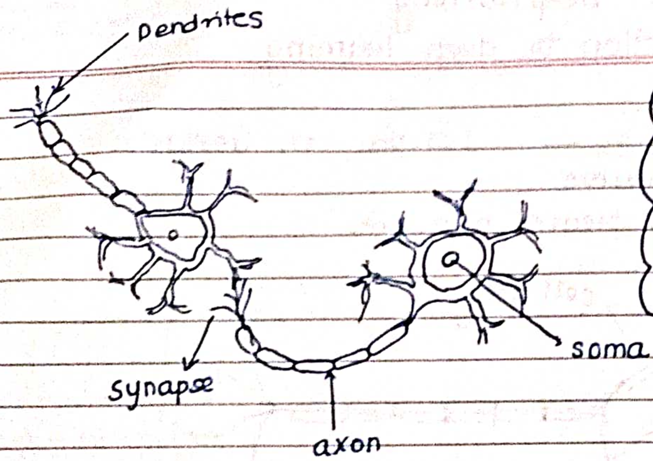
1871-1873 — Joseph Von Gerlach

- Artificial neuron
- Recurrent Neural network



- Multilayer perceptron
- Recurrent neural network.

1) Biology of neural network. (neuron).
2) Idea of computational units.
3) McCulloch-Pits unit / History of Deep learning
4) Deep learning workflow,
5) Learning types McCulloch Pitts Neuron.

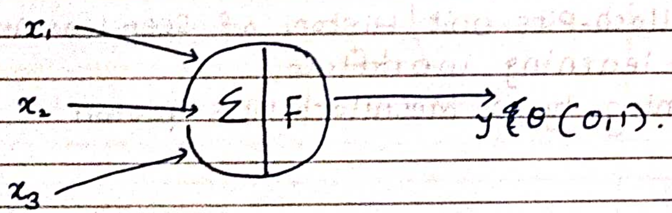


How many neurons in human brain?

86 billion

Dendrites - Input
Axon - Output
Soma - Body
Synapse -

* McChulloch pits Neural network.



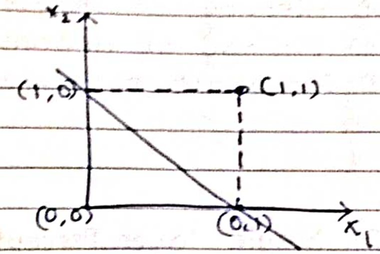
Feed Forward network.

AND GATE

x_1	x_2	y	g
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	2

Implement McCulloch Pit Neural network using OR GATE

x_1	x_2	y	g
1	1	1	2
1	0	1	1
0	1	1	1
0	0	0	0



$$\sum_{i=1}^{n} x_i w_i = x_1 w_1 + x_2 w_2 + \dots x_n w_n$$

① $x_1 = 0$ $x_2 = 0$.
Assume $w_1, w_2 = 1$.
Assume $\theta = 2$.
$g = x_1 w_1 + x_2 w_2$
$= 0 \times 1 + 0 \times 1$
$= 0$
$\therefore y = f(g(x))$
$\boxed{y = 0.}$

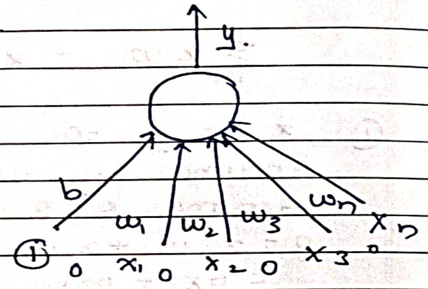
② $x_1 = 0$ $x_2 = 1$
Assume $w_1, w_2 = 1$.
$g = x_1 w_1 + x_2 w_2$
$= 0 \times 1 + 1 \times 1$
$= 0 + 1$
$= 1$
$\therefore y = f(g(x))$
$\boxed{y = 1}$

③ $x_1 = 1$, $x_2 = 0$.
Assume $w_1, w_2 = 1$
$g = x_1 w_1 + x_2 w_2$
$= 1 \times 1 + 0 \times 1$
$= 1$
$y = f(g(x))$
$\boxed{y = 1}$

④ $x_1 = 1$, $x_2 = 1$
Assume $w_1, w_2 = 1$
$g = x_1 w_1 + x_2 w_1$
$= 1 \times 1 + 1 \times 1$
$= 2$
$y = f(g(x))$
$\boxed{y = 1}$

Assignment
1) History of deeplearning from 1871
2) Biology of neuron
3) McCulloch Pits neural network.

27/7/23 Single layer Perceptron
→ (1969) Papper + Minsky
→ learning rate (α)



$$\boxed{y_{in} = b + \sum_{i=1}^{n} w_i * x_i}$$

$\Delta w_i = old + \alpha x_i t$

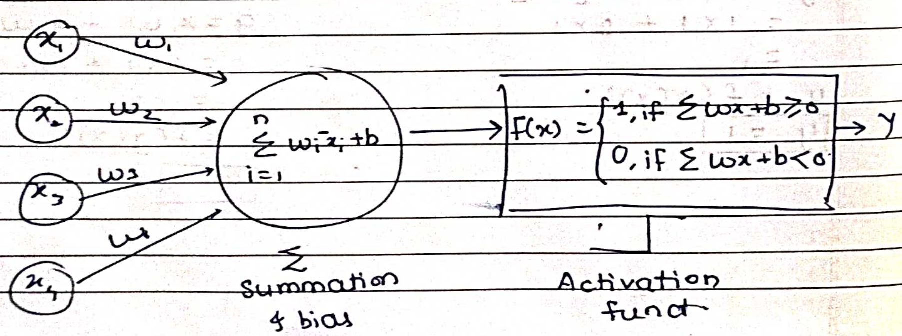
$\Delta b = \alpha t$

$\alpha = 1$

AND GATE

x_1	x_2	y
-1	-1	-1
-1	1	-1
1	-1	-1
1	1	1

$$y = f(y_{in}) \begin{cases} 1 \text{ if } y_{in} > \theta \\ 0 \text{ if } y_{in} = \theta / \theta \le y_{in} \le \theta \\ -1 \text{ if } y_{in} < \theta \end{cases}$$



Epoch 1

4)

Input x_1	x_2	target (t)	Netinput (y_{in})	actual o/p y	weight change Δw_1 (0)	Δw_2 (0)	Δb (0)	weight w_1 (0)	w_2 (0)	b (0)
-1	-1	-1	0	-1	0	0	0	0	0	0
-1	1	-1	0	-1	0	0	0	0	0	0
1	-1	-1	0	-1	0	0	0	0	0	0
1	1	1	0	-1	1	1	1	1	1	1

① $y_{in} = b + \sum_{i=1}^{n} w_i * x_i$

② $w_{(new)} = w(old) + \alpha x t$

③ $\Delta b = \alpha t$

~~taking $\theta = 2$ in above epoch~~

$$y = f(y_{in}) \begin{cases} 1 \text{ if } y_{in} > 2 \\ 0 \text{ if } -2 \le y_{in} < 0 \\ -1 \text{ if } y_{in} < \theta_2 \end{cases}$$

Epoch 2

Input x_1	x_2	target (t)	Net input (y_{in})	actual o/p y	weight change Δw_1	Δw_2	Δw_3	weight w_1 (1)	w_2 (1)	w (1)
-1	-1	-1	-1	-1	1	1	1	1	1	1
-1	1	-1	1	-1	1	1	1	1	1	1
1	-1	-1	1	-1	1	1	1	1	1	1
1	1	1	3	1	1	1	1	1	1	1

$w_1 = 1$ $w_2 = 1$

① $y_{in} = b + \sum_{i=1}^{n} w_i * x_i$

$= 1 + (1)(-1) + 1(-1)$

$= 1 - 1 - 1 = -1$

② $y_{in} = b + \sum_{i=1}^{n} w_i * x_i$

$= 1 + (-1)(1) + (1)(1)$

$= 1 - 1 + 1$

$= 1$

③ $y_{in} = b + \sum_{i=1}^{n} w_i * x_i$

$y_{in} = 1 + (1)(1) * (-1)(1)$

$= 1 + 1 - 1$

$= 1$

$w(new) = w(old) + \alpha x_i t$

$= 1 + 1 \times -1 \times -1$

$= 1 + 1$

$= 2$

④ $y_{in} = b + \sum_{i=1}^{n} w_i * x_i$

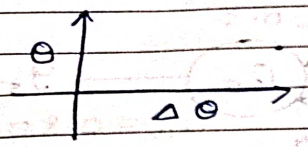
$y_{in} = 1 + (1)(1) + (1)(1)$

$= 1 + 1 + 1$

$= 3$

classmate
Date 8|8|23
Page

* Gradient Descent

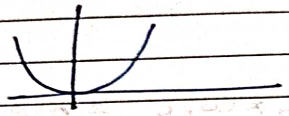


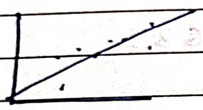
$\longrightarrow \quad \theta = (w, b)$

$\Delta\theta = (\Delta w, \Delta b)$

$\Delta w = old + \alpha\, x_i t$

$\rightarrow$ new = old + (stepsize × learning rate)



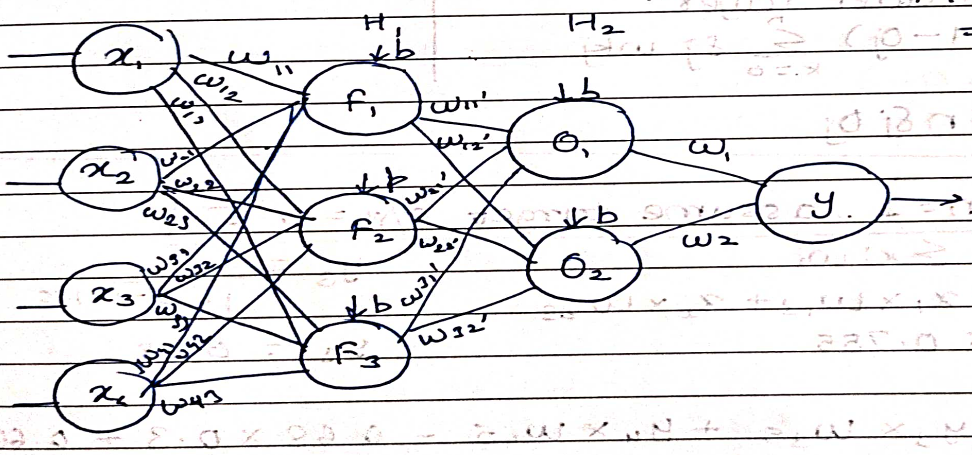


$\longrightarrow \quad y = mx + c$

* Multilayer perceptron

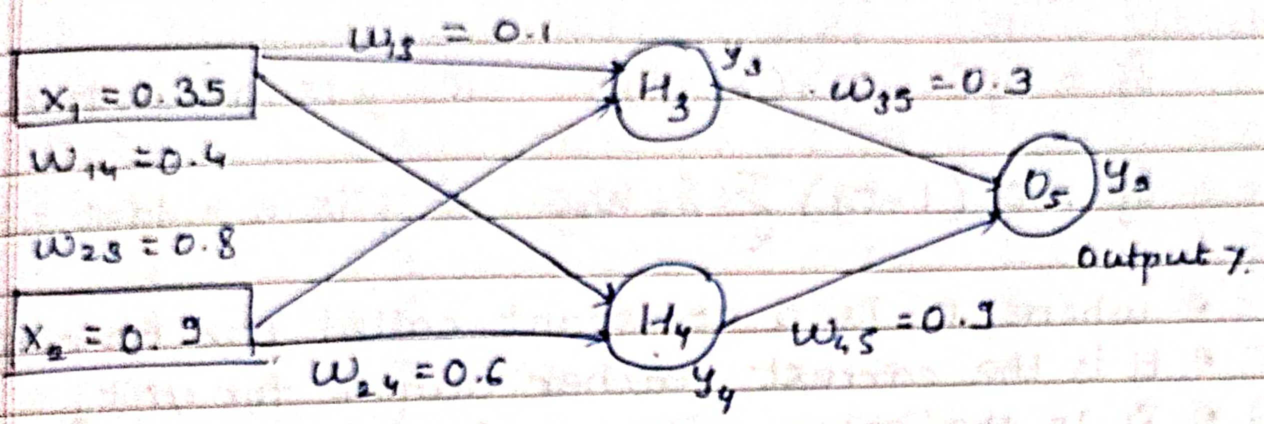
$f_i = b + \sum x_i w_i$

$= [x_1 w_{11} + x_2 w_{21} + x_3 w_{31} + x_4 w_{41}] + b.$



Assume :- (1) neurons have a sigmoid activation funct?.
perform a forward pass + a backward pass on the netw
(2) Assume, actual o/p => y = 0.5 + learning rate is 1.

Revision.



• forward Pass: $y3, y4 + y5$

$$\boxed{a_j = \sum_j (w_{ij} * x_i)}$$

$$\boxed{y_j = F(a_j) = \frac{1}{1+e^{-a_j}}}$$

$a_1 = (w_{13} * x_1) + (w_{23} * x_2)$
$= (0.1 * 0.35) + (0.8 \times 0.9)$
$= 0.755$

$y_3 = f(a_1)$
$= \frac{1}{1+e^{-0.755}}$
$\boxed{y_3 = 0.68.}$

$a_2 = (w_{14} * x_1) + (w_{24} * x_2)$
$= (0.4 * 0.35) + (0.6 * 0.9)$
$= 0.6637$

$y_4 = f(a_2) = \frac{1}{1+e^{-0.68}}$
$\boxed{y_4 = 0.6637}$

$a_3 = (w_{35} * y_3) + (w_{45} * y_4)$
$= (0.3 * 0.68) + (0.9 * 0.6637) = 0.801$

$y_5 = f(a_3)$
$= \frac{1}{1+e^{0.801}}$
$\boxed{y_5 = 0.69.}$

$\boxed{\text{Network output} = 0.69.}$

But expected output $= 0.5$

$\boxed{\therefore \text{ Error} = y_{target} - y_5 = 0.5 - 0.69 = -0.19}$

- Each weight changed by:

$$\Delta w_{ji} = \eta \delta_j O_i$$

$$\delta_j = O_j(1-O_j)(t_j - O_j) \quad \text{if } j \text{ is an output unit}$$

$$\delta_j = O_j(1-O_j)\sum_k \delta_k W_{kj} \quad \text{if } j \text{ is a hidden unit.}$$

- where η is a constant called the learning rate
- t_j is the correst teacher output for unit j
- δ_j is the error measure for unit j.

1) Backward Pass: δ_3, δ_4 & δ_5.

For output unit:

$$\delta_5 = y_5(1-y_5)(y_{target} - y_5)$$

$$= 0.69(1-0.69)(0.5-0.69) = -0.0406.$$

For hidden unit:

$$\delta_3 = y_3(1-y_3) w_{35} * \delta_5.$$

$$= 0.68(1-0.68)(0.3 * -0.0406) = -0.0082$$

$$\Delta w_n = \eta \delta_j O_i$$

① $\Delta w_{45} = \eta \delta_5 y_4 = 1 * -0.0406 * 0.6637 = -0.0269.$

$$W_{45}(new) = \Delta W_{45} + W_{45}(old)$$

$$= -0.0269 + (0.9) = 0.8731$$

② $\Delta w_{14} = \eta \delta_4 x_1 = 1 * -0.0082 * 0.35 = -0.00287$

$$W_{14}(new) = \Delta w_{14} + W_{14}(old) = -0.00287 + 0.4 = 0.3971$$

Similarly, update all other weights

i	j	w_{ij}	δ_j	x_i	η	Updated w_{ij}
1	3	0.1	-0.00265	0.35	1	0.0991
2	3	0.8	-0.00265	0.9	1	0.7976
1	4	0.4	-0.0082	0.35	1	0.3971
2	4	0.6	-0.0082	0.9	1	0.5926
3	5	0.3	-0.0406	0.68	1	0.2724
4	5	0.9	-0.0406	0.6637	1	0.8731.

(3) $\Delta w_{13} = \eta \delta_3 x_1$

$= 1 * (-0.0082)(0.35)$

$= 1 * (-0.00287)$

$\Delta w_{13} = -0.00287$

$w_{13}(new) = \Delta w_{13} + w_{13}(old)$.

$= -0.00287 + 0.1$

$= 0.09713 \approxeq 0.0991$

(4) $\Delta W_{23} = \eta \delta_3 x_2 = 1 * (-0.00265) * (0.9)$

$= -0.002385$

$W_{23}(new) = \Delta W_{23} + W_{23}(old)$

$= -0.002385 + 0.8 = 0.7976$.

(5) $\Delta w_{24} = \eta \delta_4 x_2 = 1 * (-0.0082) * 0.9 = -0.00738$

$W_{24}(new) = \Delta W_{24} + W_{24}(old)$

$= -0.00738 + 0.6 = 0.5926$.

(6) $\Delta W_{35} = \eta \delta_5 y_3 = 1 * (-0.0406) * (0.68) = -0.0276$

$W_{35}(new) = \Delta W_{35} + W_{35}(old)$

$= -0.0276 + 0.3 = 0.2724$.

forward Phase II

$a_1 = (w_{13}(new) * x_1) + (w_{23(new)} * x_2)$
$= (0.0991 * 0.35) + (0.7976 * 0.9)$.
$= 0.034685 + 0.71784$
$= 0.7525$

$y_3 = f(a_1) = \frac{1}{1+e^{-0.7525}}$
$= \frac{1}{1 + 0.4711}$
$= \frac{1}{1.4712} = \cancel{0.9550}\ 0.6797$

~~$y_4 = f$~~
$a_2 = (w_{14}(new) * x_1) + (w_{24}(new) * x_2)$
$= (0.3971 * 0.35) + (0.5926 * 0.9$.
$= 0.1389 + 0.53334$
$= 0.6722$

$y_4 = f(a_2) = \frac{1}{1+e^{-0.6722}} = \frac{1}{1+0.510} = \frac{1}{1.5105}$
$y_4 = 0.6620$

$a_3 = (w_{35}(new) * y_3) + (w_{45}(new) * y_4)$
$= (0.2724 * 0.68) + (0.8731 * 0.6637)$.
$= 0.1852 + 0.5794$.
$= 0.7646$.

$y_5 = f(a_3) = \frac{1}{1+e^{-0.7646}} = \frac{1}{1+0.4655} = \frac{1}{1.4655}$
$y_5 = 0.6829$