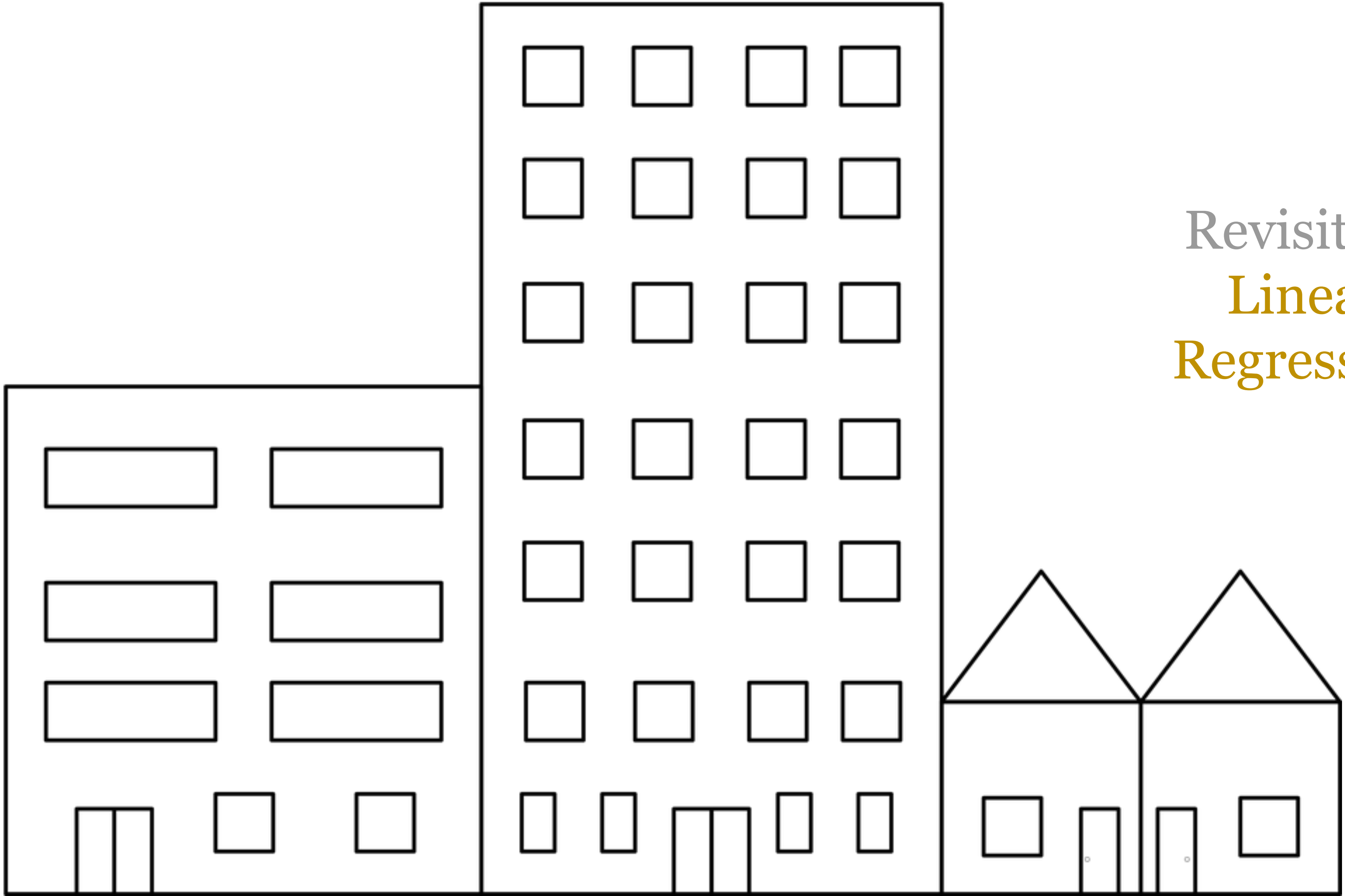
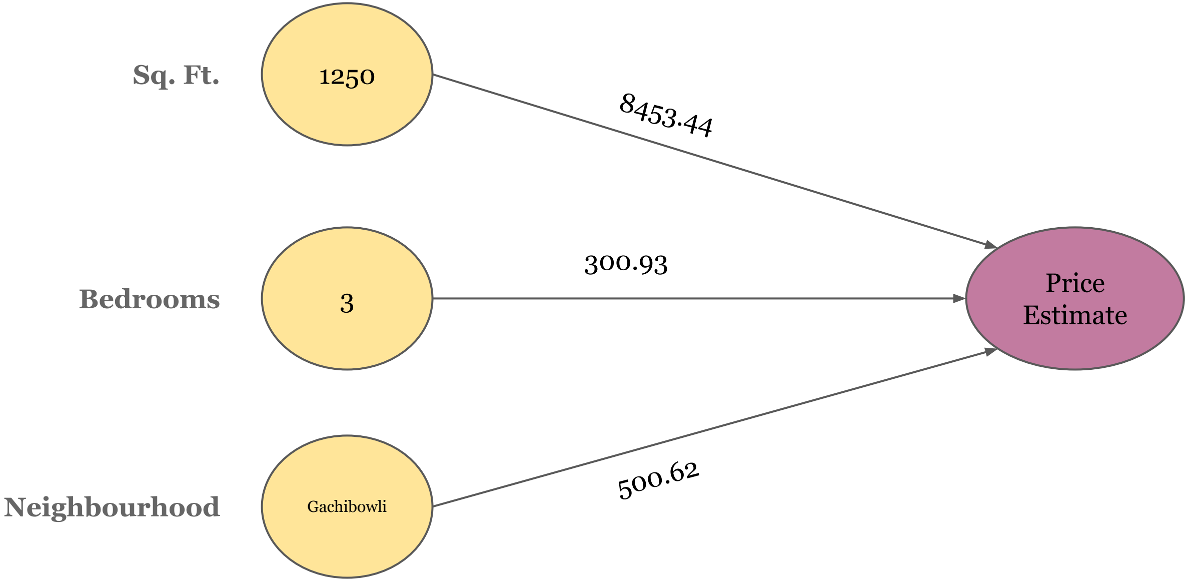


Revisiting Linear Regression

Sq. Ft	Neighbourhood	Bedrooms	Price ('000)
2000	Gachibowli	3	180
1750	Jubilee Hills	3	210
1100	Kukatpally	2	55
900	Gachibowli	2	72
1245	KPHB	3	60
...	...	...	...
...	...	...	...

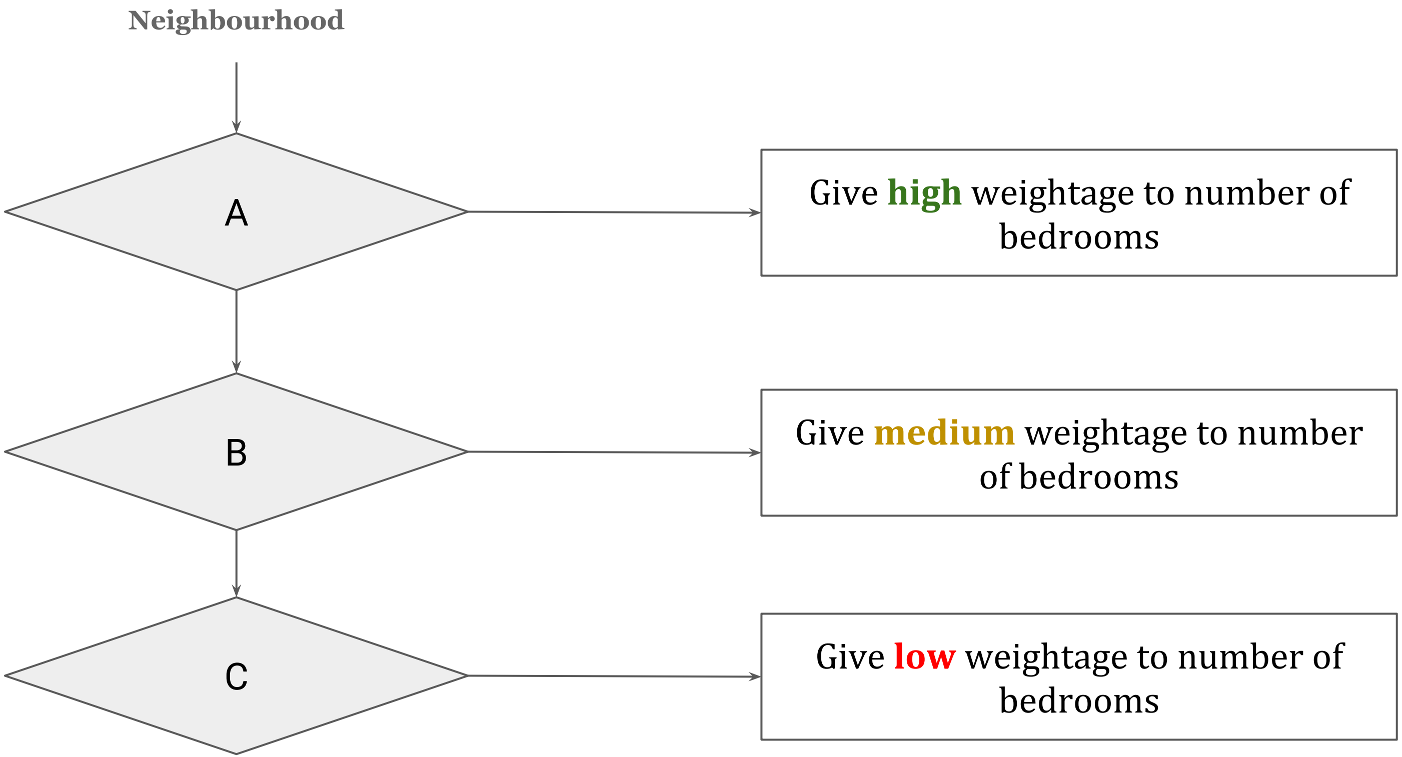
Sq. Ft	Neighbourhood	Bedrooms	Price ('000)
1250	Gachibowli	3	???



$$y = w_1 x_1 + w_2 x_2 + w_3 x_3 + b$$

Not this simple!!!

There are usually lots of Ifs and buts...



Neighbourhood

A

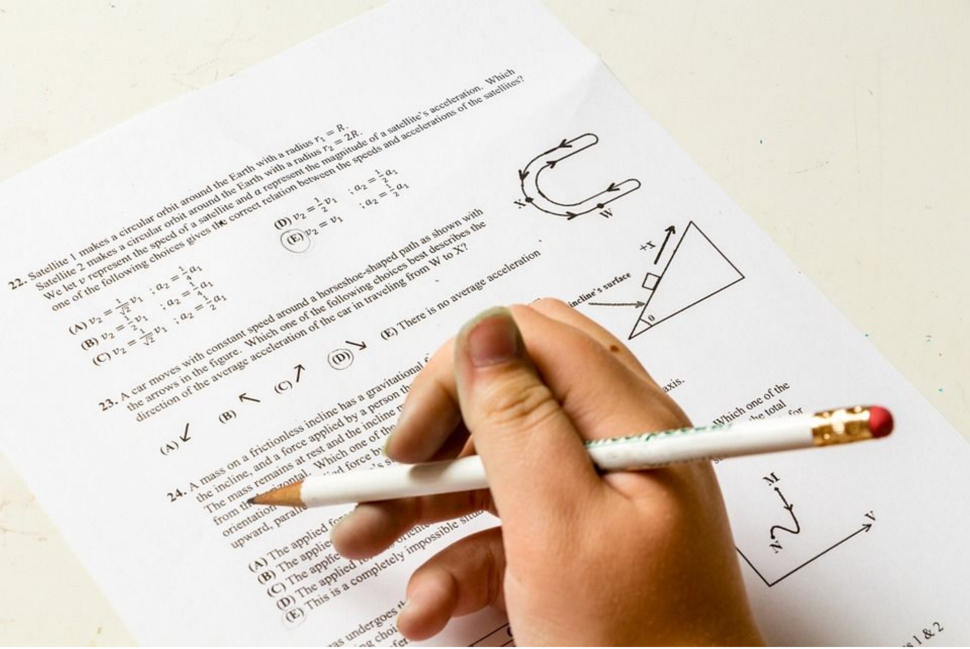
B

C

Give **high** weightage to number of bedrooms

Give **medium** weightage to number of bedrooms

Give **low** weightage to number of bedrooms



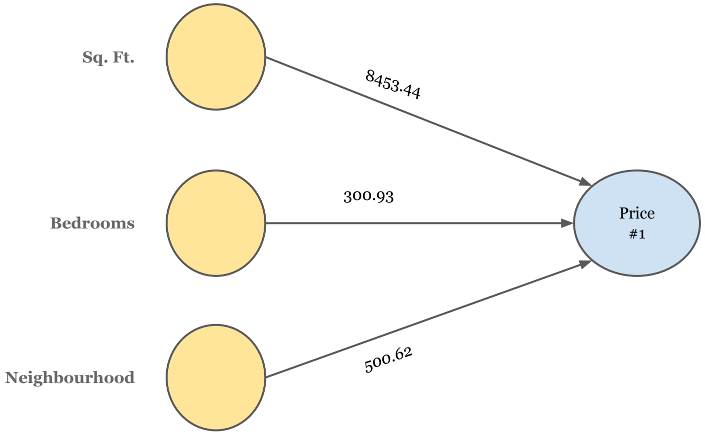
22. Satellite 1 makes a circular orbit around the Earth with a radius $r_1 = R$.
Satellite 2 makes a circular orbit around the Earth with a radius $r_2 = 2R$.
We let v represent the speed of a satellite and a represent the magnitude of a satellite's acceleration. Which one of the following choices gives the correct relation between the speeds and accelerations of the satellites?

(A) $v_2 = \frac{1}{\sqrt{2}}v_1$; $a_2 = \frac{1}{4}a_1$
(B) $v_2 = \frac{1}{2}v_1$; $a_2 = \frac{1}{4}a_1$
(C) $v_2 = \frac{1}{\sqrt{2}}v_1$; $a_2 = \frac{1}{2}a_1$
(D) $v_2 = \frac{1}{2}v_1$; $a_2 = \frac{1}{2}a_1$
[✓ (E)] $v_2 = v_1$; $a_2 = \frac{1}{2}a_1$

23. A car moves with constant speed around a horseshoe-shaped path as shown with the arrows in the figure. Which one of the following choices best describes the direction of the average acceleration of the car in traveling from W to X?

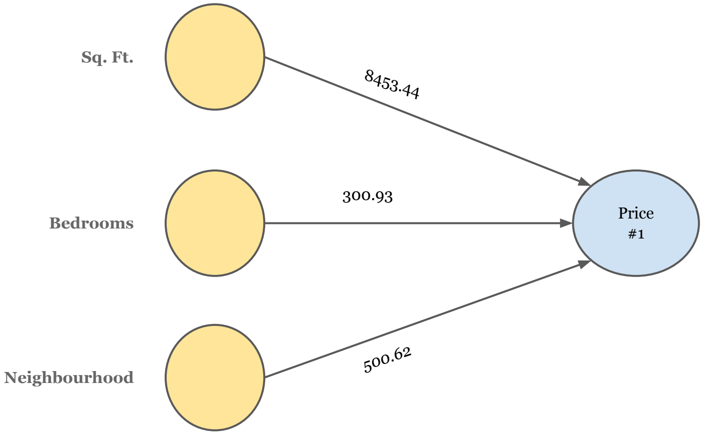
(A) ↙ (B) ↖ (C) ↗ [✓ (D)] ↘ (E) There is no average acceleration

How do we capture complex logic?

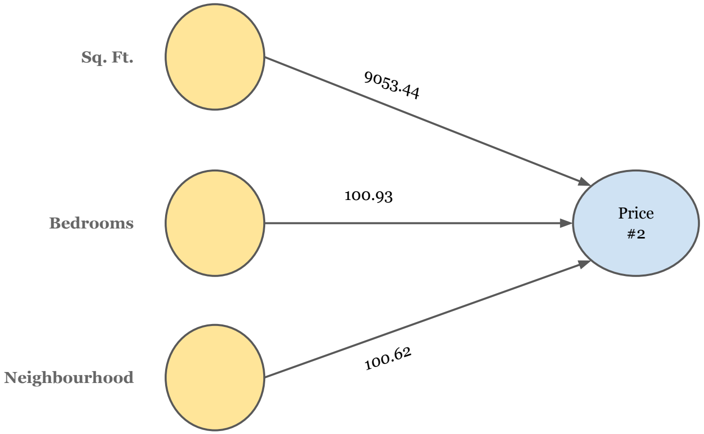


First Set of Weights

Learn some logic in data



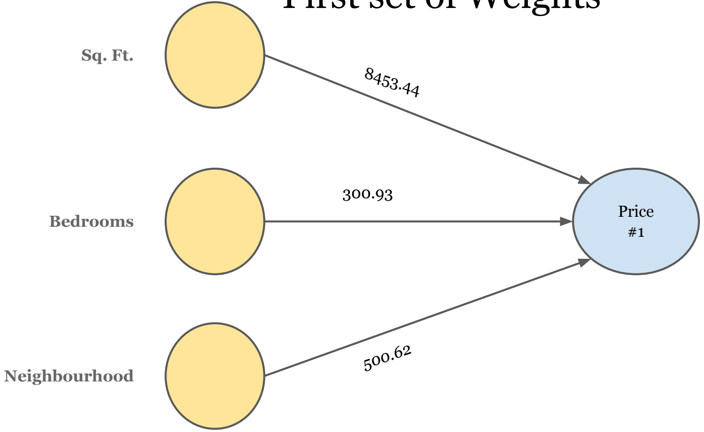
First Set of Weights



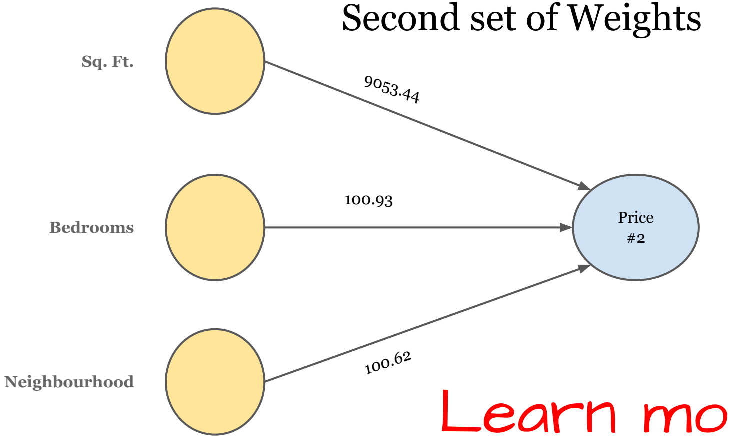
Second Set of Weights

Learn ANOTHER
logic in data

First set of Weights

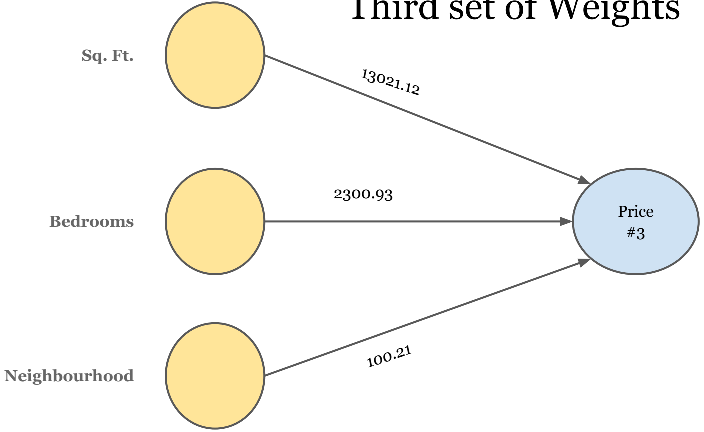


Second set of Weights

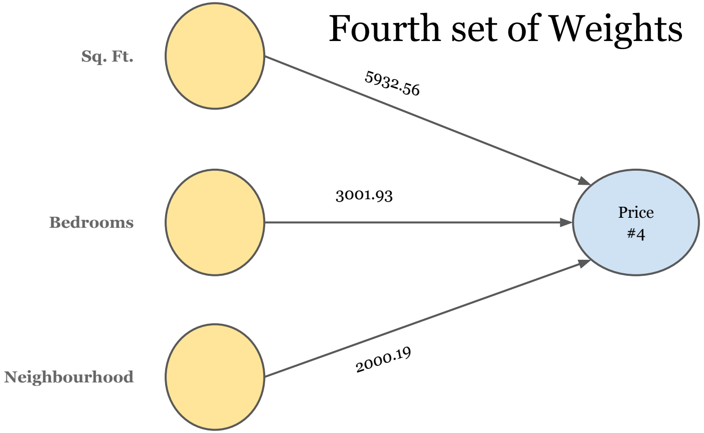


Learn more and more

Third set of Weights

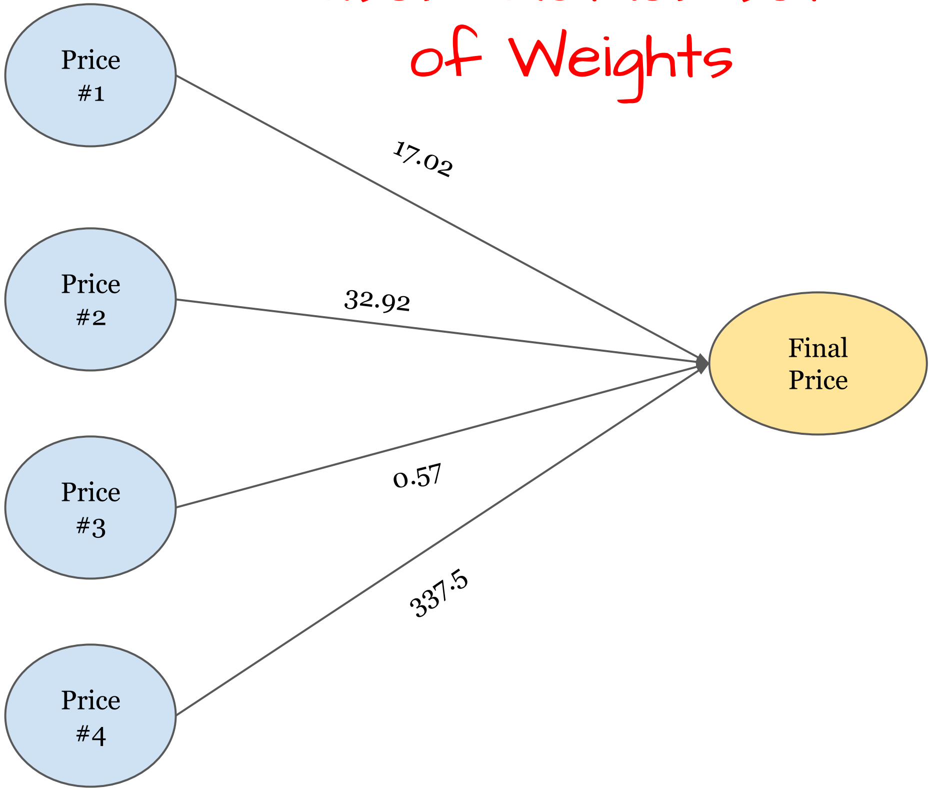


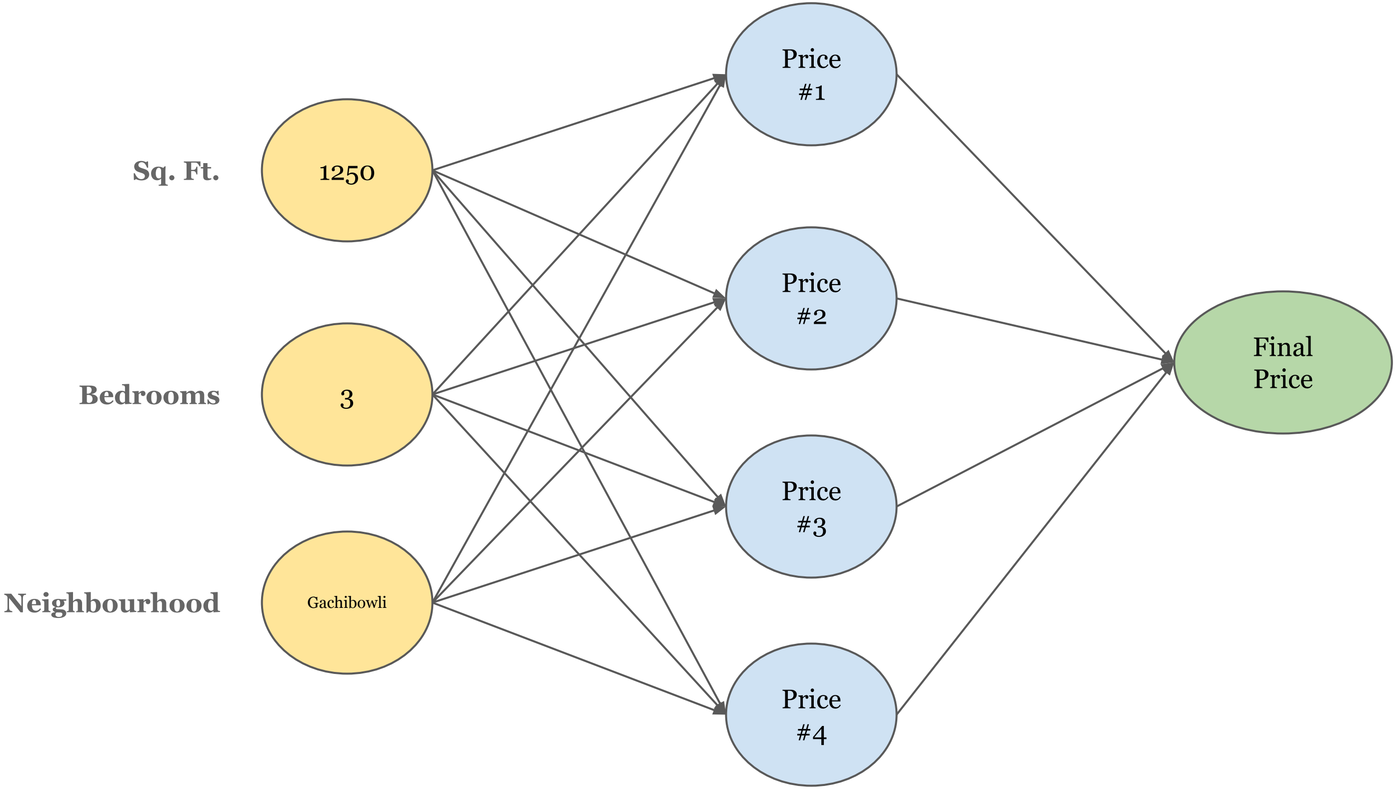
Fourth set of Weights



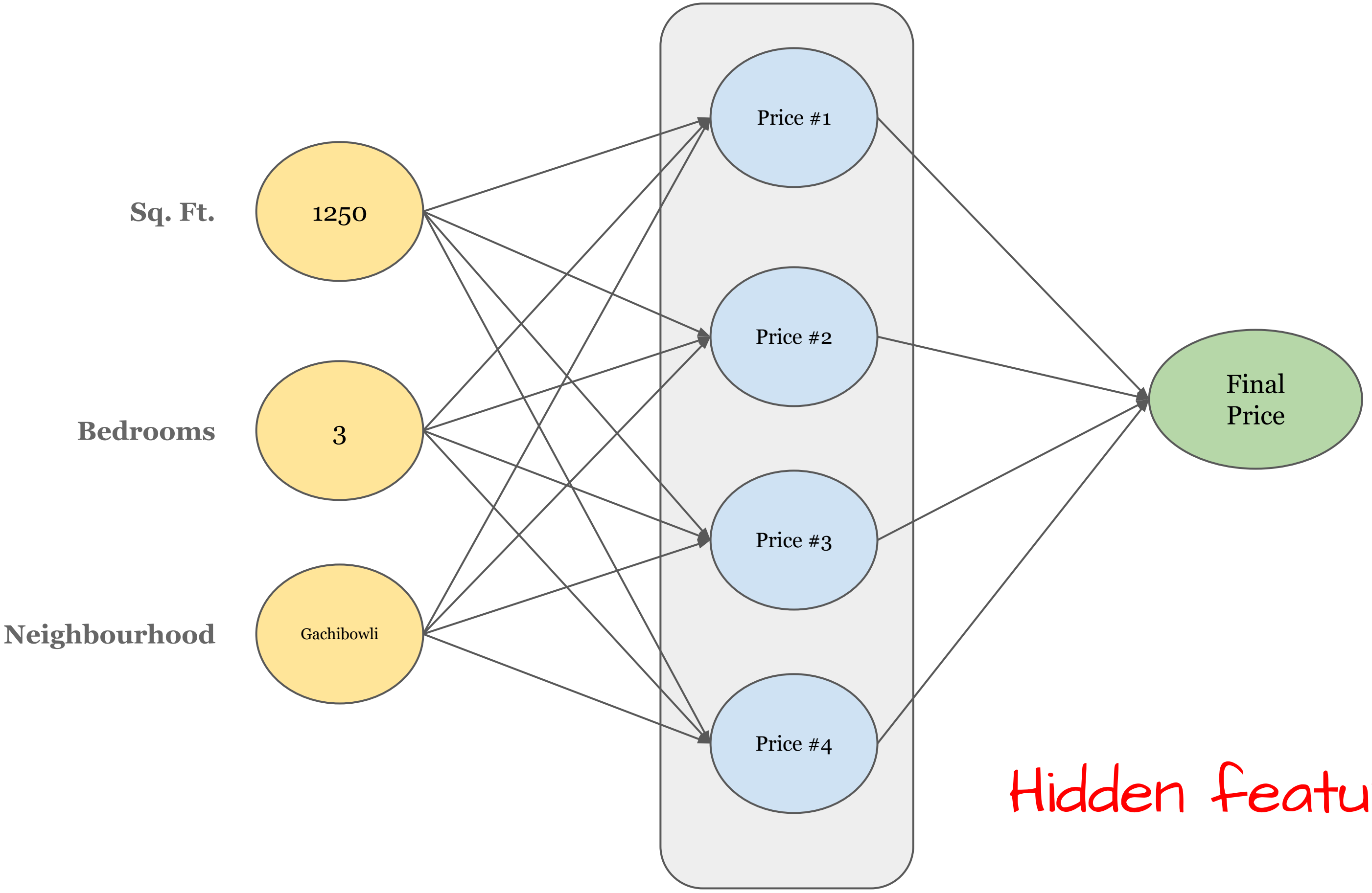
What to do with all the Prices?

User another set of Weights




Sq. Ft.
1250
Bedrooms
3
Neighbourhood
Gachibowli
Price #1
Price #2
Price #3
Price #4
Final Price

What do Price# 1, 2, 3 and 4 represent?



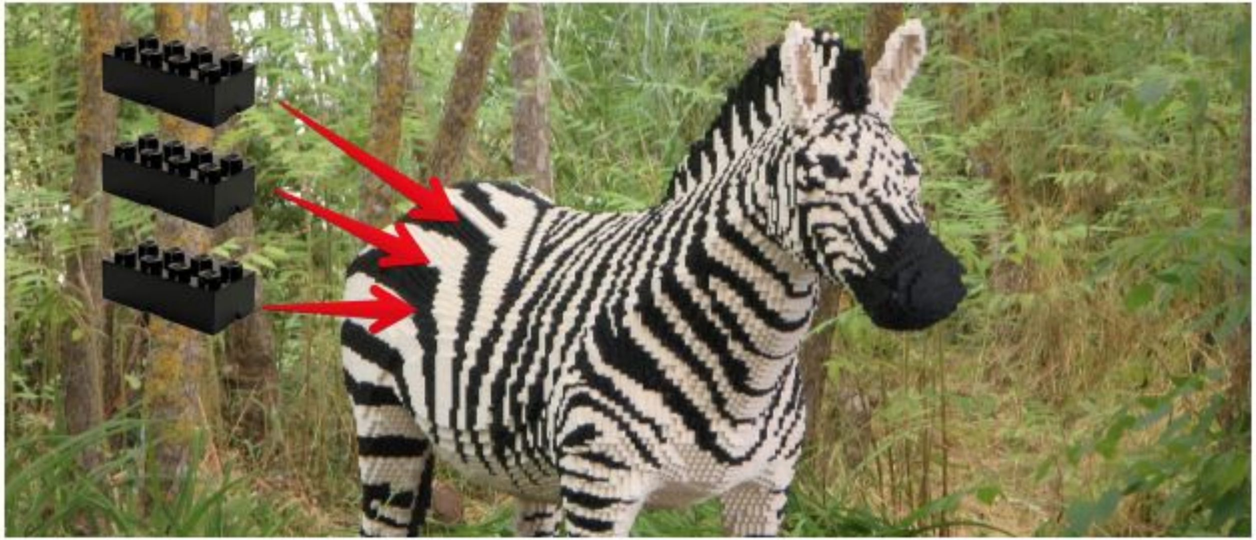
Hidden features

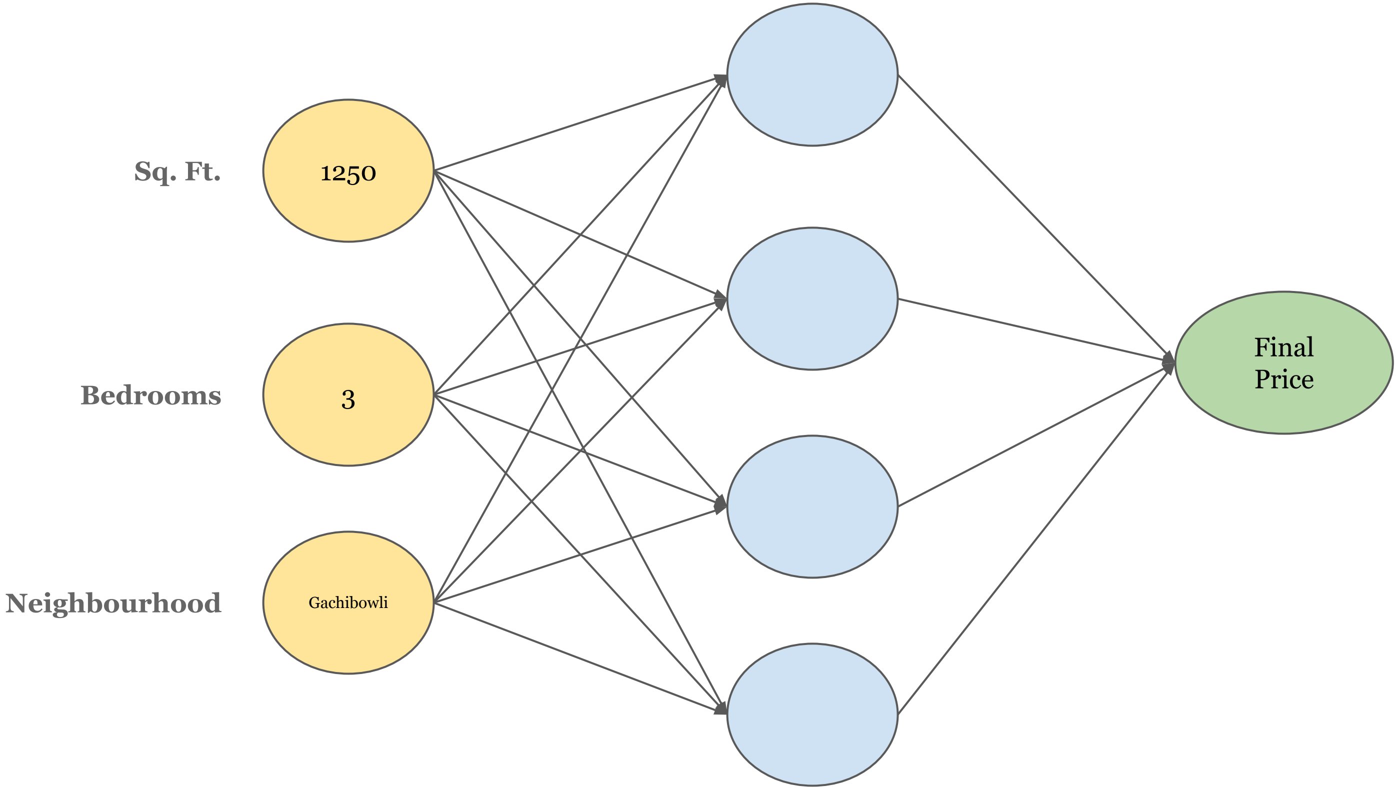
What are these hidden features?

What are these hidden features?

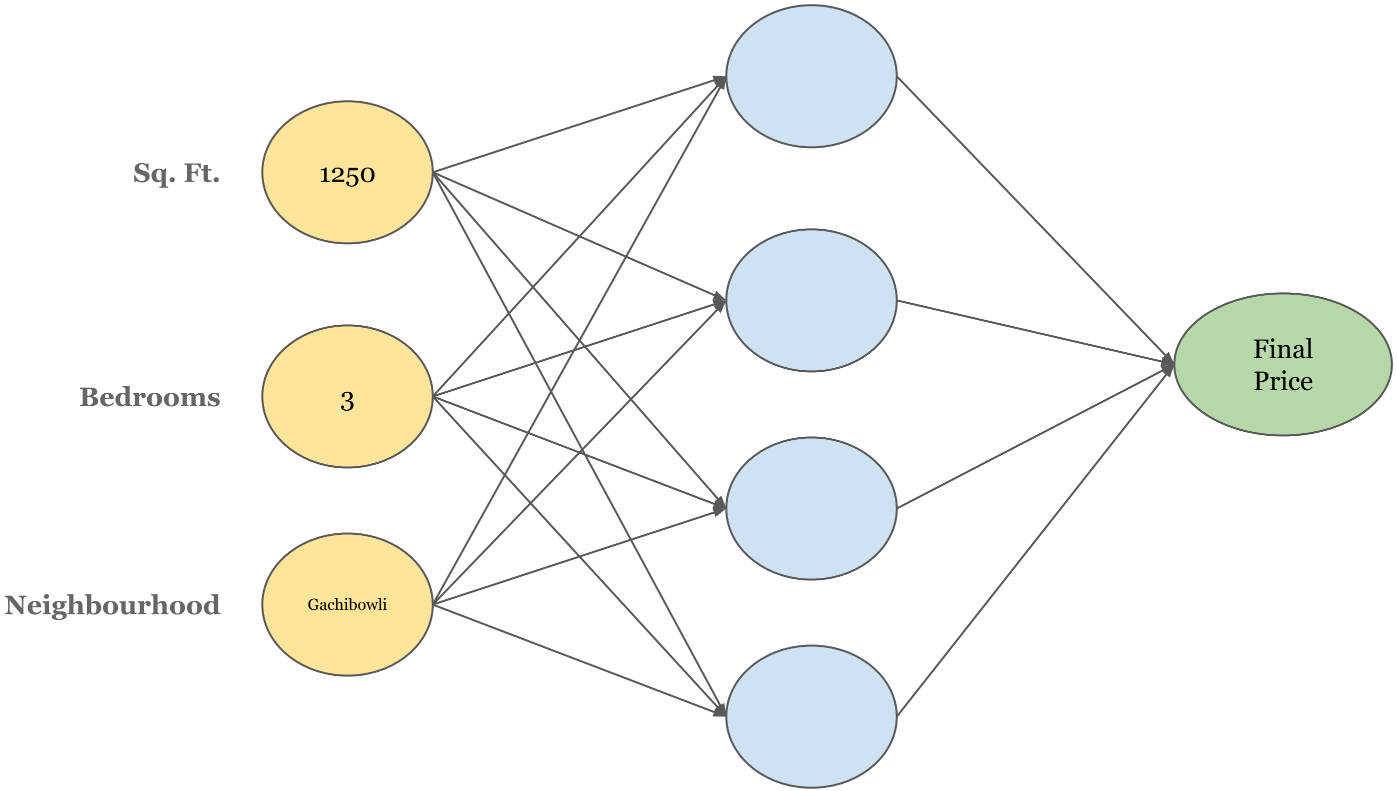
We, humans,   do
not really
understand them

It's like 'Lego'

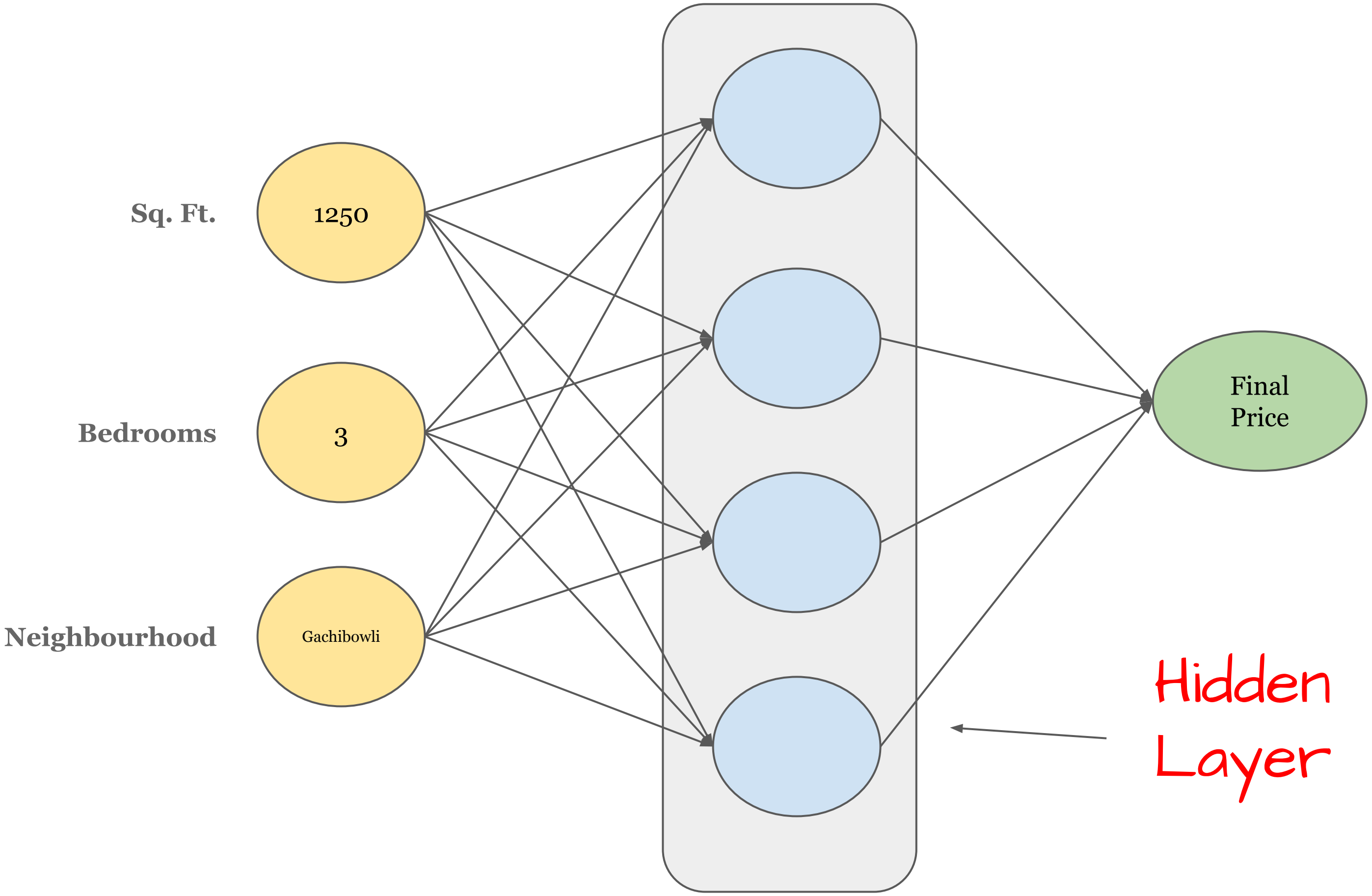




What is this structure?



Neural Network



Sq. Ft.

1250

Bedrooms

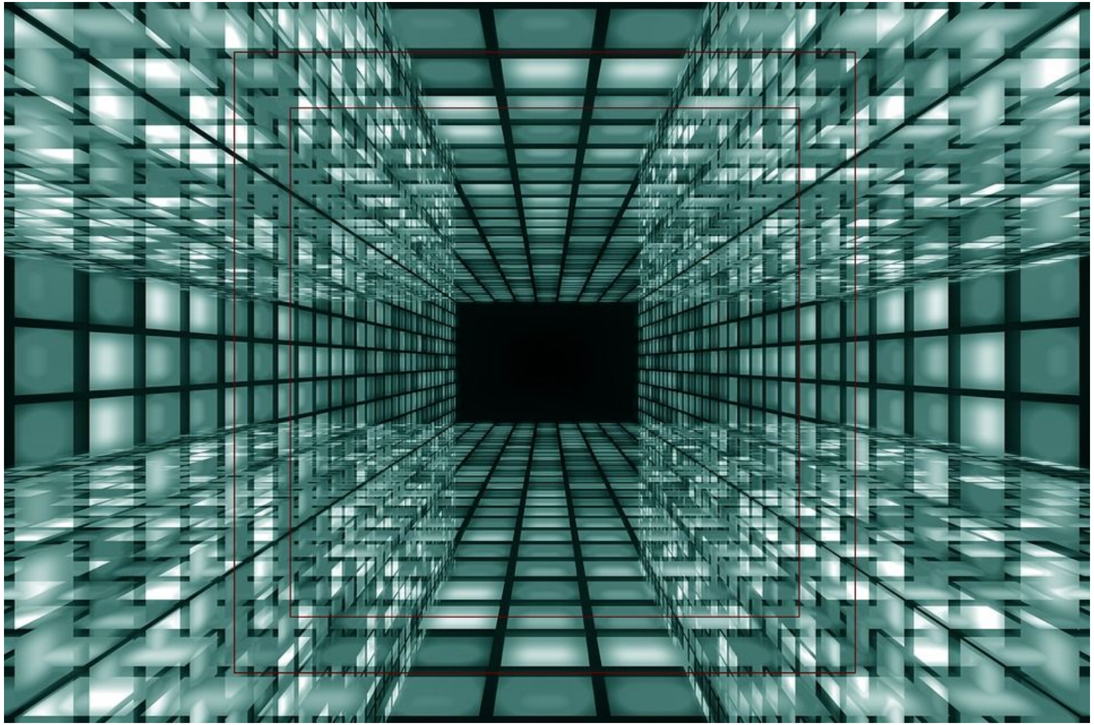
3

Neighbourhood

Gachibowli

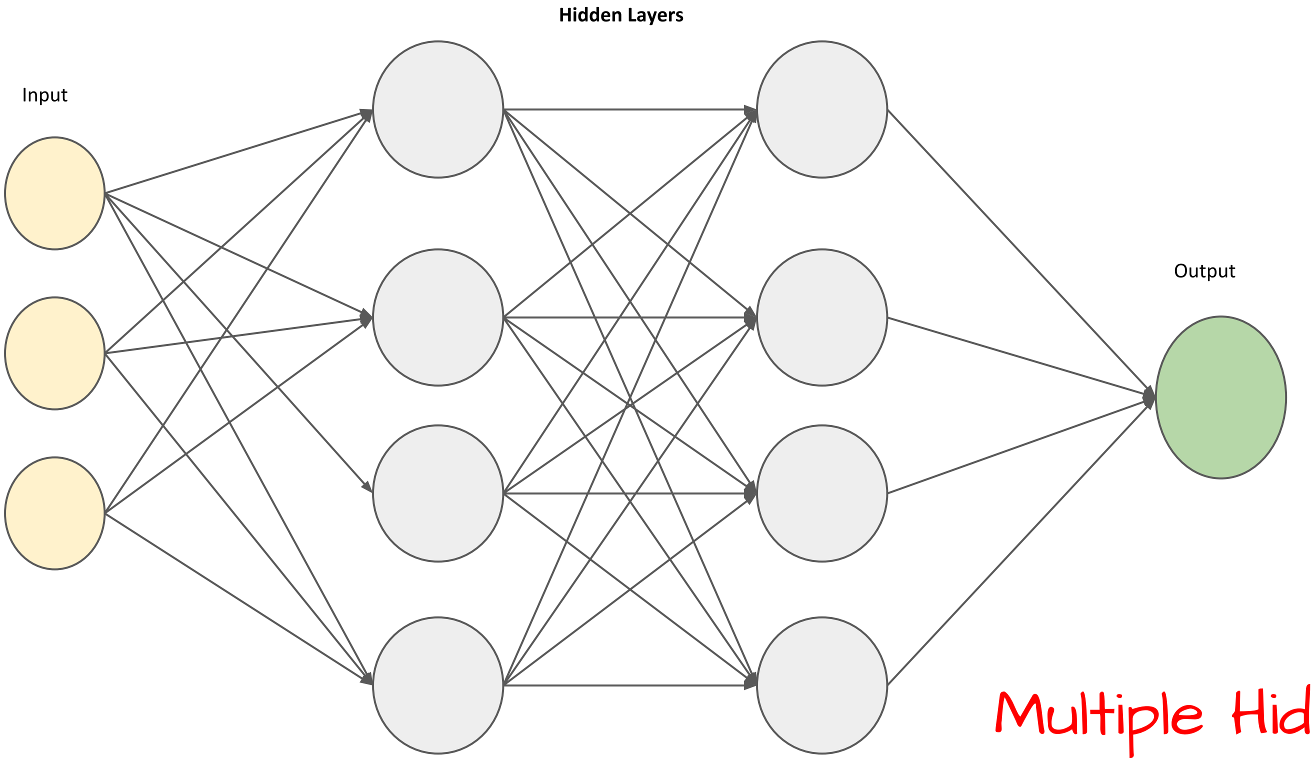
Final
Price

Hidden
Layer



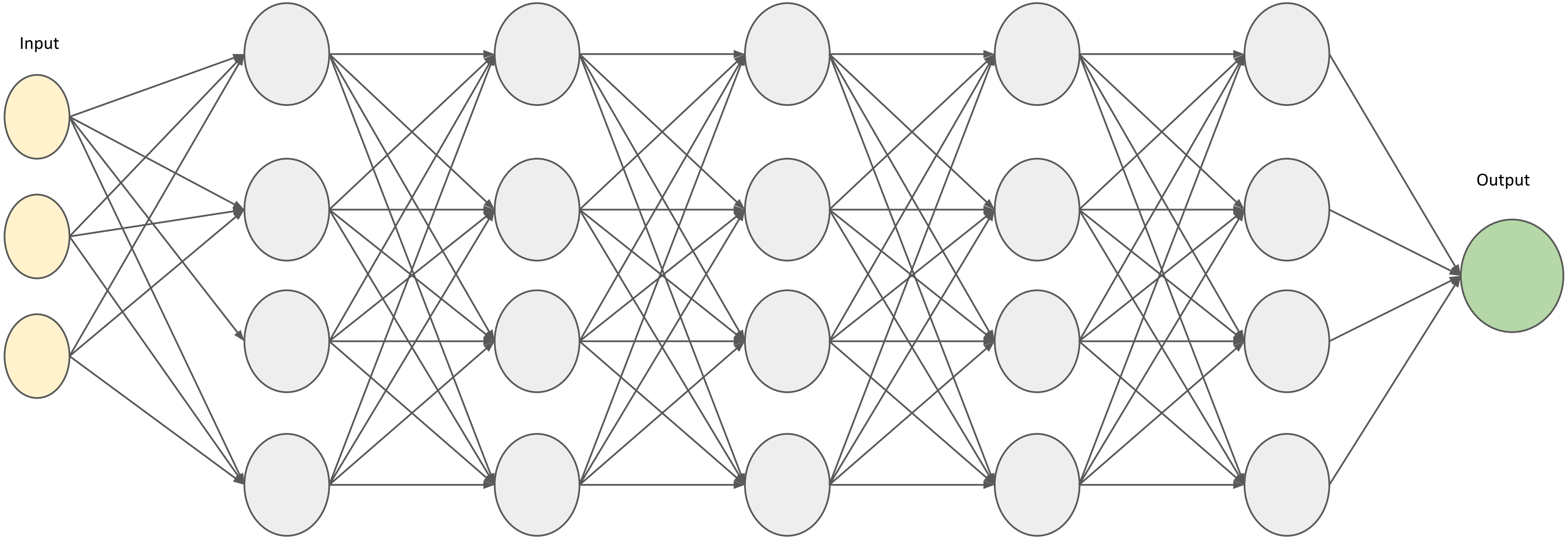
What is
Deep Learning

Deep Neural Network



Multiple Hidden Layers

Deep Neural Network

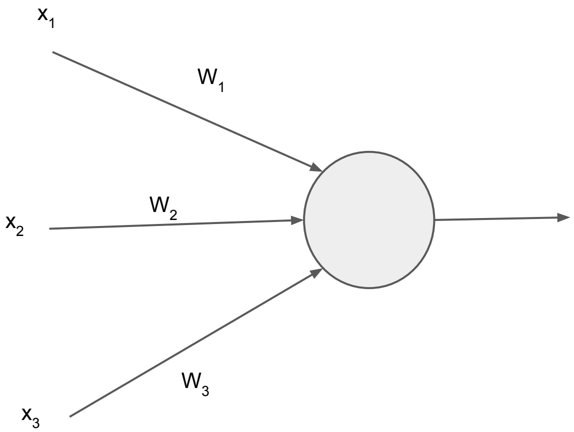


Hidden Layers

Adding non-linearity to Decision making

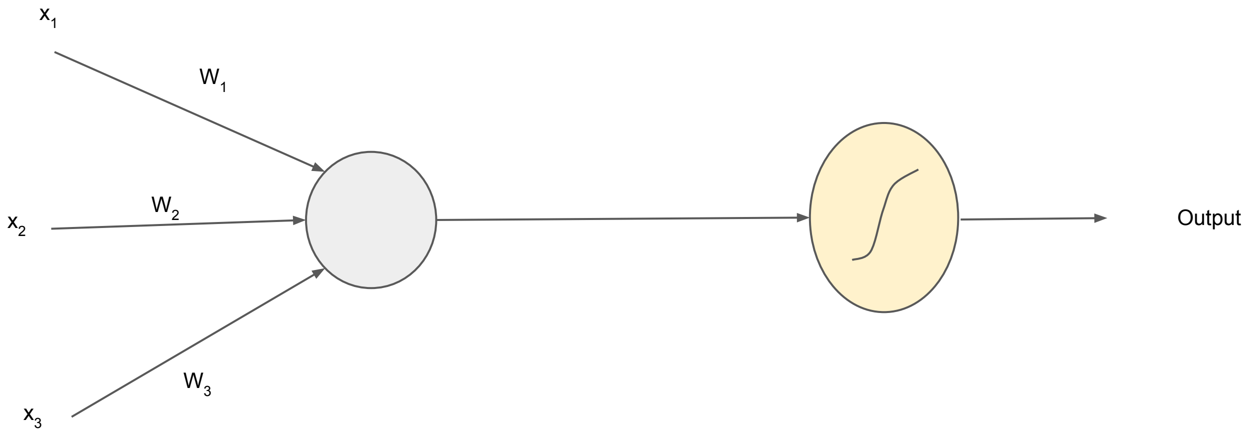
Activation function

Neuron output

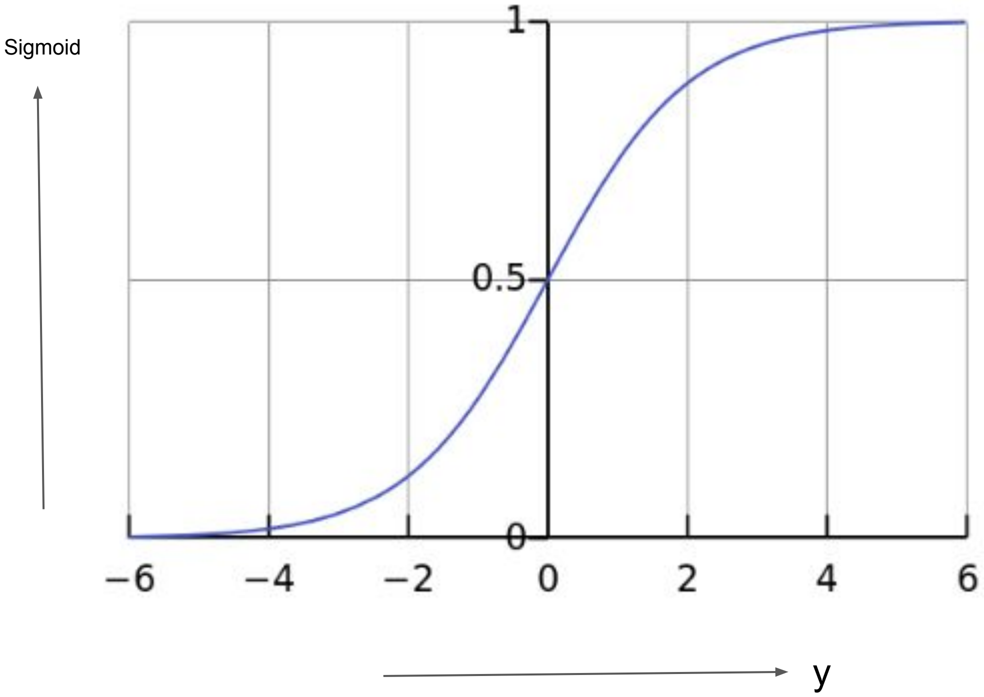


$$y = w_1x_1 + w_2x_2 + w_3x_3 + b$$

Controlling Neuron firing



Sigmoid Function

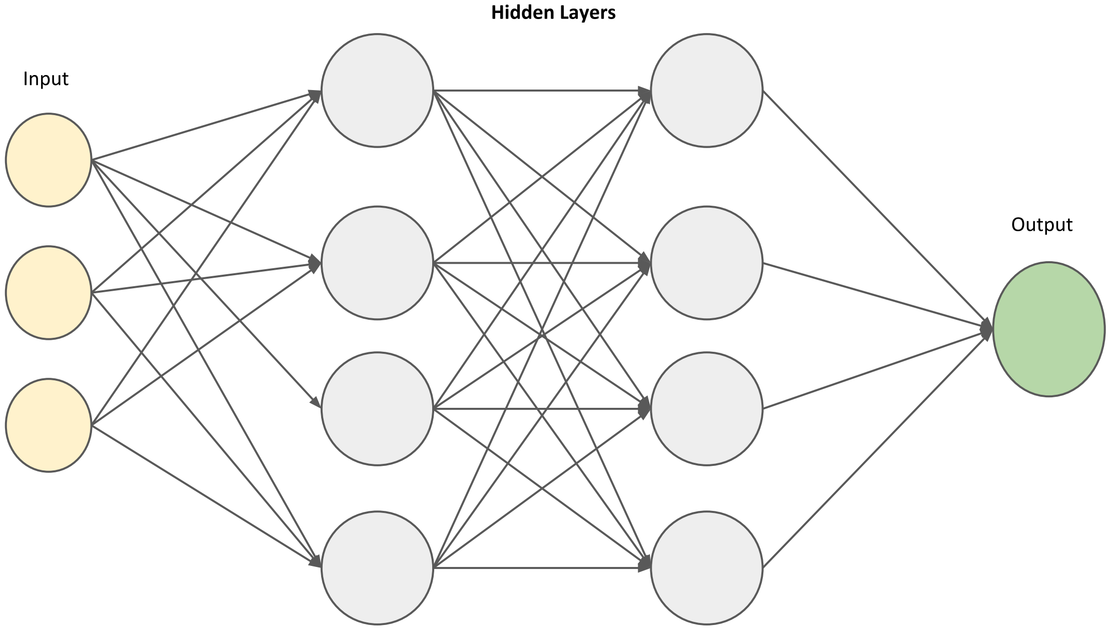


$$sigmoid(y) = \frac{1}{1 + e^{-y}}$$

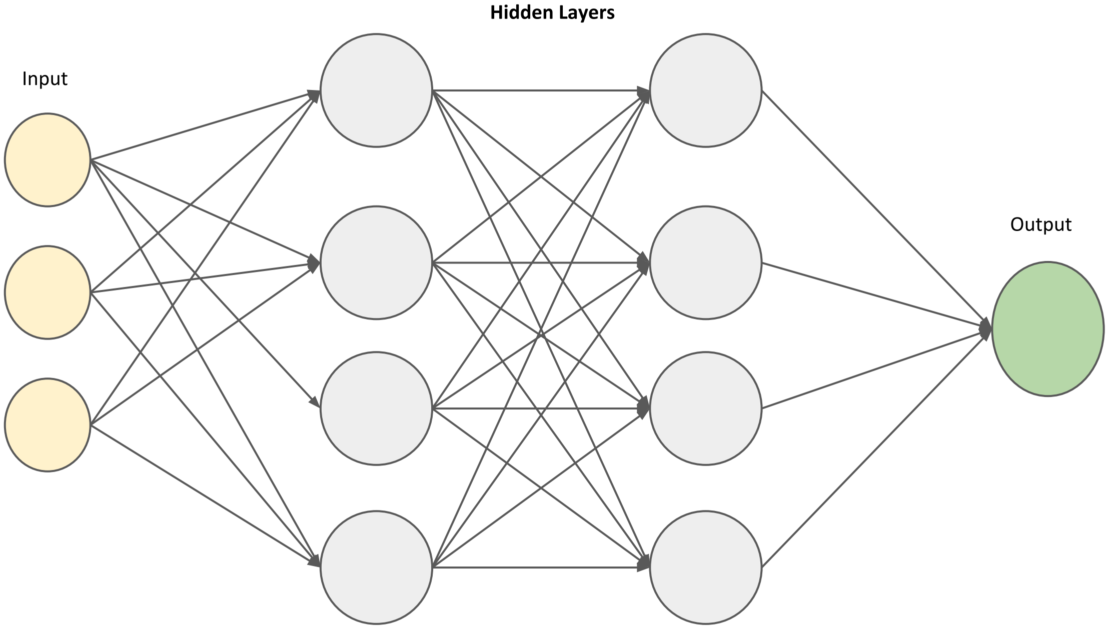
$$y = \sum wx + b$$

How do we update weights?

Number of weights

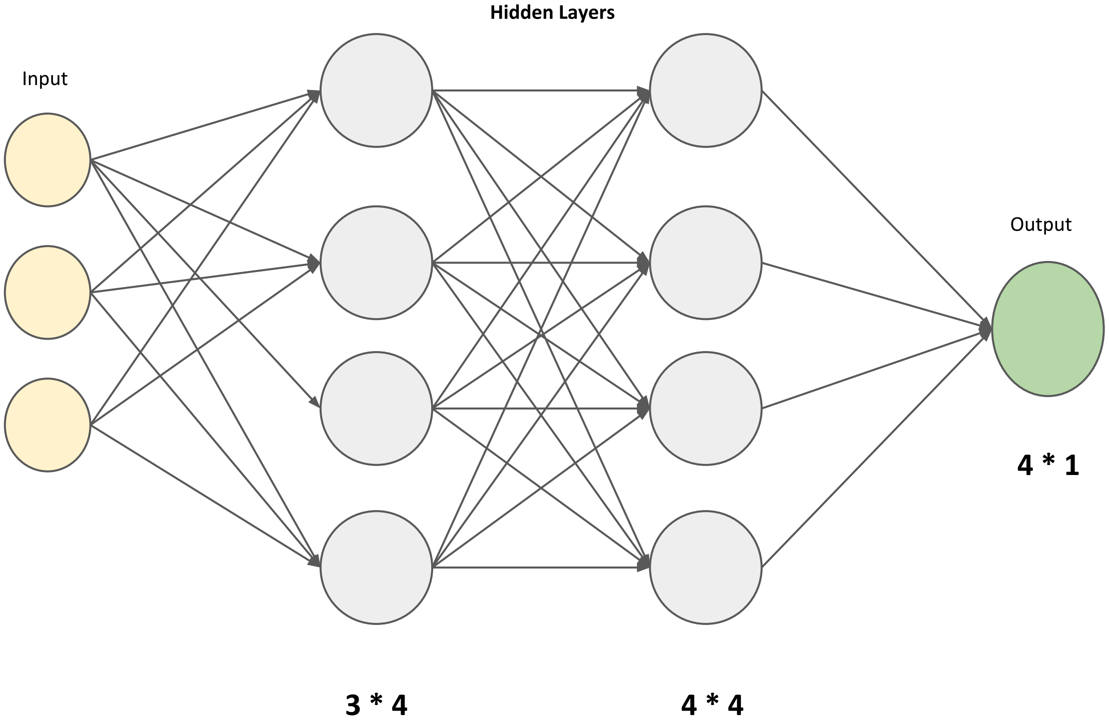


Number of weights



3 * 4

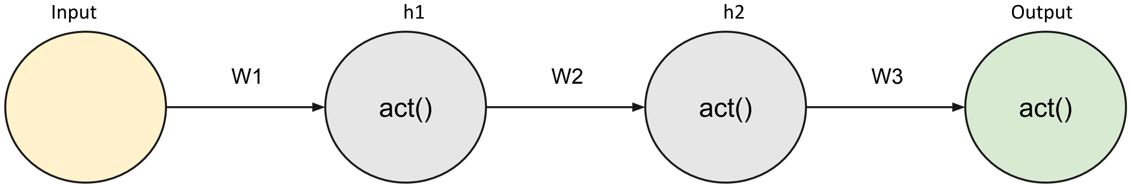
Number of weights



Function Derivative

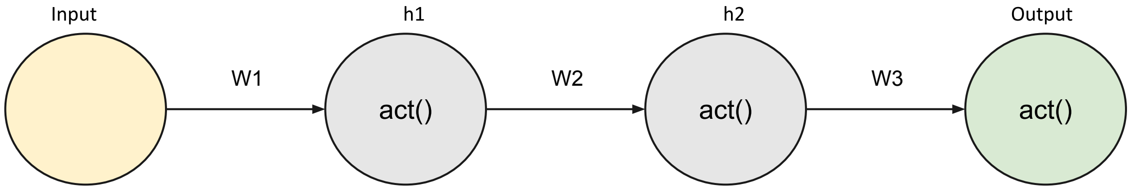
$$w_{new} = w_{old} - \eta \frac{d}{dw} J(w_{old})$$

Output in Neural Network



$$Output = act(w3 * h2)$$

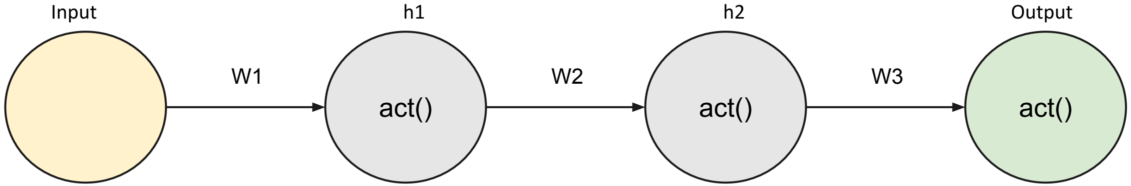
Output in Neural Network



$$Output = act(w3 * h2)$$

$$h2 = act(w2 * h1)$$

Output in Neural Network

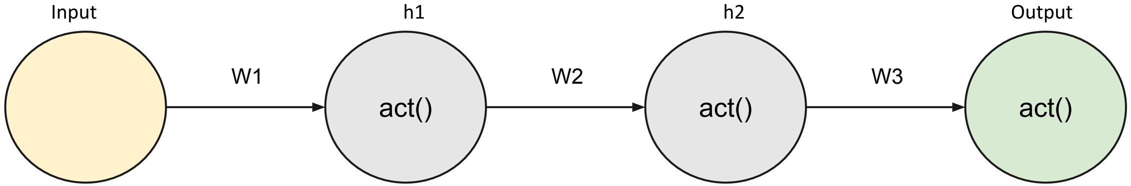


$$Output = act(w3 * h2)$$

$$h2 = act(w2 * h1)$$

$$h1 = act(w1 * input)$$

Output in Neural Network



$$Output = act(w3 * h2)$$

$$h2 = act(w2 * h1)$$

$$h1 = act(w1 * input)$$

$$Output = act(w3 * act(w2 * act(w1 * input)))$$

Derivative of Output function w.r.t w1

Derivative of Output function w.r.t w1

$$Chain\ rule \rightarrow \frac{d\,f(g(x))}{dx} = \frac{df}{dg} * \frac{dg}{dx}$$

Derivative of Output function w.r.t w1

$$\mathit{Chain\ rule} \rightarrow \frac{d\,f(g(x))}{dx} = \frac{df}{dg} * \frac{dg}{dx}$$

$$f(w1) \; = output(h2(h1(w1)))$$

Derivative of Output function w.r.t w1

$$\mathit{Chain\ rule} \rightarrow \frac{d\,f(g(x))}{dx} = \frac{df}{dg} * \frac{dg}{dx}$$

$$f(w1) = output(h2(h1(w1)))$$

$$\frac{d\,f(w1)}{d\,w1} = \frac{d\,output}{d\,h2} * \frac{d\,h2}{d\,h1} * \frac{d\,h1}{d\,w1}$$

Loss function w.r.t Weights

Loss function w.r.t Weights

$$\frac{d\,LOSS}{d\,w3} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,w3}$$

Loss function w.r.t Weights

$$\frac{d\,LOSS}{d\,w2} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,h2} * \frac{d\,h2}{d\,w2}$$

$$\frac{d\,LOSS}{d\,w3} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,w3}$$

Loss function w.r.t Weights

$$\frac{d\,LOSS}{d\,w1} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,h2} * \frac{d\,h2}{d\,h1} * \frac{d\,h1}{d\,w1}$$

$$\frac{d\,LOSS}{d\,w2} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,h2} * \frac{d\,h2}{d\,w2}$$

$$\frac{d\,LOSS}{d\,w3} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,w3}$$

Loss function w.r.t Weights

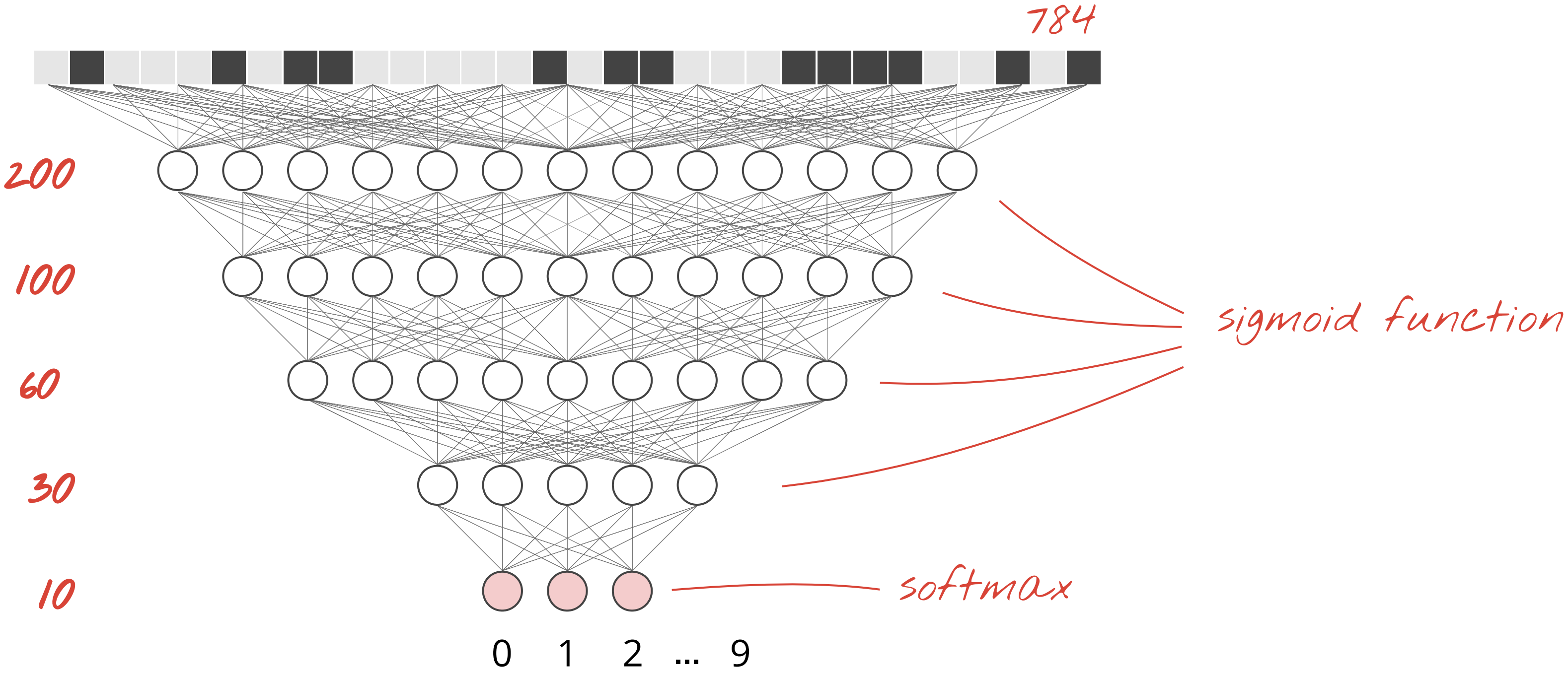
$$\frac{d\,LOSS}{d\,w1} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,h2} * \frac{d\,h2}{d\,h1} * \frac{d\,h1}{d\,w1}$$

$$\frac{d\,LOSS}{d\,w2} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,h2} * \frac{d\,h2}{d\,w2}$$

$$\frac{d\,LOSS}{d\,w3} = \frac{d\,Loss}{d\,output} * \frac{d\,output}{d\,w3}$$

Backpropagation Algorithm

Applying Deep Learning on MNIST


784
200
100
60
30
10
sigmoid function
softmax
0
1
2
...
9