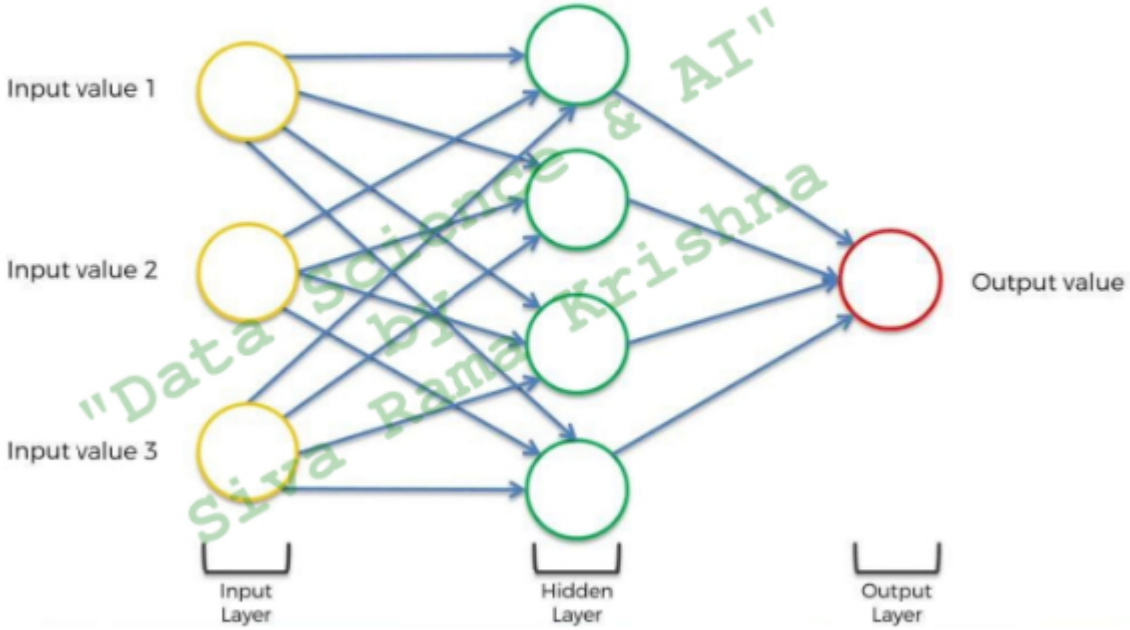
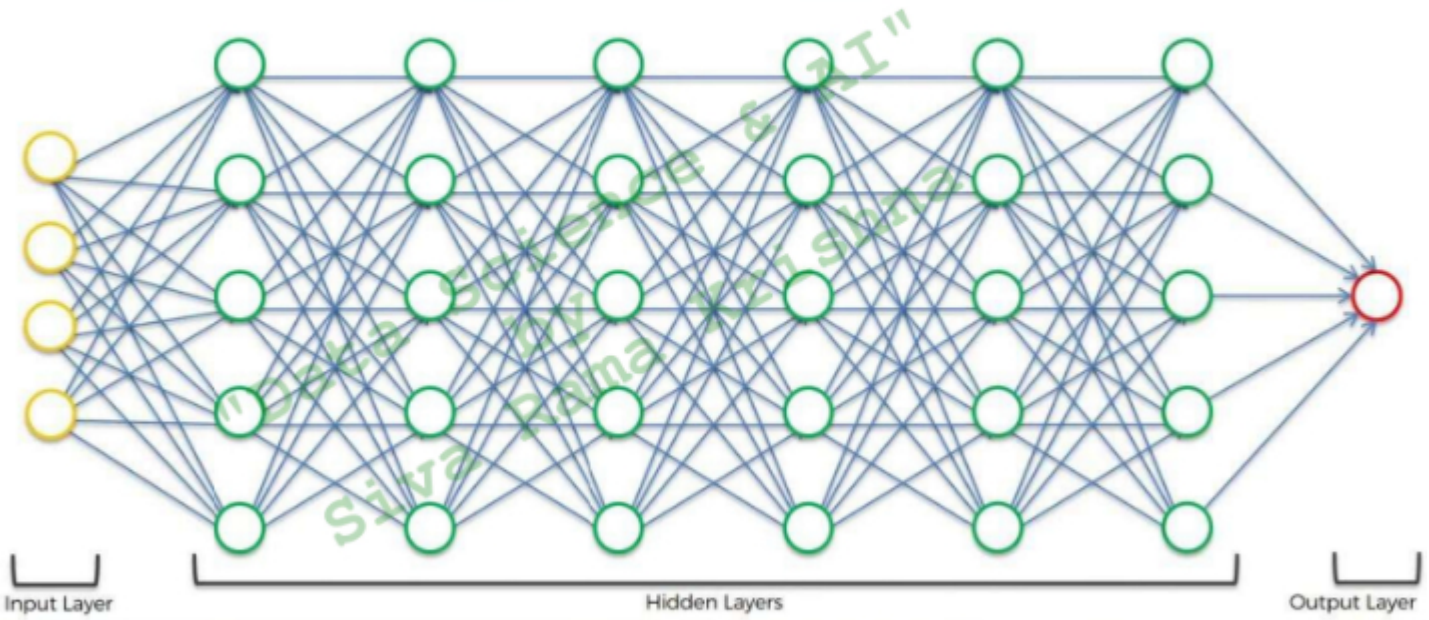
What is Deep Learning?

What is Deep Learning?



What is Deep Learning?



The Neuron

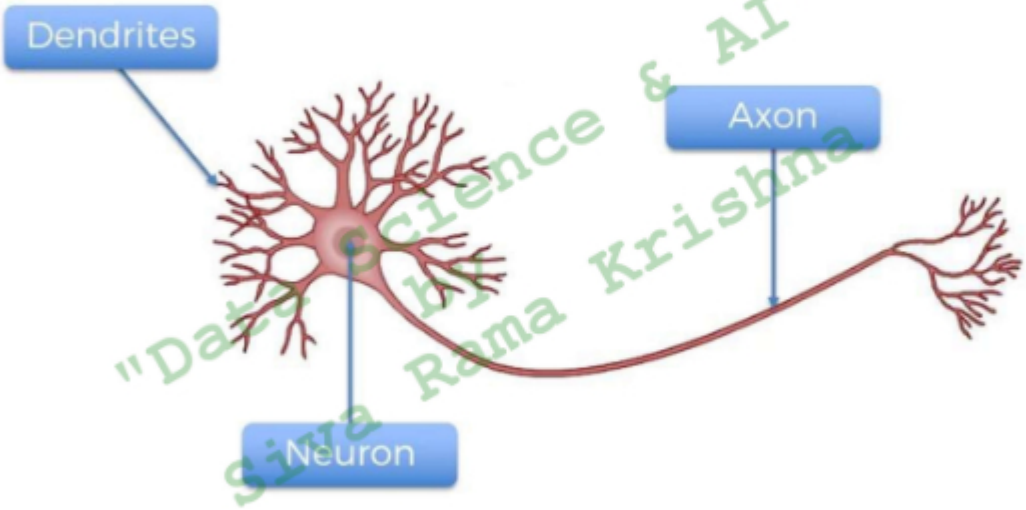
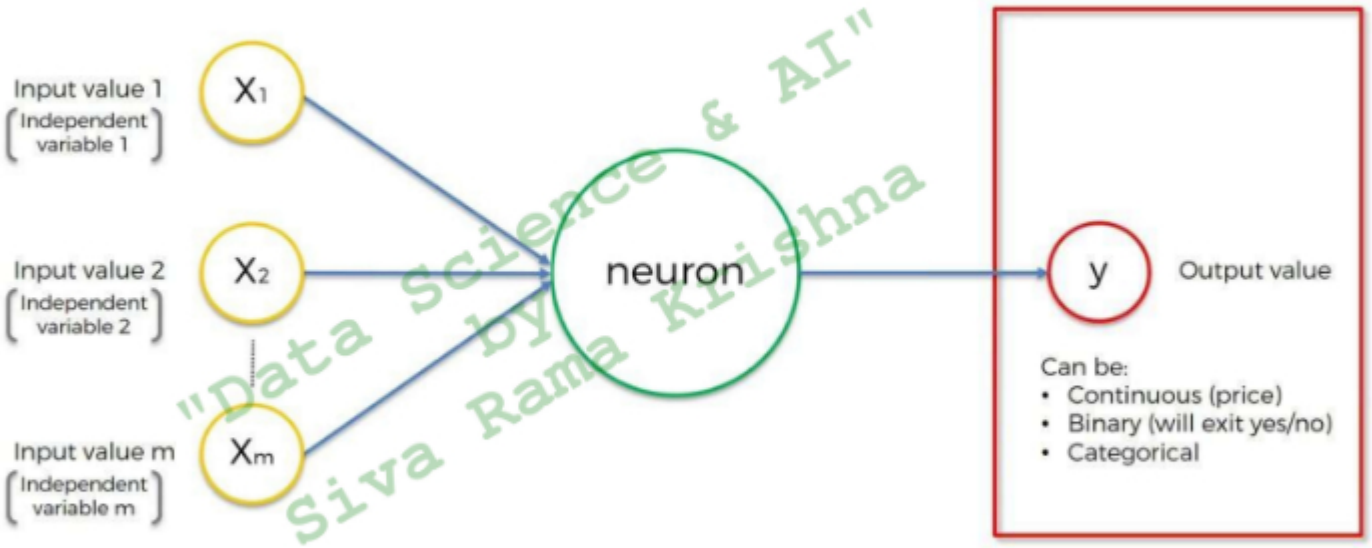
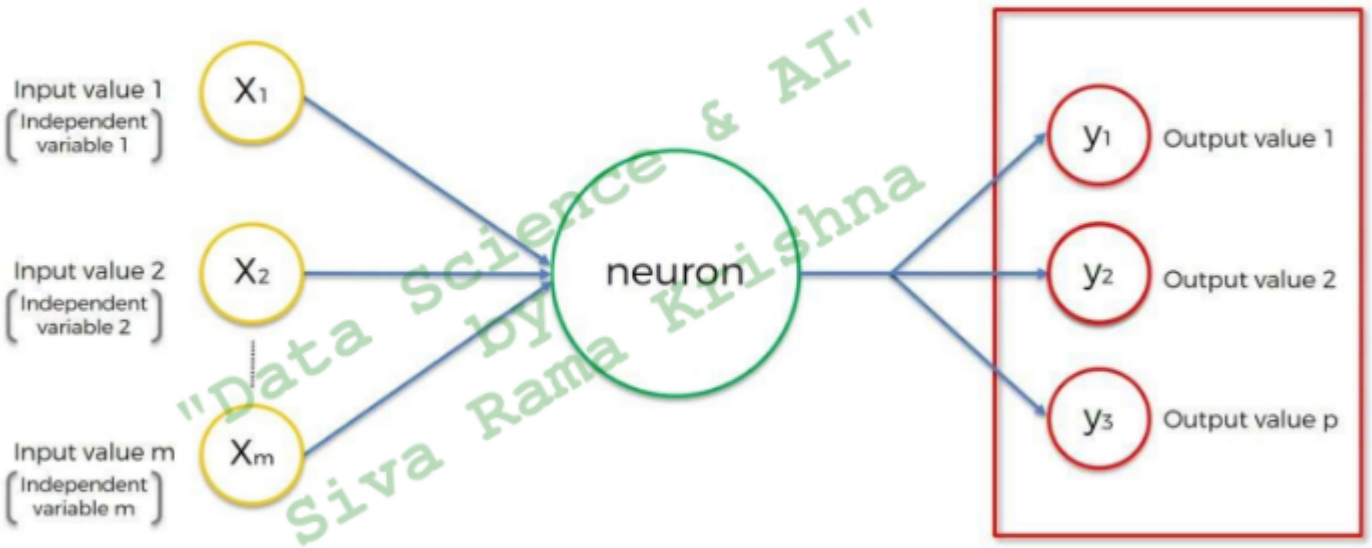


Image Source: Wikipedia

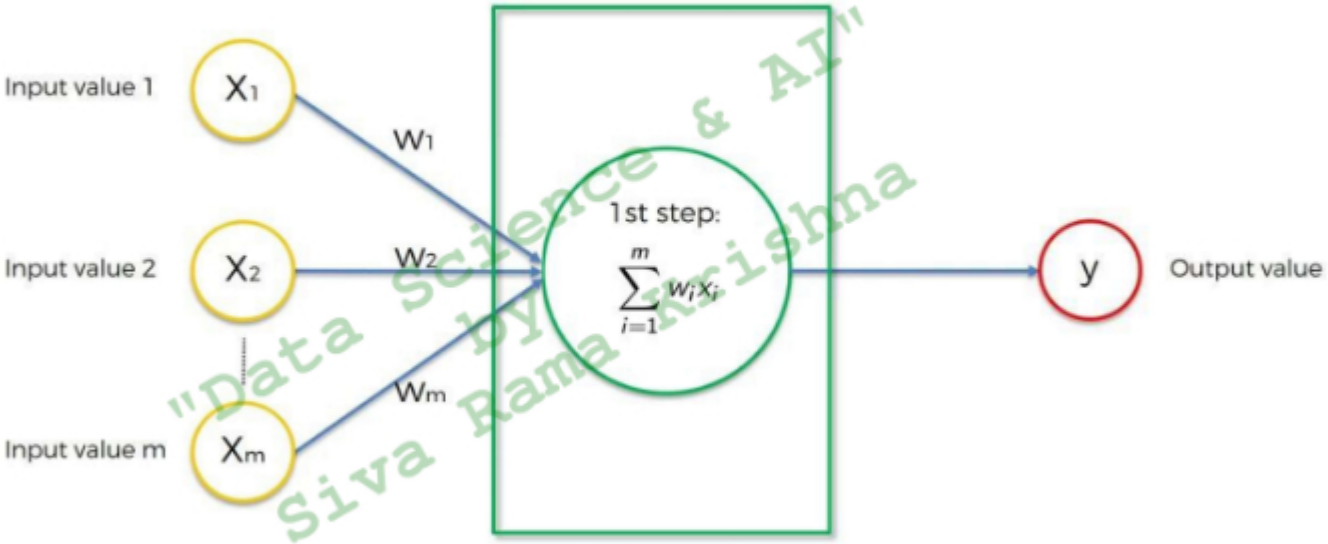
The Neuron



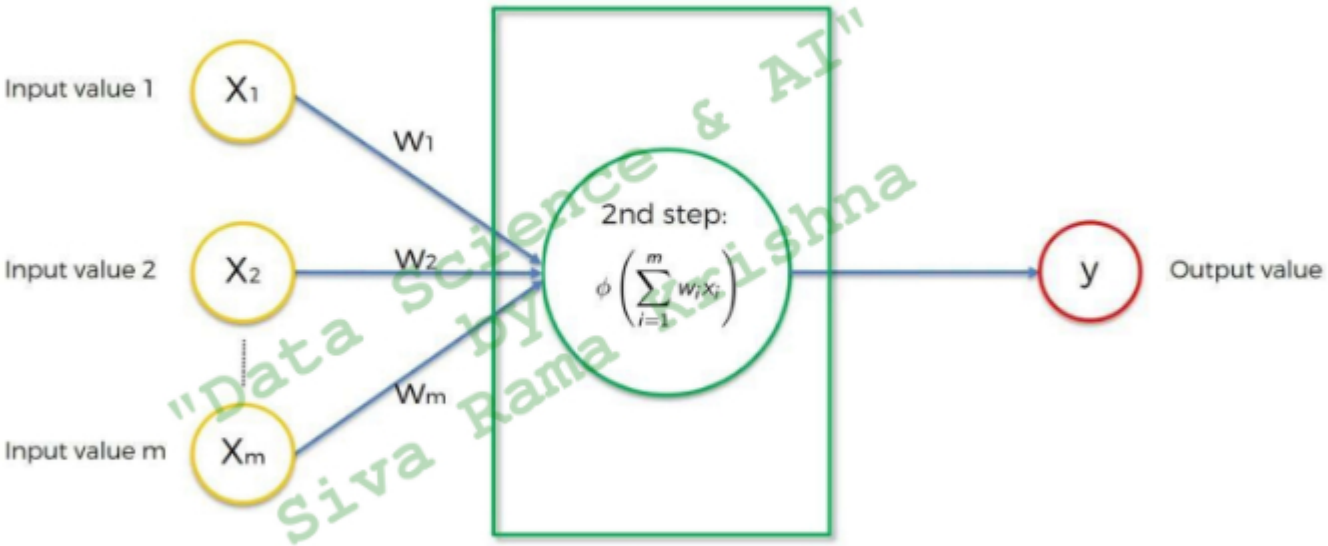
The Neuron

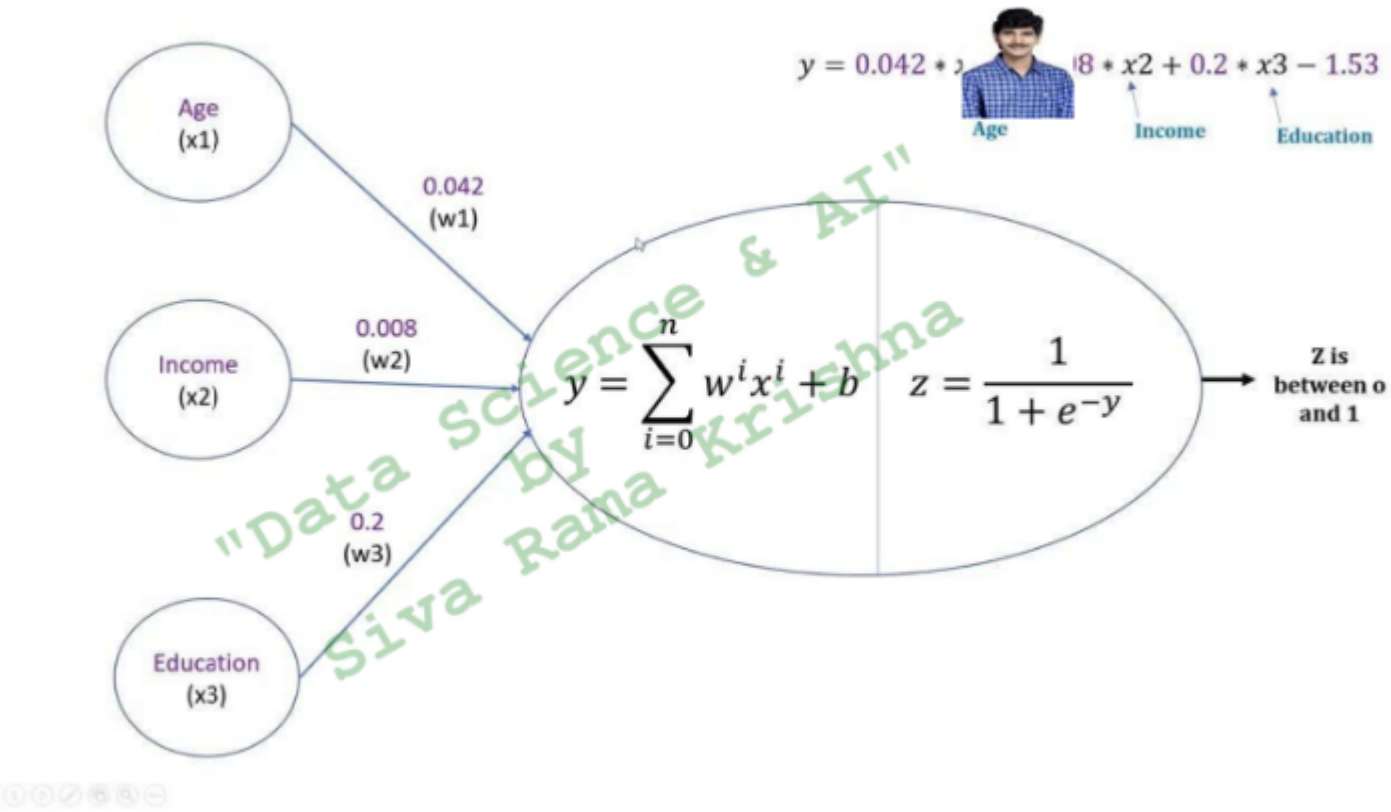


The Neuron

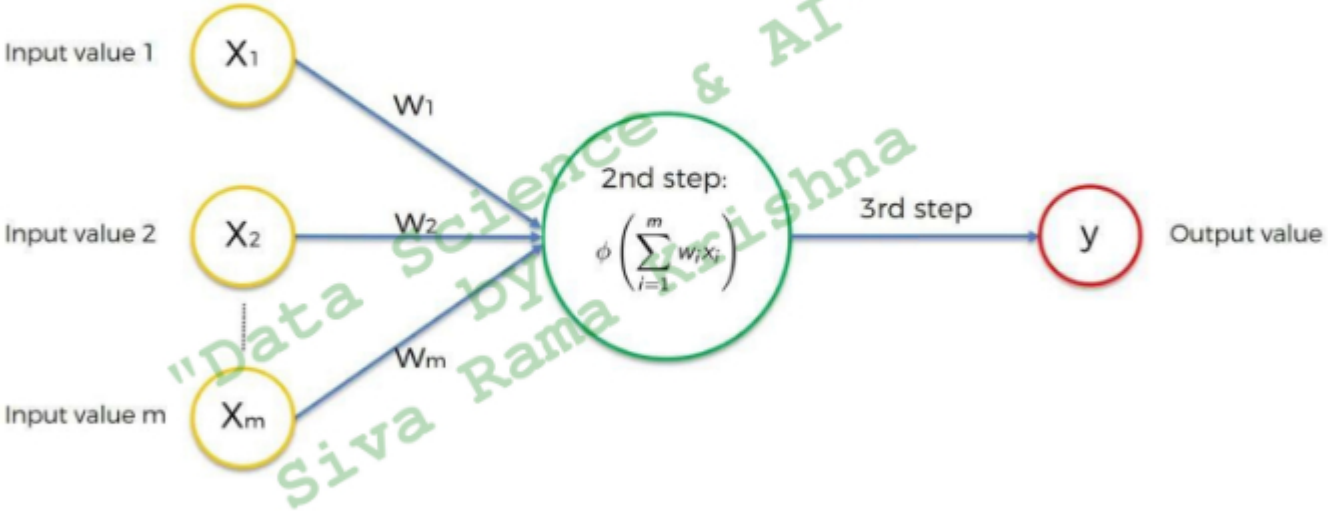


The Neuron

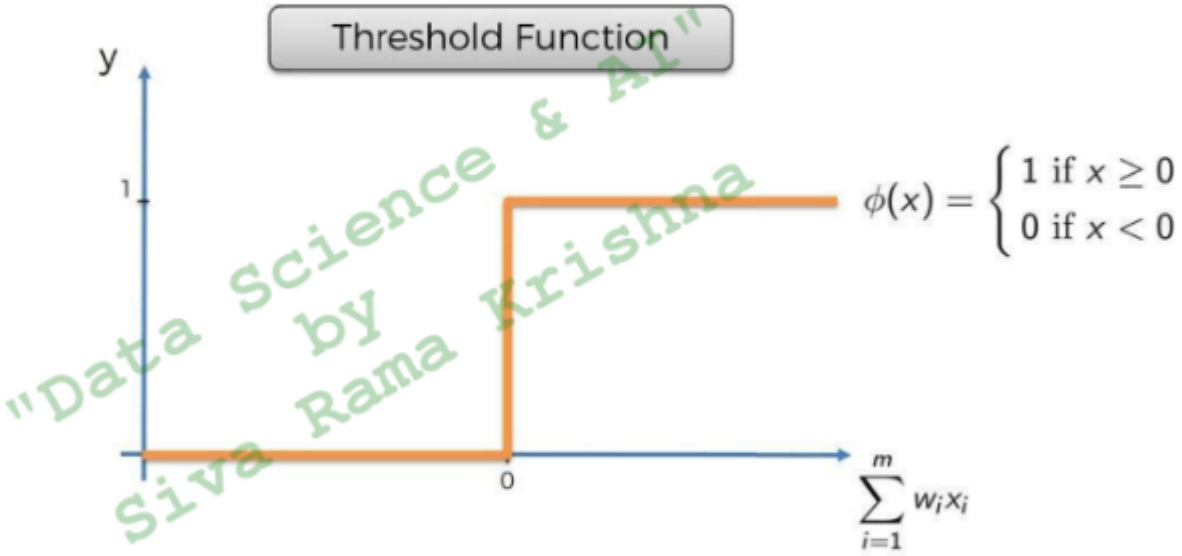



$y = 0.042 * x1 + 0.008 * x2 + 0.2 * x3 - 1.53$
Age
Income
Education
Age
(x1)
0.042
(w1)
Income
(x2)
0.008
(w2)
0.2
(w3)
Education
(x3)
$y = \sum_{i=0}^{n} w^i x^i + b$
$z = \frac{1}{1 + e^{-y}}$
Z is between o and 1
"Data Science & AI" by Siva Rama Krishna

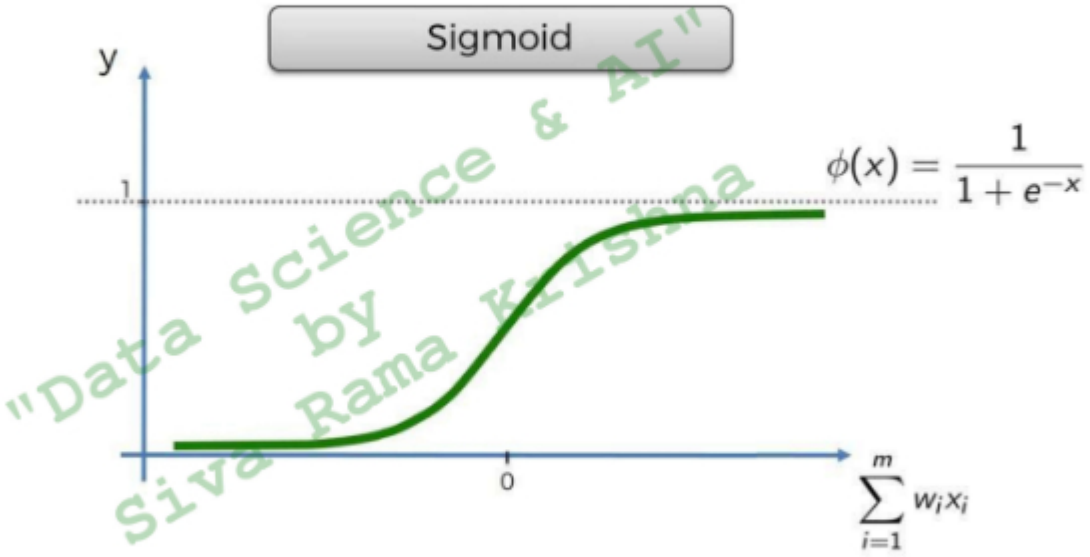
The Activation Function



The Activation Function

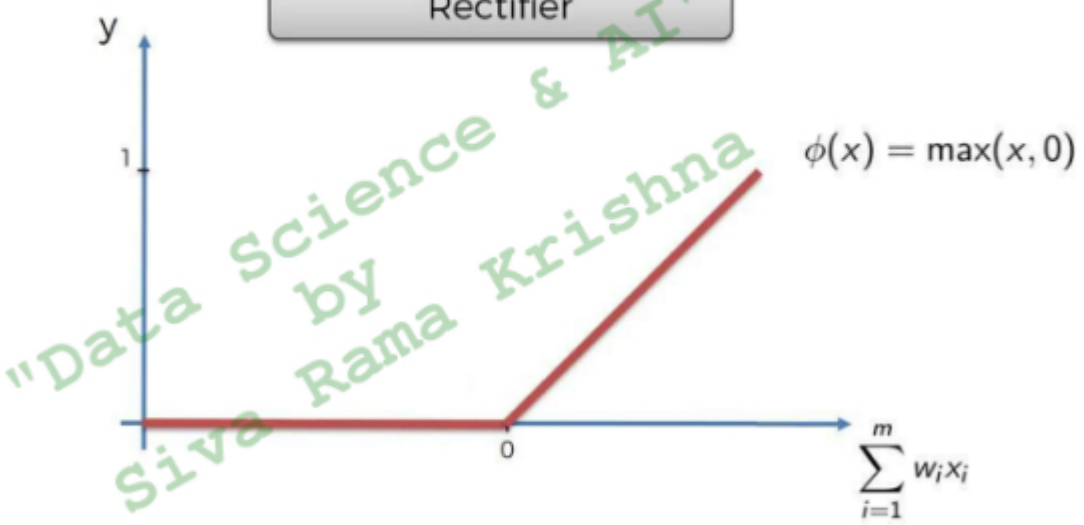


The Activation Function



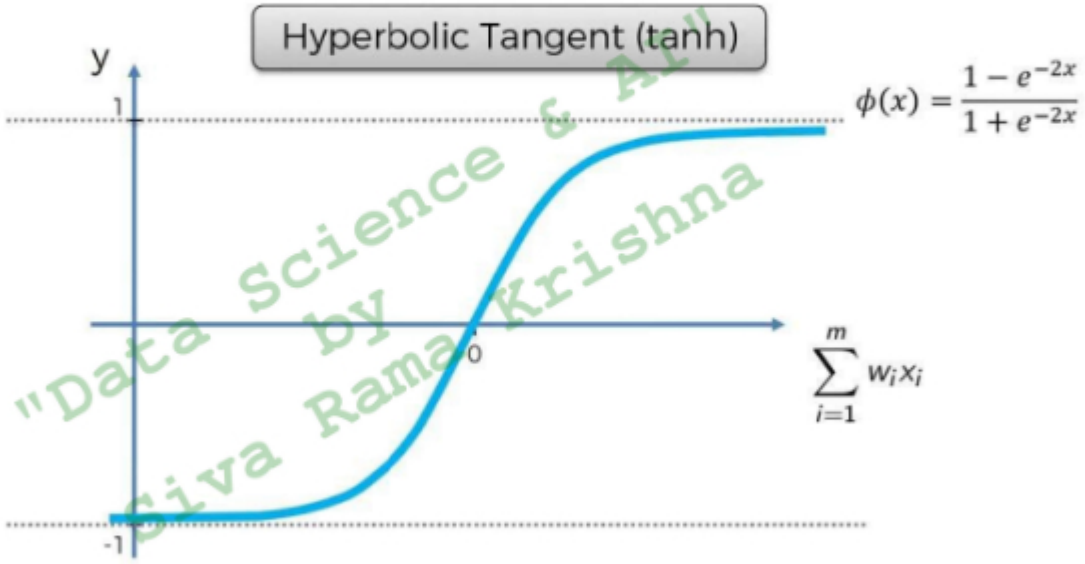
The Activation Function

Rectifier

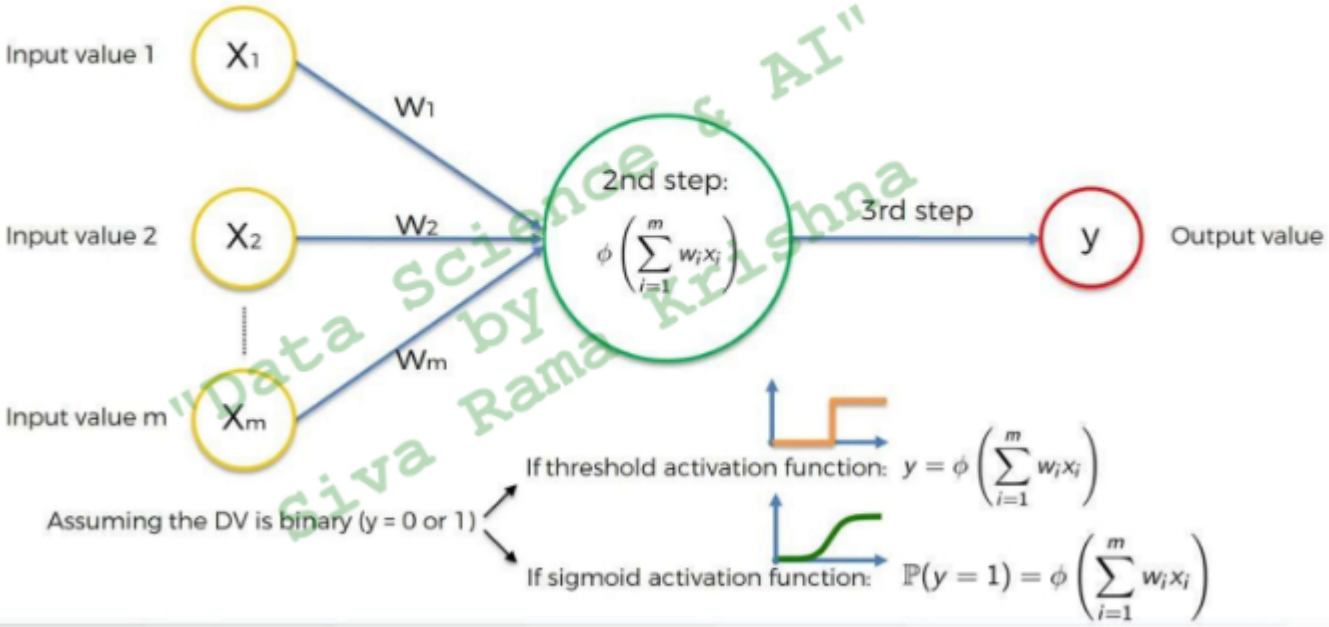


$$\phi(x) = \max(x, 0)$$

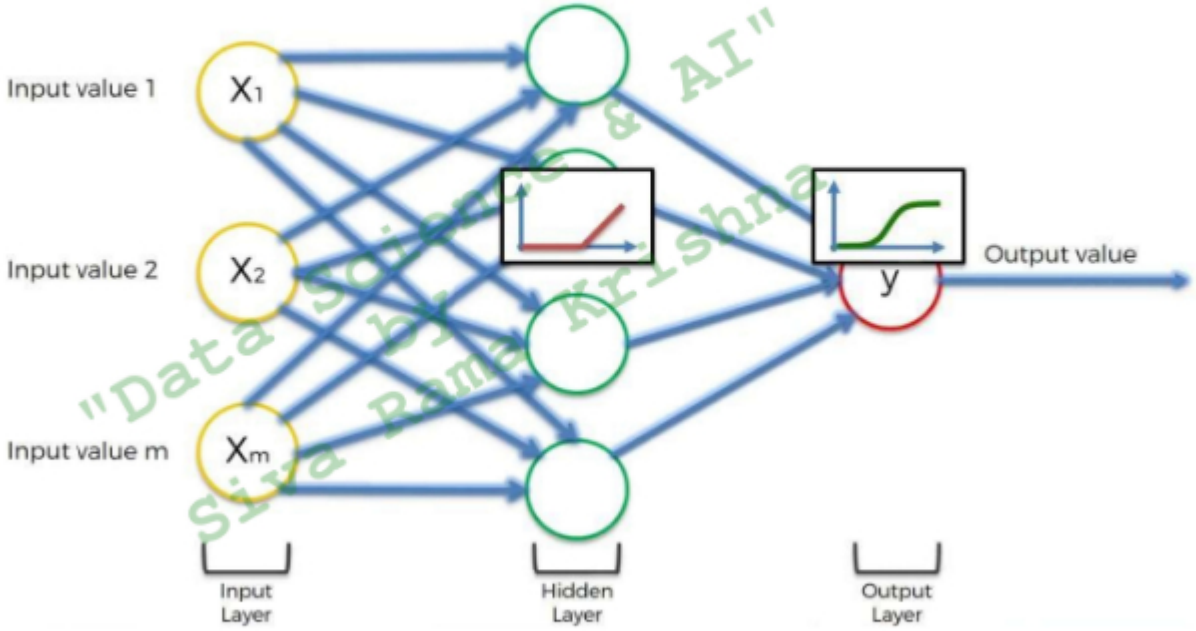
The Activation Function

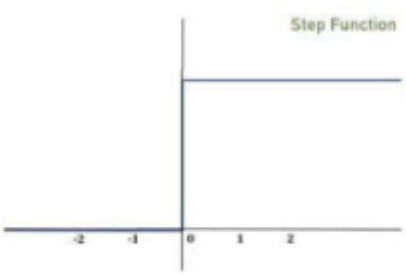


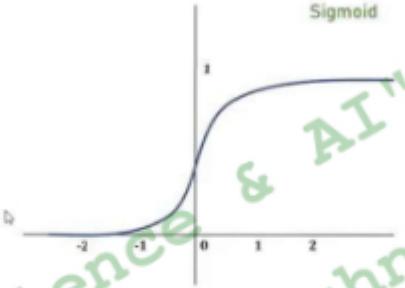
The Activation Function

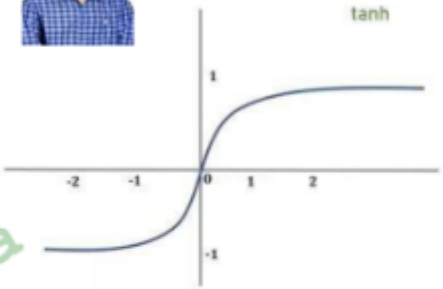


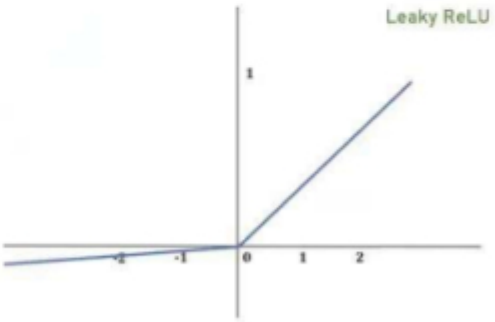
The Activation Function







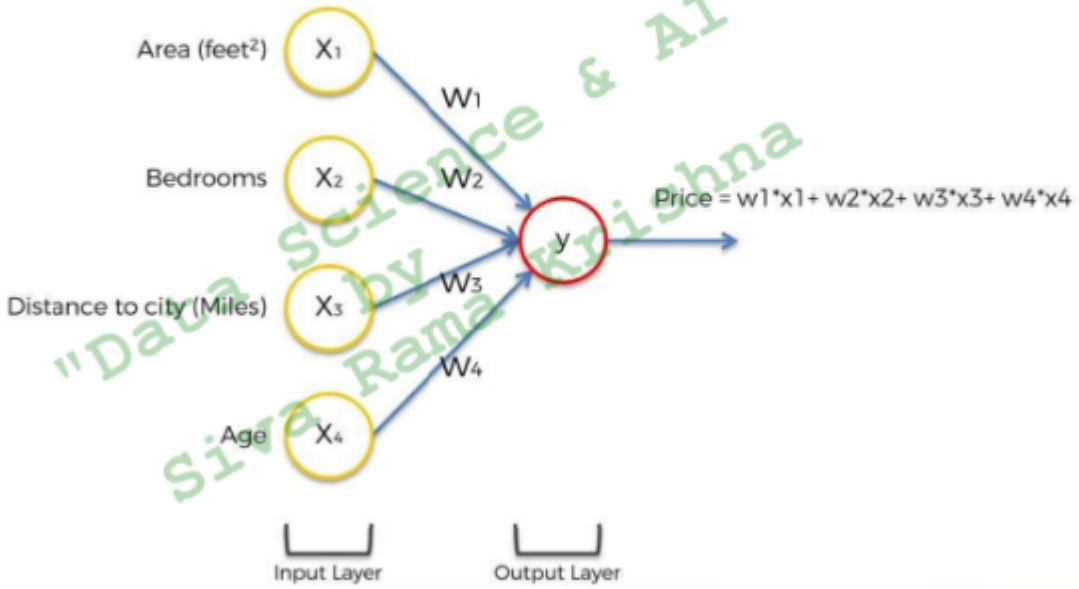




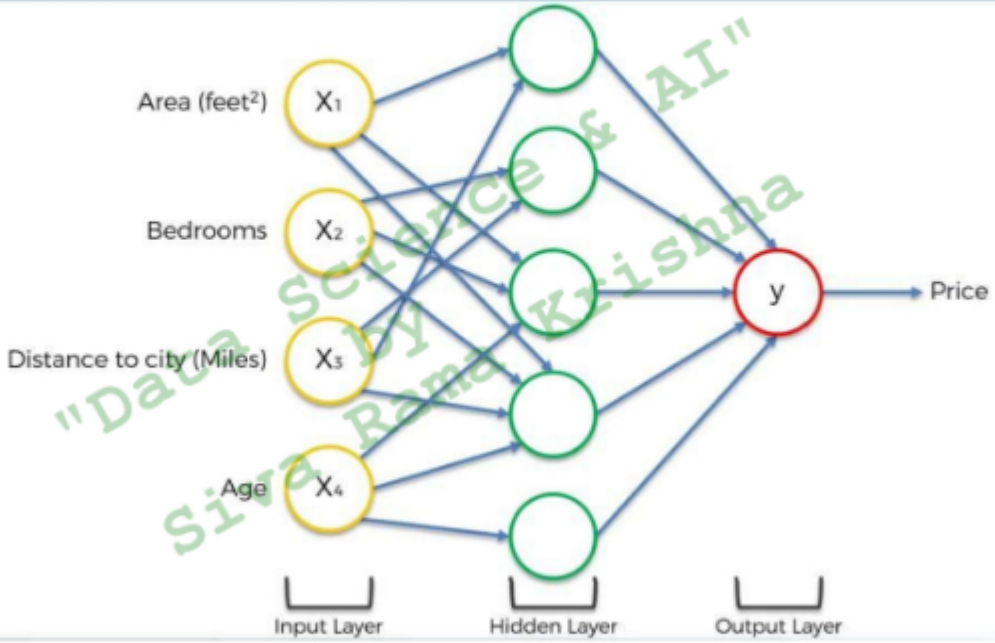
For hidden layers if you are not sure which activation function to use, just use ReLU as your default choice

Use sigmoid in output layer.

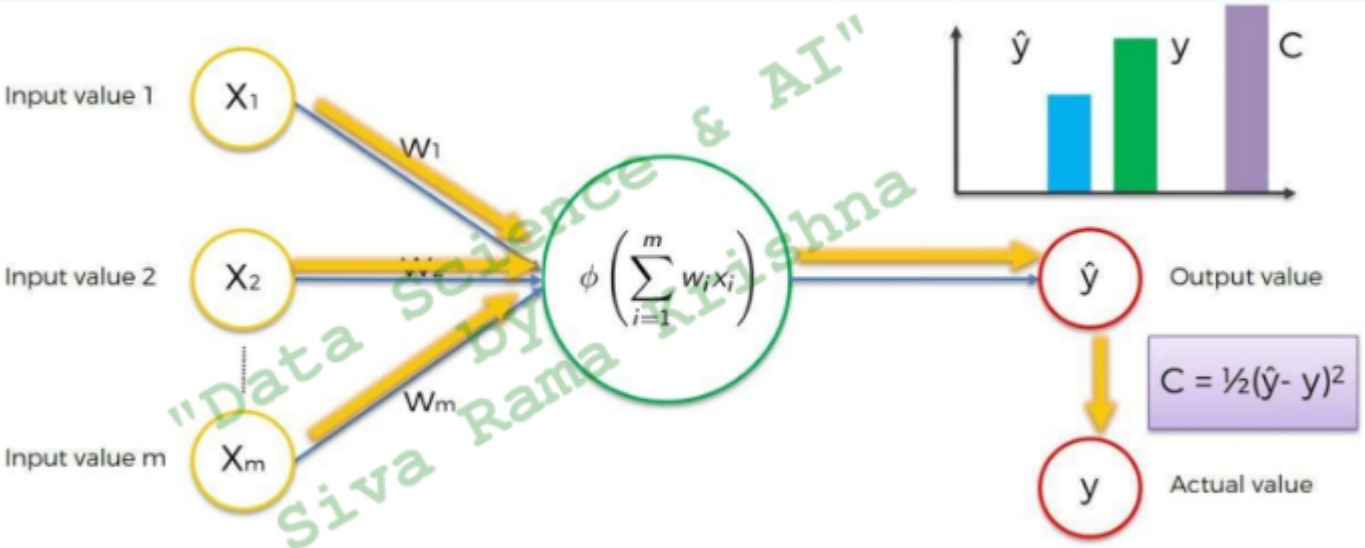
How Do Neural Networks  ork?



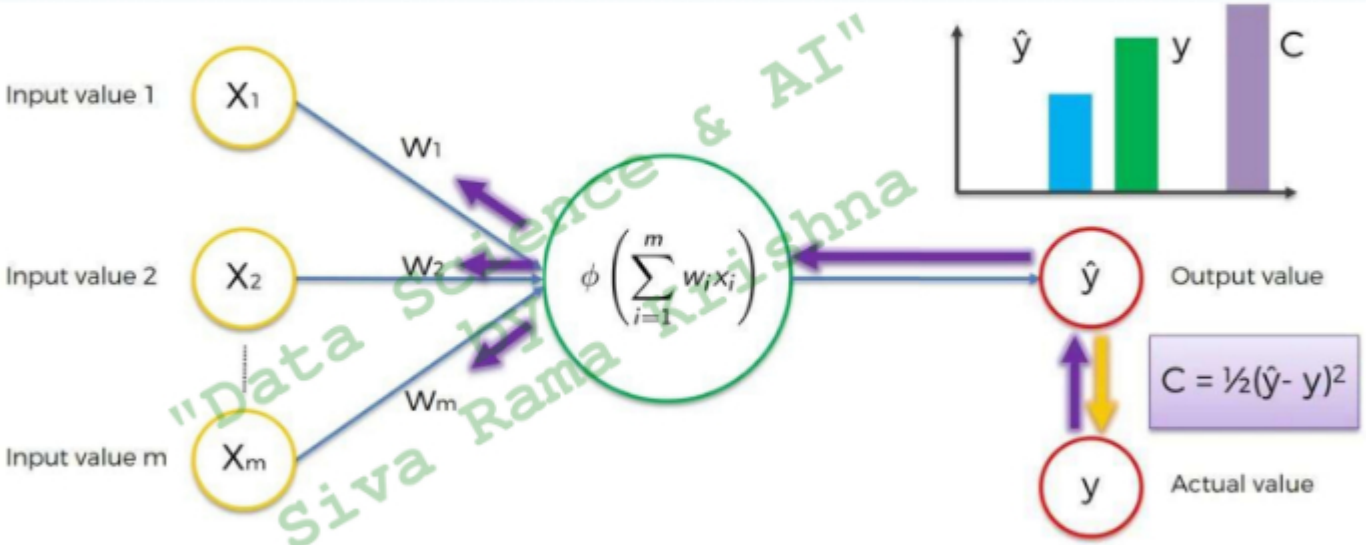
How Do Neural Networks  ork?



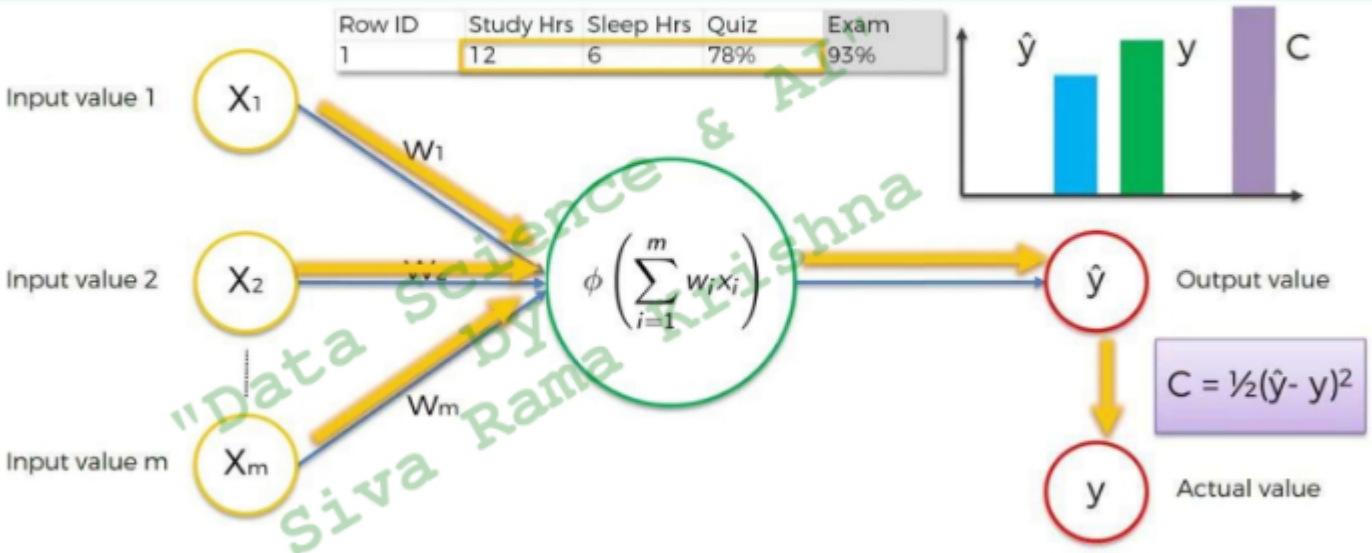
How do Neural Networks  rn?



How do Neural Networks  rn?



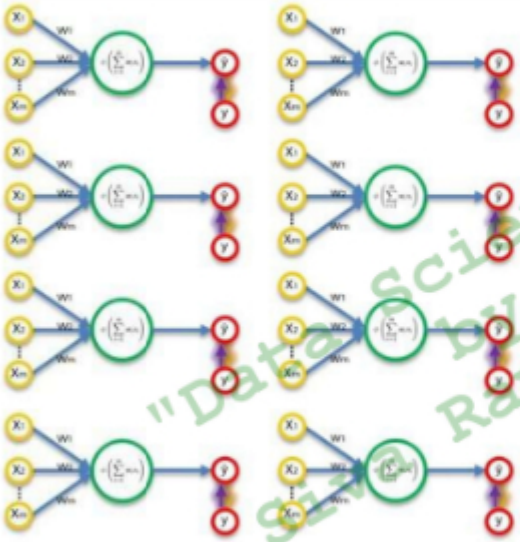
How do Neural Networks  rn?



How do Neural Networks  rn?

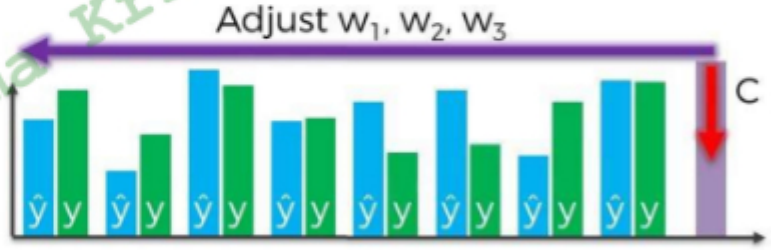
Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%

How do Neural Networks  rn?

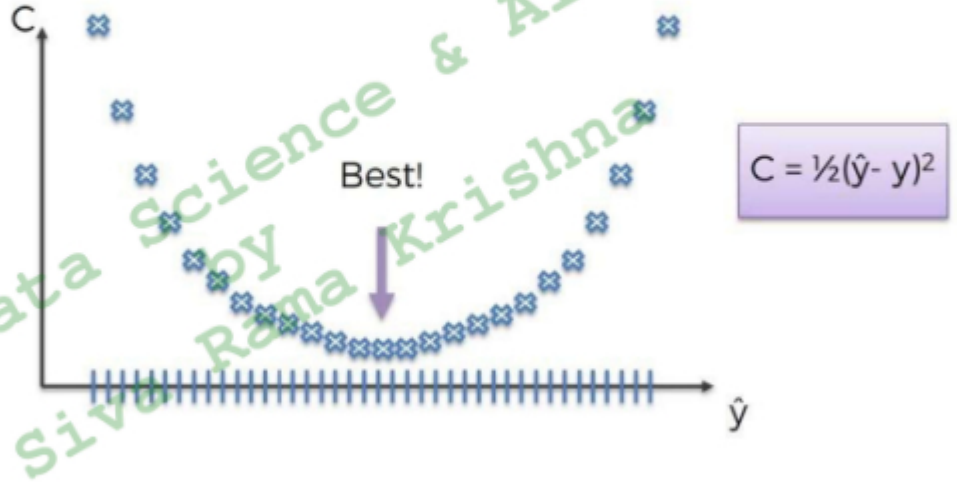


Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%
2	22	6.5	24%	68%
3	115	4	100%	95%
4	31	9	67%	75%
5	0	10	58%	51%
6	5	8	78%	60%
7	92	6	82%	89%
8	57	8	91%	97%

$$C = \sum \tfrac{1}{2}(\hat{y} - y)^2$$

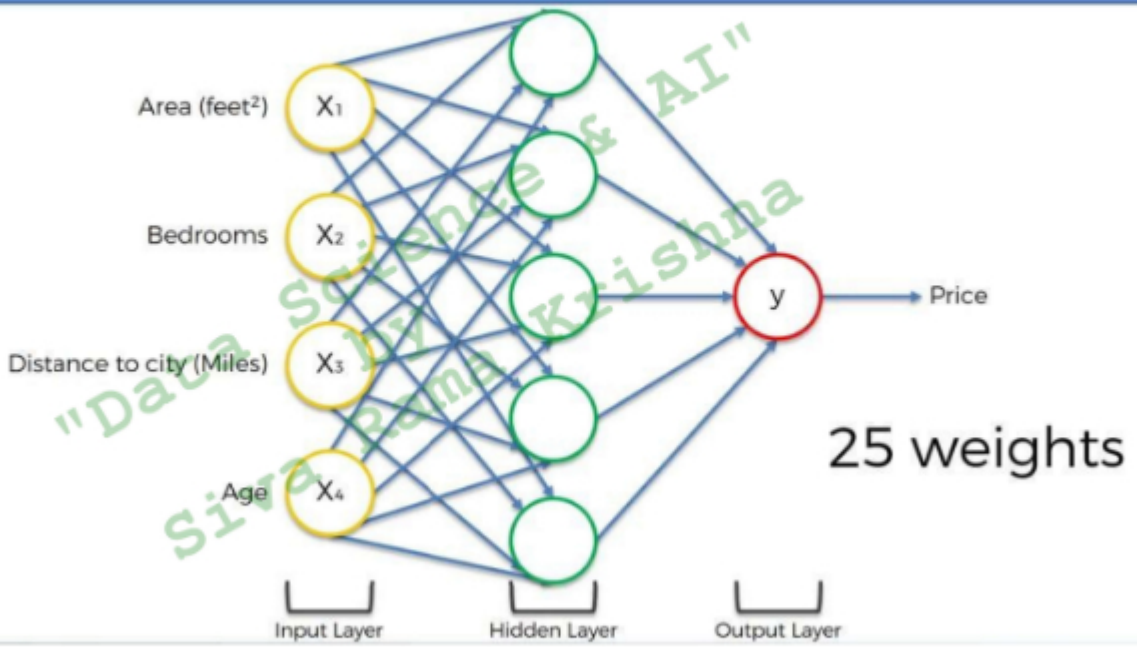


Gradient Descent



$C = \frac{1}{2}(\hat{y} - y)^2$

Gradient Descent



25 weights

Gradient Descent

$$C = \tfrac{1}{2}(\hat{y} - y)^2$$

Gradient Descent

$C = \frac{1}{2}(\hat{y} - y)^2$

Gradient Descent

$C = \frac{1}{2}(\hat{y} - y)^2$

Gradient Descent

$C = \frac{1}{2}(\hat{y} - y)^2$

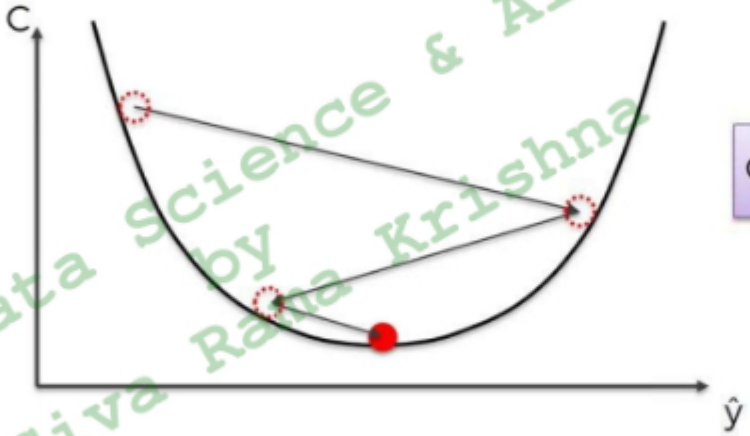
Gradient Descent

$C = \frac{1}{2}(\hat{y} - y)^2$

Gradient Descent

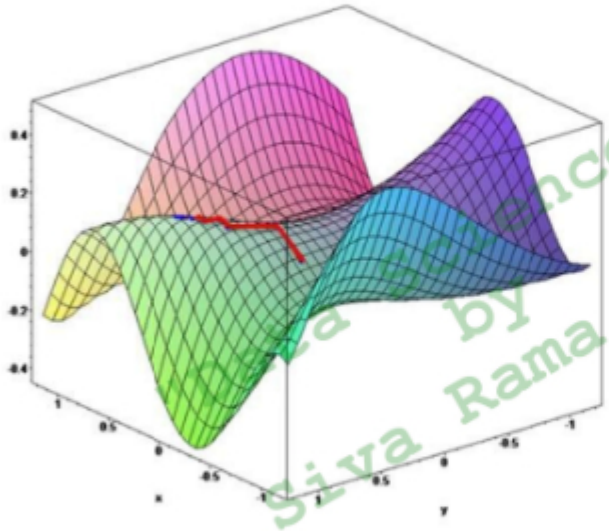
$C = \frac{1}{2}(\hat{y} - y)^2$

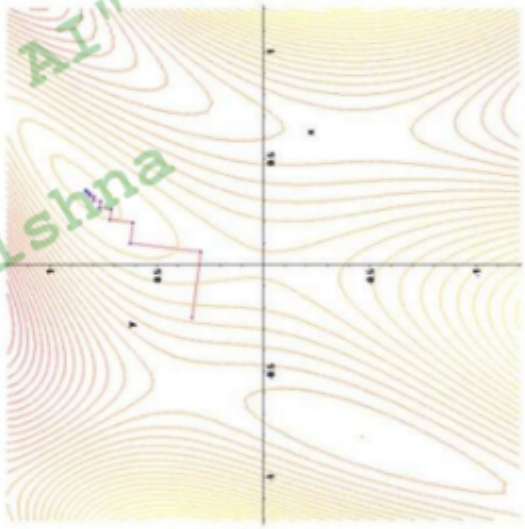
Gradient Descent



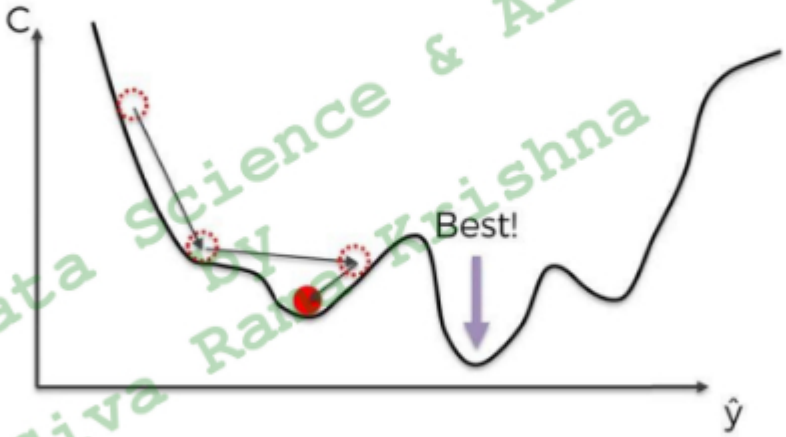
$C = \frac{1}{2}(\hat{y} - y)^2$

Gradient Descent

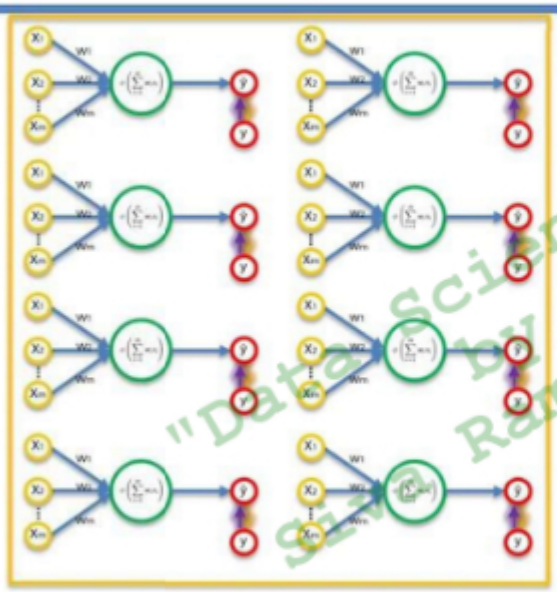




Stochastic Gradient Desc



Stochastic Gradient Desc



Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%
2	22	6.5	24%	68%
3	115	4	100%	95%
4	31	9	67%	75%
5	0	10	58%	51%
6	5	8	78%	60%
7	92	6	82%	89%
8	57	8	91%	97%

$$C = \sum \tfrac{1}{2}(\hat{y} - y)^2$$

Stochastic Gradient Desc

Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%
2	22	6.5	24%	68%
3	115	4	100%	95%
4	31	9	67%	75%
5	0	10	58%	51%
6	5	8	78%	60%
7	92	6	82%	89%
8	57	8	91%	97%

Upd w's

Batch Gradient Descent

Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%
2	22	6.5	24%	68%
3	115	4	100%	95%
4	31	9	67%	75%
5	0	10	58%	51%
6	5	8	78%	60%
7	92	6	82%	89%
8	57	8	91%	97%

Upd w's
Upd w's
Upd w's
Upd w's
Upd w's
Upd w's
Upd w's
Upd w's

Stochastic Gradient Descent

Gradient Descent

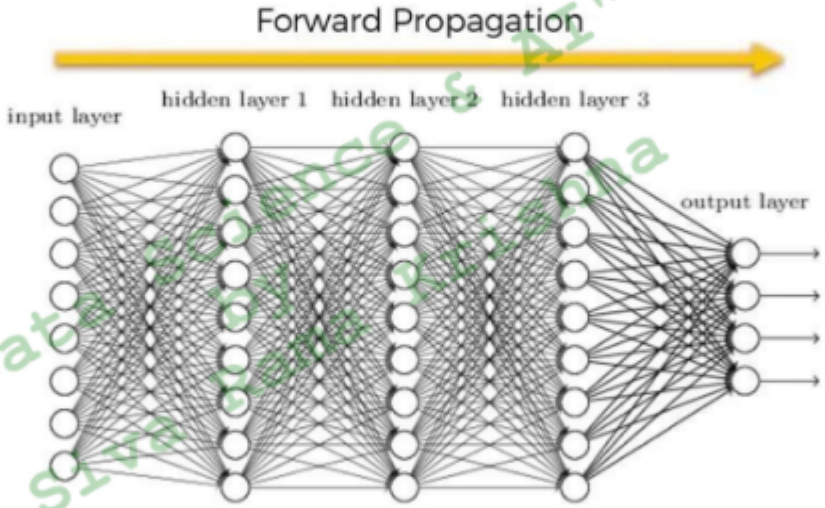


Image Source: neuralnetworksanddeeplearning.com

Gradient Descent

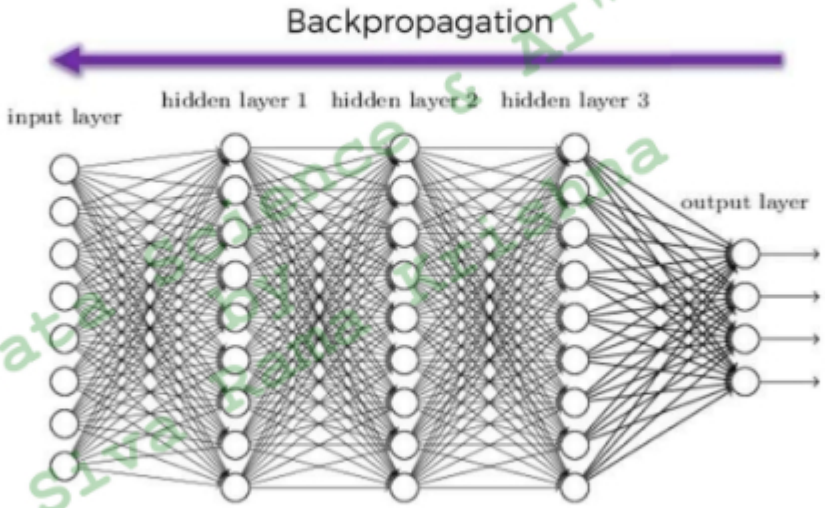


Image Source: neuralnetworksanddeeplearning.com