



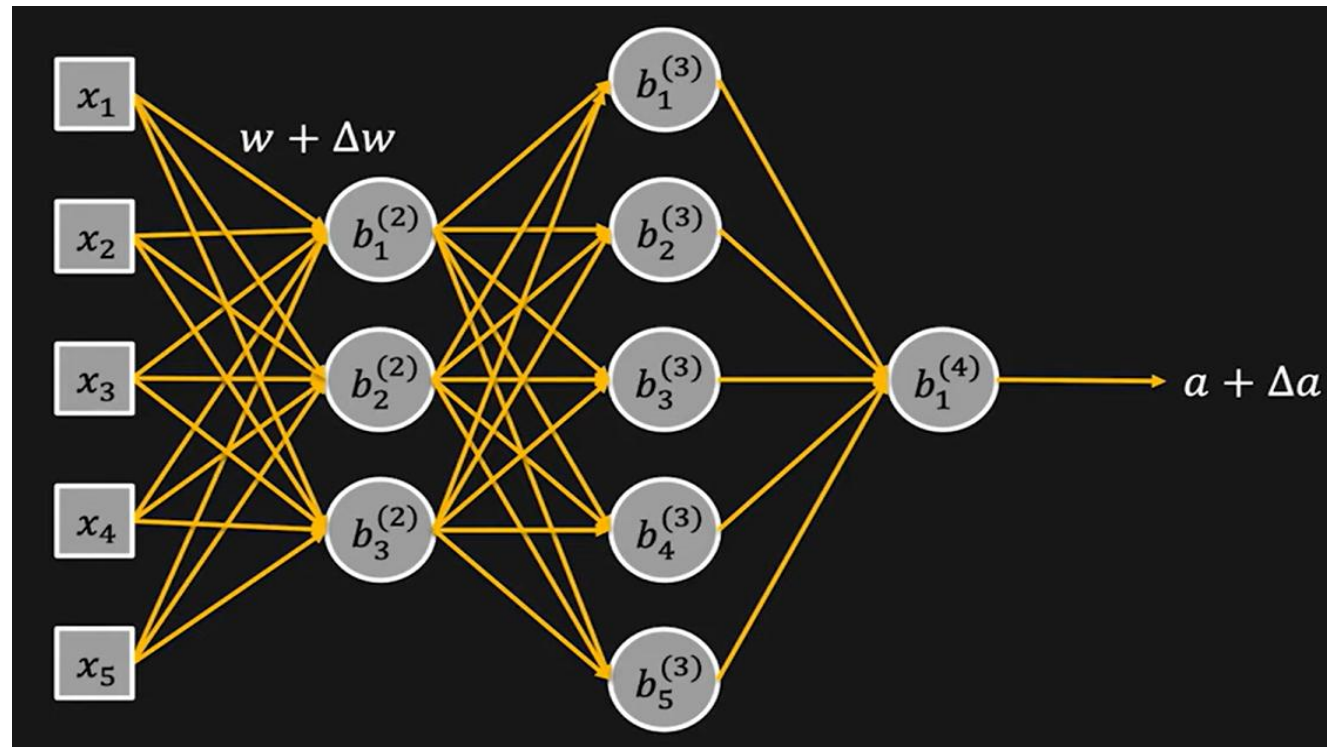
# Neural Network

---

RECOGNITION

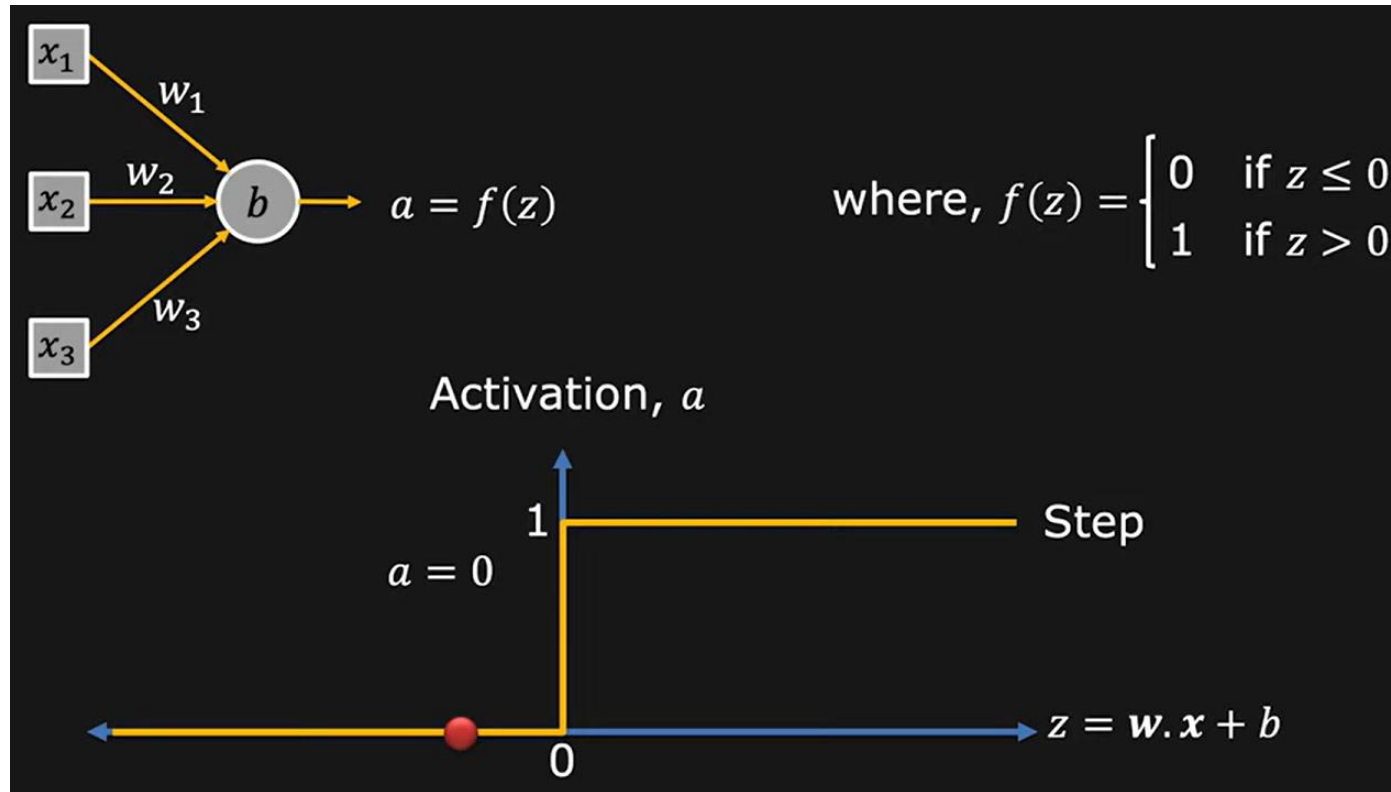
# Varying the weights & biases

---



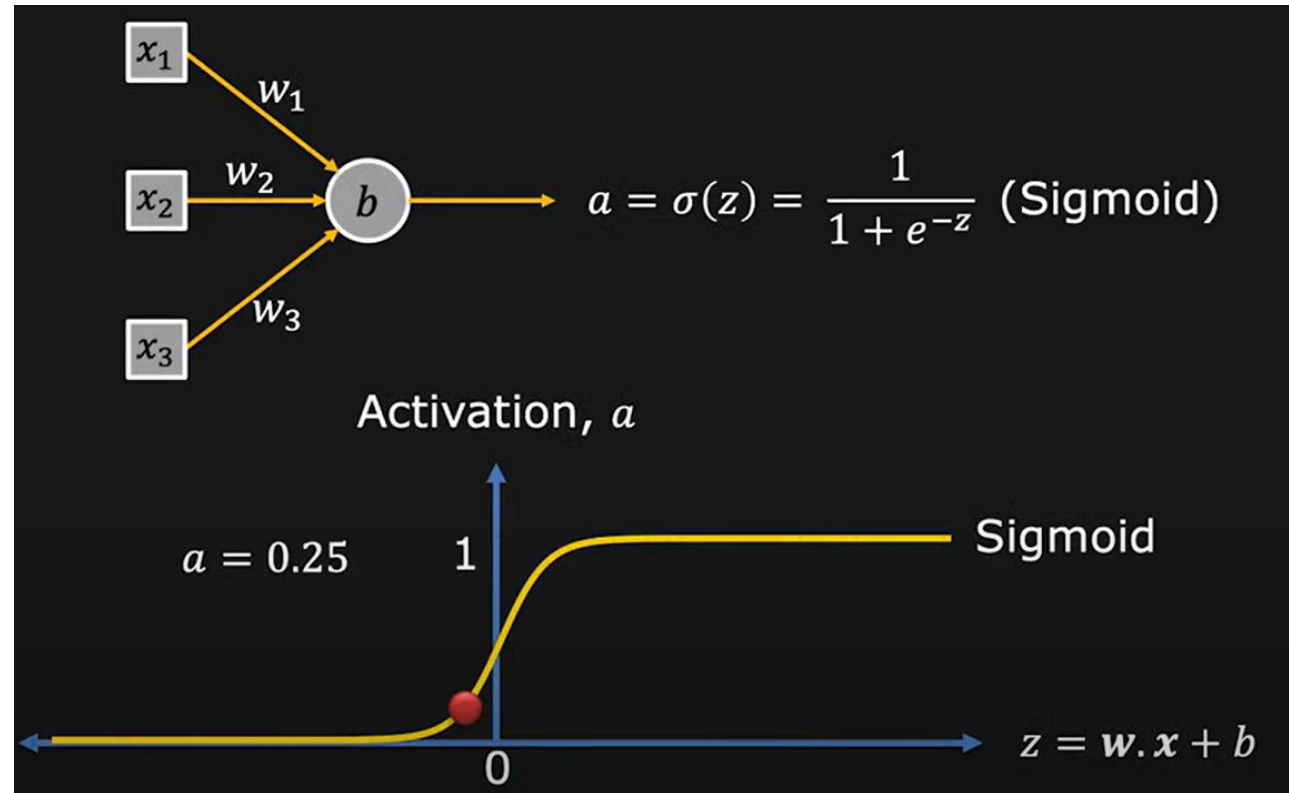
# Varying the weights & biases

---



# Sigmoid Neuron

---



# Recognizing Characters

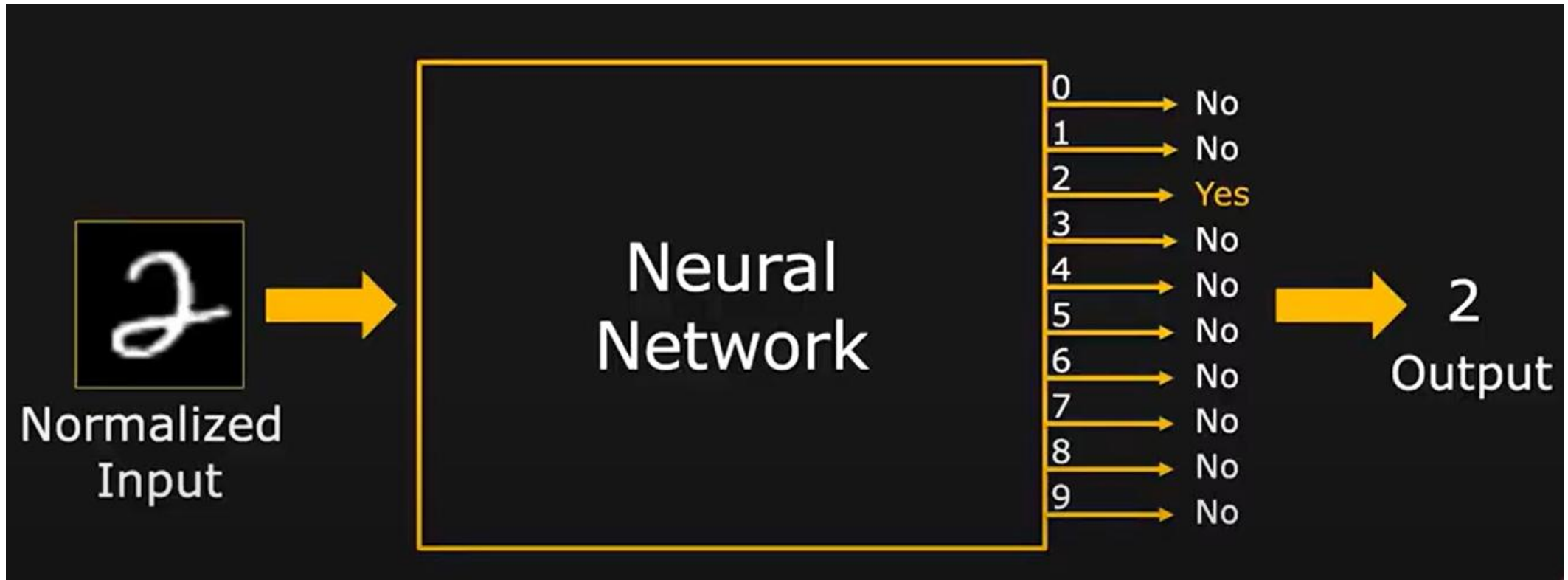
---





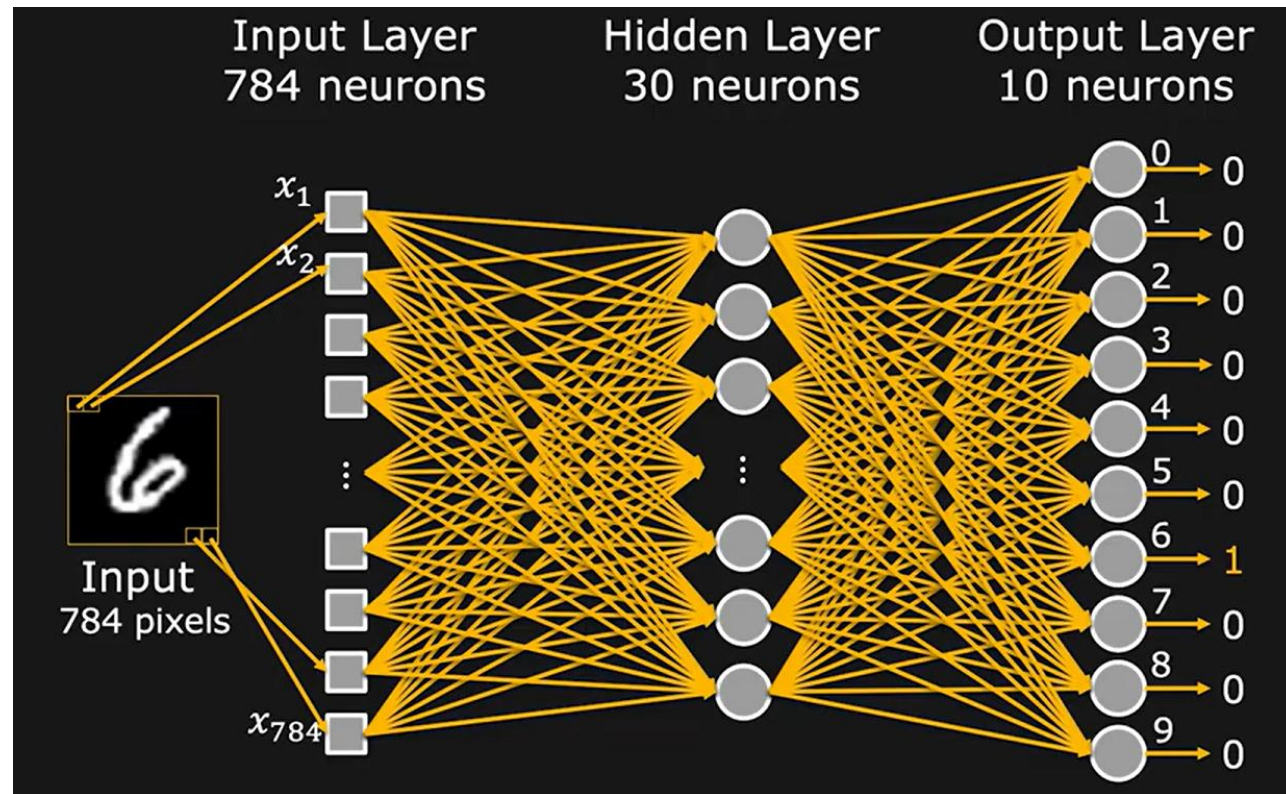
# Recognizing Characters

---

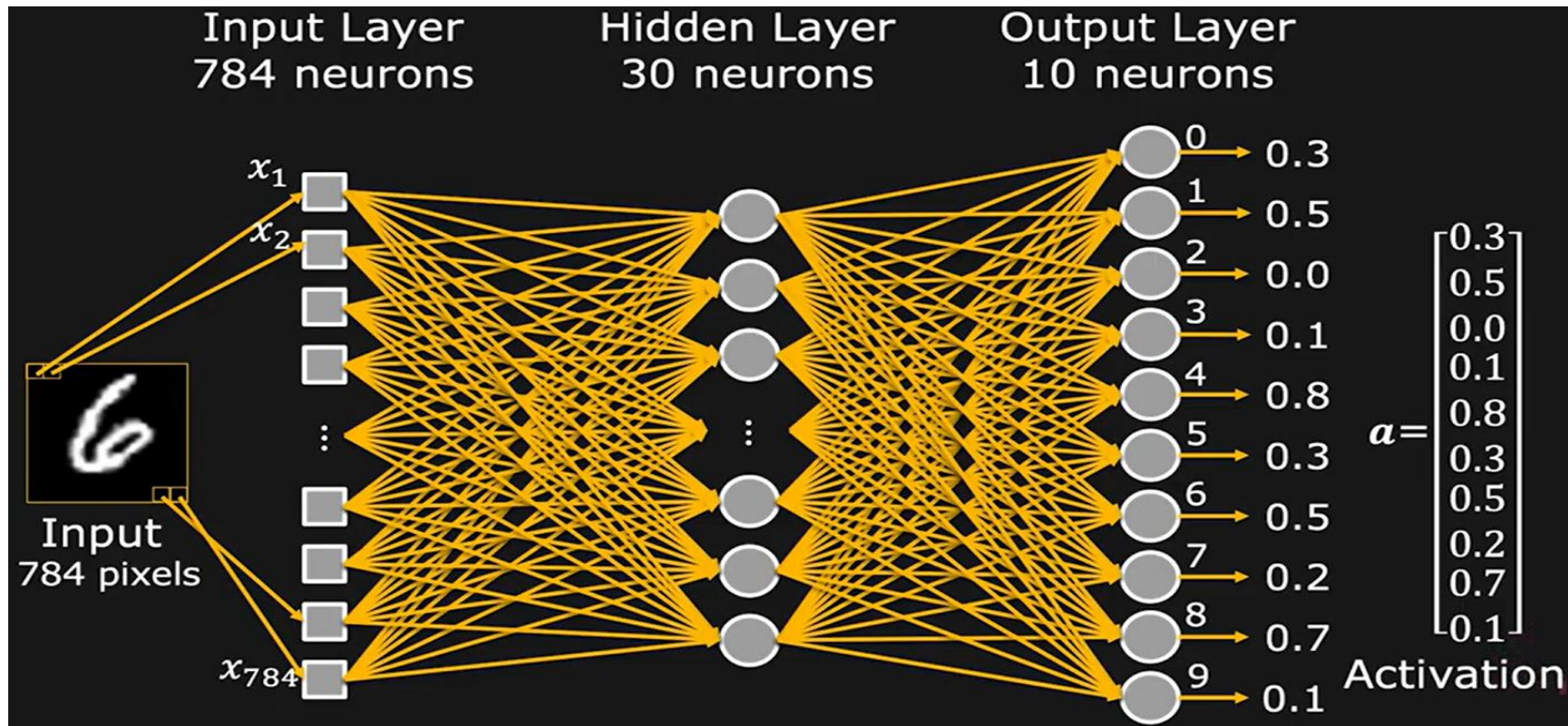


# Character Recognition Network

---













# Network with Random Weights & Biases





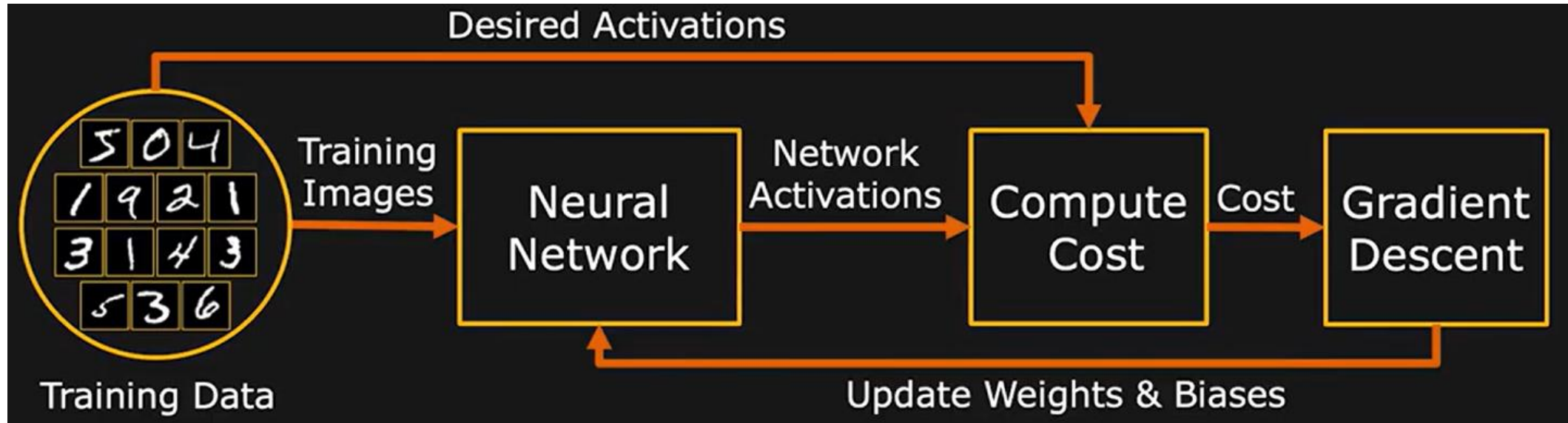
# Training data

Sample Training Data - MNIST Dataset (60,000 images)

Training Image $x$										
Label	5	0	4	1	9	2	1	3	1	4
Desired Activation $\hat{a}(x)$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

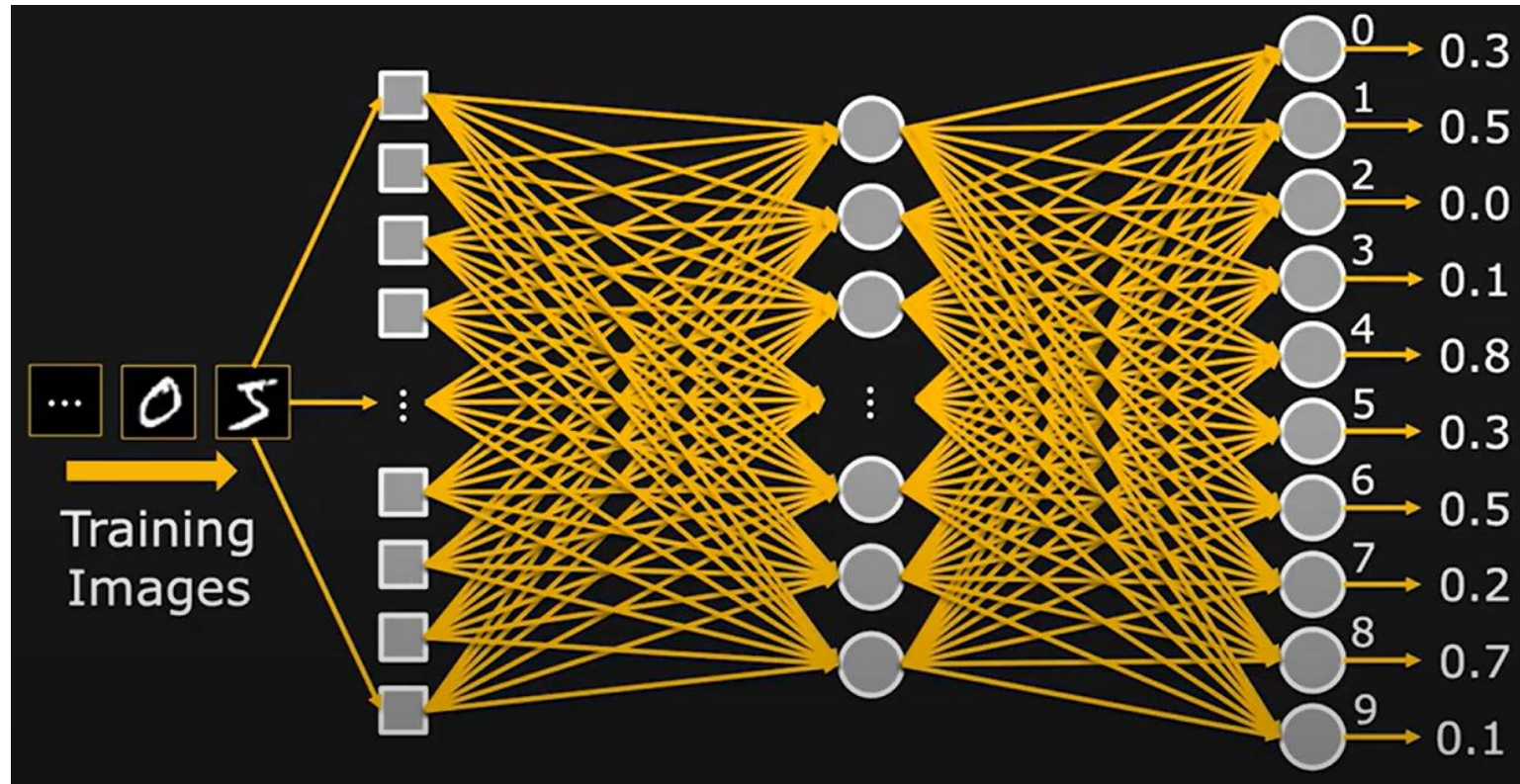
# Training Process

---



# Compute Activation

---



# Computing Cost

---

For a **Single** Training Image:

$$C_x(\mathbf{w}, \mathbf{b}) = \|\hat{\mathbf{a}}(x) - \mathbf{a}(x|\mathbf{w}, \mathbf{b})\|^2$$

Desired Activation

Network Activation

For  $\Sigma$

$$C_x = \left\| \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0.3 \\ 0.5 \\ 0.0 \\ 0.1 \\ 0.8 \\ 0.3 \\ 0.5 \\ 0.2 \\ 0.7 \\ 0.1 \end{bmatrix} \right\|^2 = 2.27$$

For **Entire** Training Data:

$$C(\mathbf{w}, \mathbf{b}) = \frac{1}{n} \sum_x C_x(\mathbf{w}, \mathbf{b})$$

Number of Training Images



# Training Process

---

1. Initialize Network with random weights and biases values.
2. Compute Network activation for each training image.
3. Compute cost for entire training data.
4. Update Wights and Biases using Gradient Descent.