Weights & Biases

30 November 2024 20:39

Linear Regression

weight | coefficients and bias (intercept)

(Radio analogy)

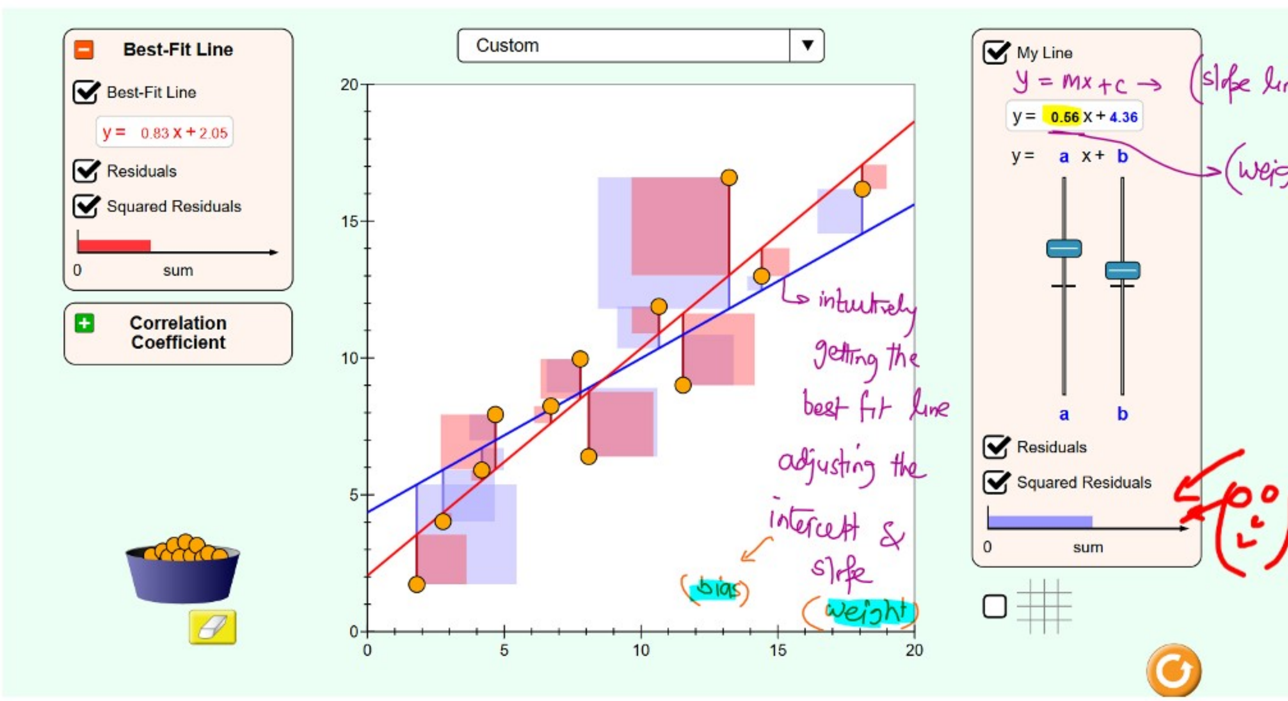
Weights and Biases

* Linear Regression → first algorithm ✓

$$y = \left(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_n x_n\right)$$

intercept (constant)

coefficients



[COLORADO APP to demo Linear Regression]

Imitation Game

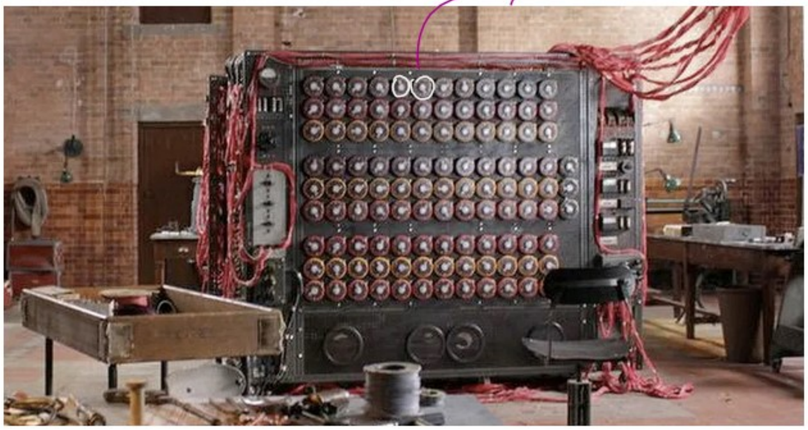
↳ Alan Turing

↳ (ENIGMA) → Encrypt

→ Decrypt.

(knobs & dials)

(Knobs & dials)



Weights

$x_1 \mid x_2 \mid x_3 \mid x_4$

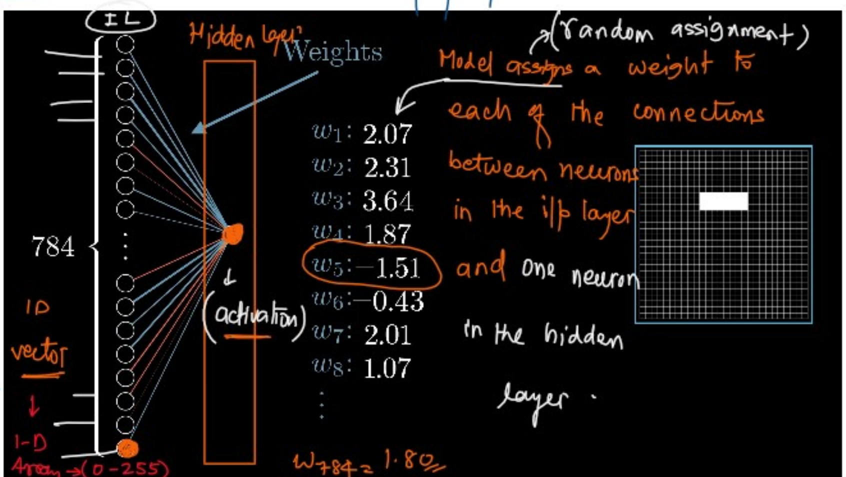
2

← 28 →

28*28=784

1D vector

$x_1, x_2, \ldots x_{784}$



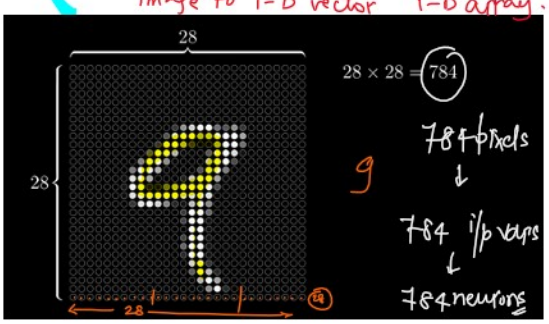
$\sum x_i w_i + b \longrightarrow \boxed{AF} \longrightarrow$ activation

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$

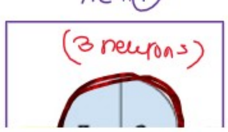
$y_{pred} - y_{actual}$ → error → dL

GDA

Flattening the below pixelated image to 1-D vector 1-D array.

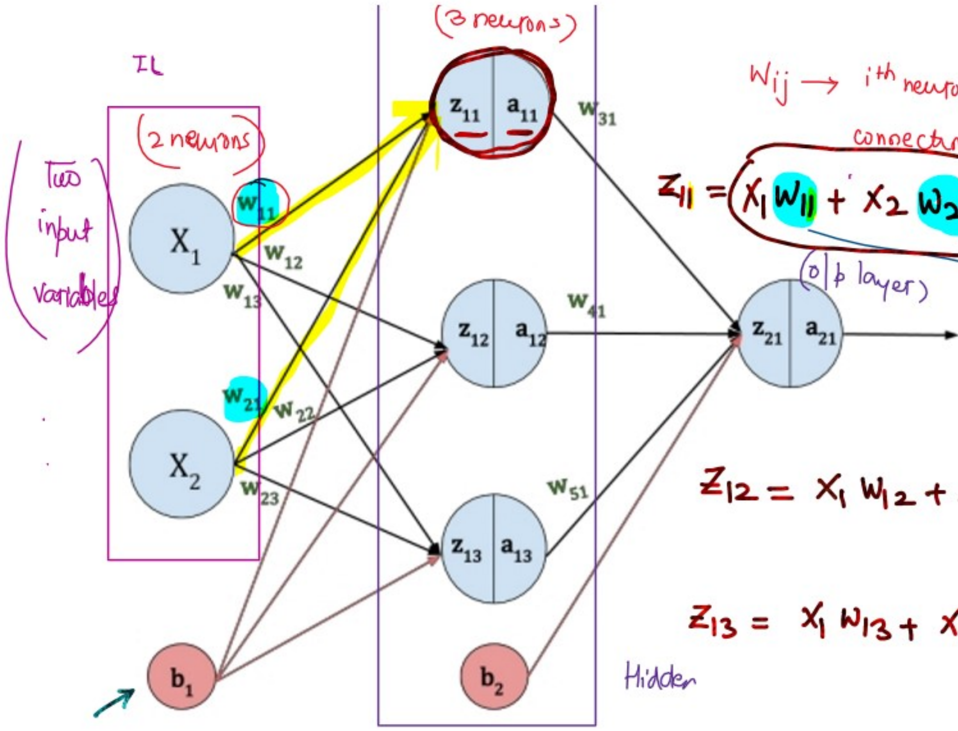


HL#1



IL

$W_{ij} \rightarrow$ i^{th} neuron in IL



$W_{ij} \to$ i^{th} neuron in IL connecting to j^{th} neuron in HL.

$$Z_{11} = x_1 w_{11} + x_2 w_{21} + b_1$$

→ bias

(weights)

(o/p layer)

$$Z_{12} = x_1 w_{12} + x_2 w_{22} + b_1$$

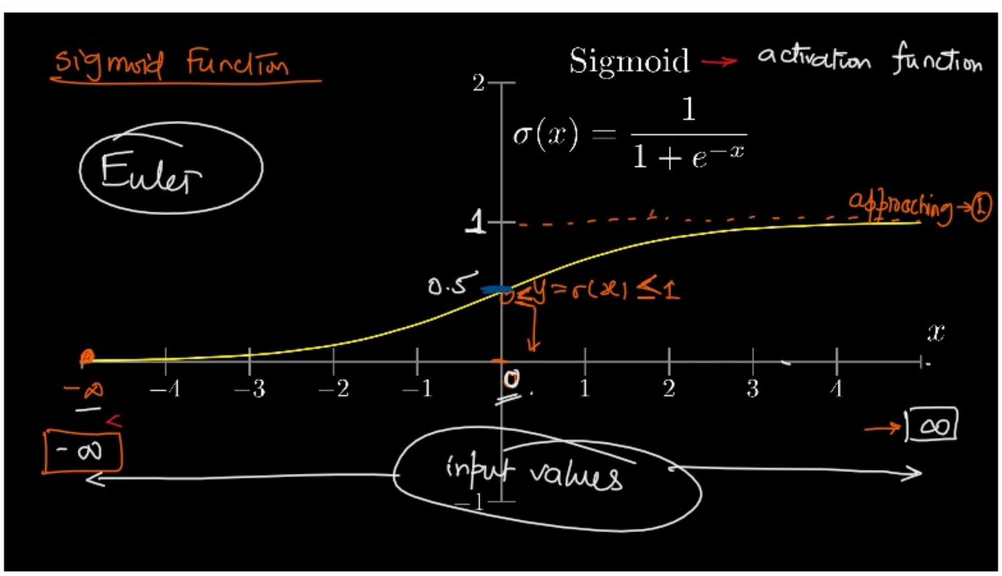
$$Z_{13} = x_1 w_{13} + x_2 w_{23} + b_1$$

where Z_{11}, Z_{12}, Z_{13} are the intermediate neuron value.

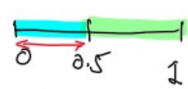
Activation function (sigmoid)

$$a_{11} = \sigma(Z_{11}) \qquad a_{12} = \sigma(Z_{12}) \qquad a_{13} = \sigma(Z_{13})$$

activation functions



logits



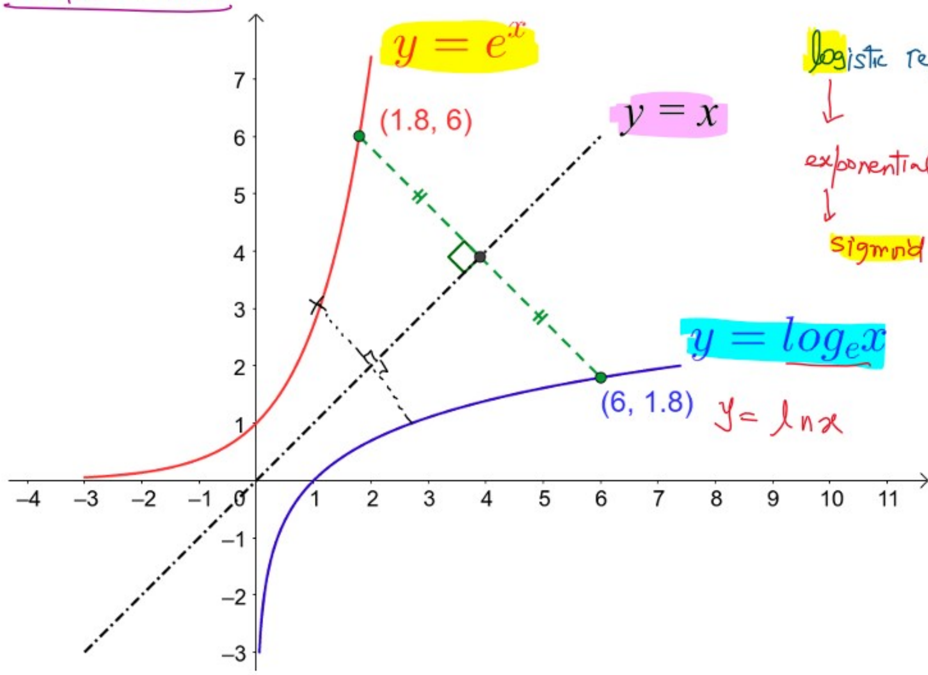
$0 \le \sigma(x) < 0.5 \to$ class#0

$0.5 \le \sigma(x) \le 1 \to$ class#1

$\sigma(x) = \left(\frac{1}{1+e^{-x}}\right)$: sigmoid function.

Exponential Function

Inverse Functions



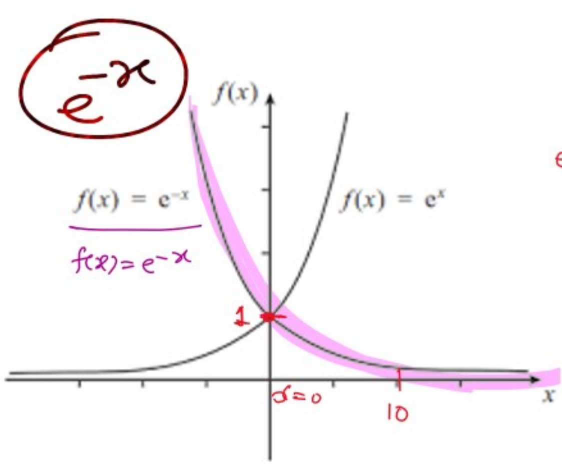
logistic regression

↓

exponential

↓

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



$e^{-x} \rightarrow$ Put $x=0$ $\quad e^0 = \boxed{1}$

$e^{-10} \triangleq (2.71)^{-10}$

$\triangleq$

e^(-10) =

0.00004539992 $\triangleq \underline{0}$

$\triangleq 4.5 \times 10^{-5} \rightarrow 0$

$e^{-\infty} \rightarrow \frac{1}{e^{\infty}} \rightarrow \frac{1}{\cancel{\infty}} \underline{0}$

Sigmoid function

$$\sigma(x) = \left(\frac{1}{1+e^{-x}}\right)$$

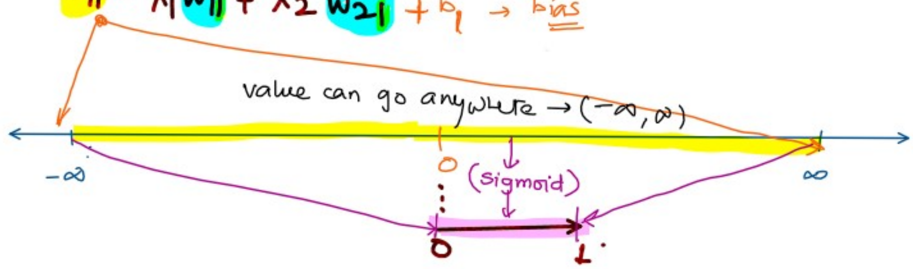
Put $x=0 \rightarrow \sigma(0) = \frac{1}{1+e^{-0}} = \frac{1}{1+e^{0}} = \frac{1}{1+1} = \frac{1}{2} = \underline{\underline{0.5}}$.

Put $x \rightarrow \infty$ $\quad \sigma(\infty) = \frac{1}{1+e^{-\infty}} = \frac{1}{1+\cancel{\left(\frac{1}{e^{\infty}}\right)}^{\,0}} = 1$

Put $x \rightarrow -\infty$ $\quad \sigma(-\infty) = \frac{1}{1+e^{-(-\infty)}} = \frac{1}{1+e^{\infty}} = \frac{1}{1+\infty} = \frac{1}{\infty} \rightarrow \underline{\underline{0}}$

connecting to jth neuron in HL.

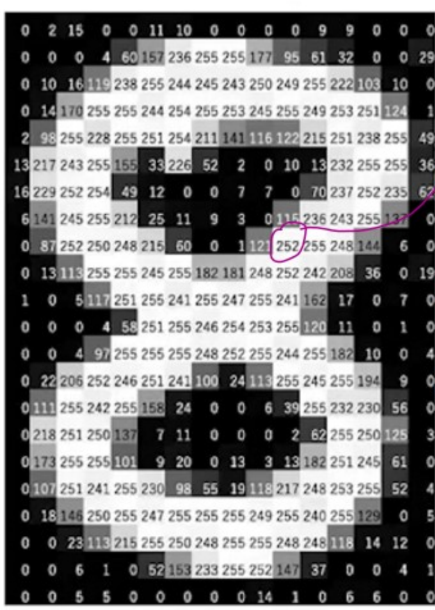
$Z_{11} = X_1 W_{11} + X_2 W_{21} + b_1 \rightarrow$ bias



(grayscale)

8 bit data type $\rightarrow 2^8 = \underline{256}$

0 – 255 → 256 combinations



$X_{100} = 252$
$*$
$w_{100} = -1.8$
$b_{100} = 0.5$

$\xrightarrow{\sigma(x)} \boxed{0 \text{ to } 1}$

$\sigma(x) = \left(\frac{1}{1+e^{-x}}\right)$

$\sigma(0) = 0.5$

$\sigma($

$= 252*(-1.8) + 0.5$

$=$ 252*(-1.8) + 0.5=**-453.1**

$(-453.1) \approx \underline{\underline{-453}}$

$\sigma(-453) = \frac{1}{1+e^{-(-453)}}$

$= \frac{1}{1+\cancel{e^{453}}\,\infty} = \underline{\underline{0}}$

e^453 =
5.437513e+196

$5.4 * 10^{196} \approx \infty$

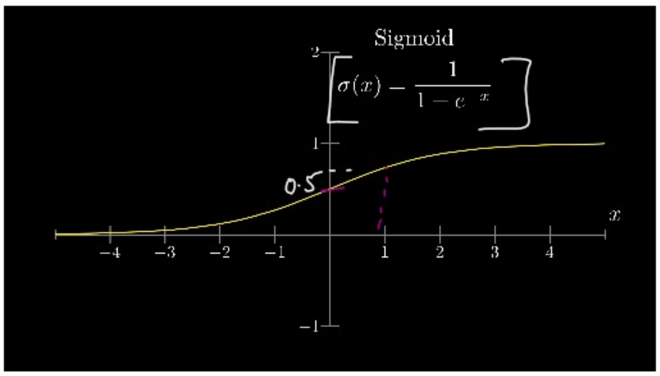
$= \frac{1}{\infty} = 0$..

$5.4 * 10^{196} \doteq \infty$

$= \frac{1}{\infty} = 0$

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

- $x < 0 \rightarrow$ 0 to 0.5
- $x = 0 \rightarrow \underline{0.5}$
- $x > 0 \rightarrow$ 0.5 to 1



$x = 0 \rightarrow 0.5$

$x = 1 \rightarrow \frac{1}{1+e^{-1}} =$ 1/(1+exp(-1))=

1 / (1 + (e^(-1))) =
0.73105857863

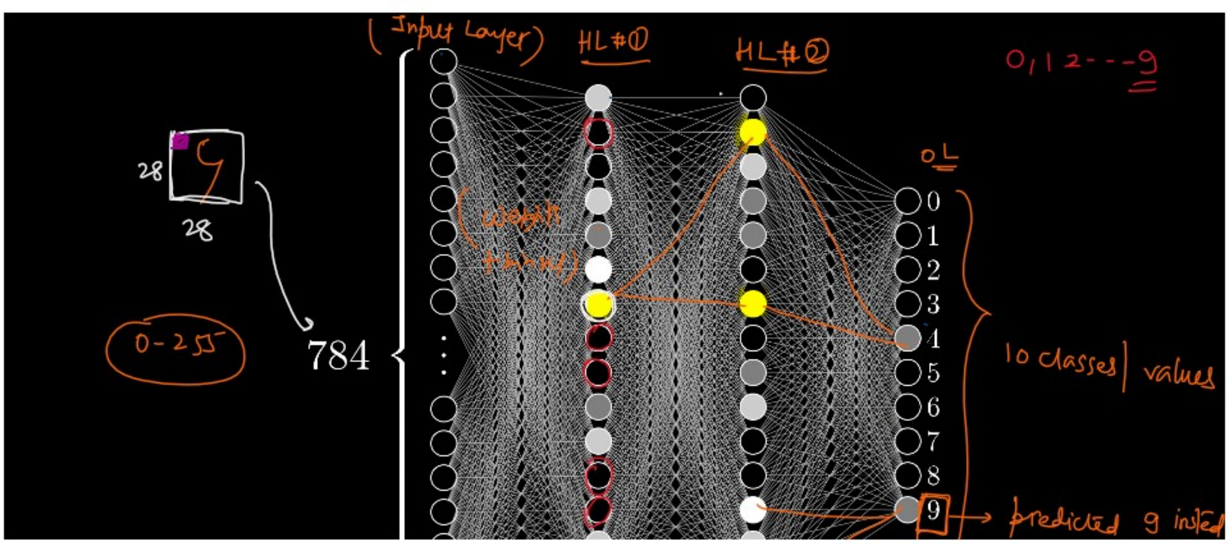
$\simeq \boxed{0.73}$ 35 times

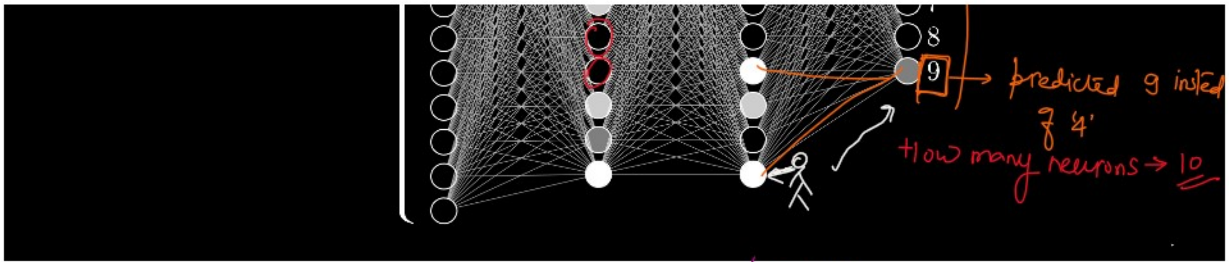
* 10 times

$x = 10 \rightarrow \frac{1}{1+e^{-10}} \triangleq \boxed{0.99}$

1 / (1 + (e^(-10))) =
0.99995460213

0.99/0.73=1.3562





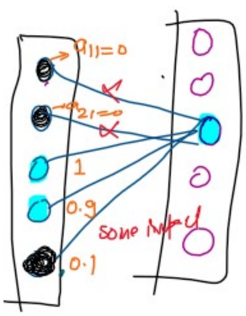
(Fired)

Full-on active

0 ⟷ 1

If the neuron in the first hidden layer is **ON**, a positive weight (ex: $\underset{IL \leftarrow \; \rightarrow HL}{w_{11}} = 2.07$) suggests that the neuron in the second hidden layer should also be **ON**

Similarly a negative weight (ex. $w_{71} = -1.84$) suggests that the neuron in the second layer should be off.



$a_{11} * w_{12} + a_{21} * w_{12} + \cdots$