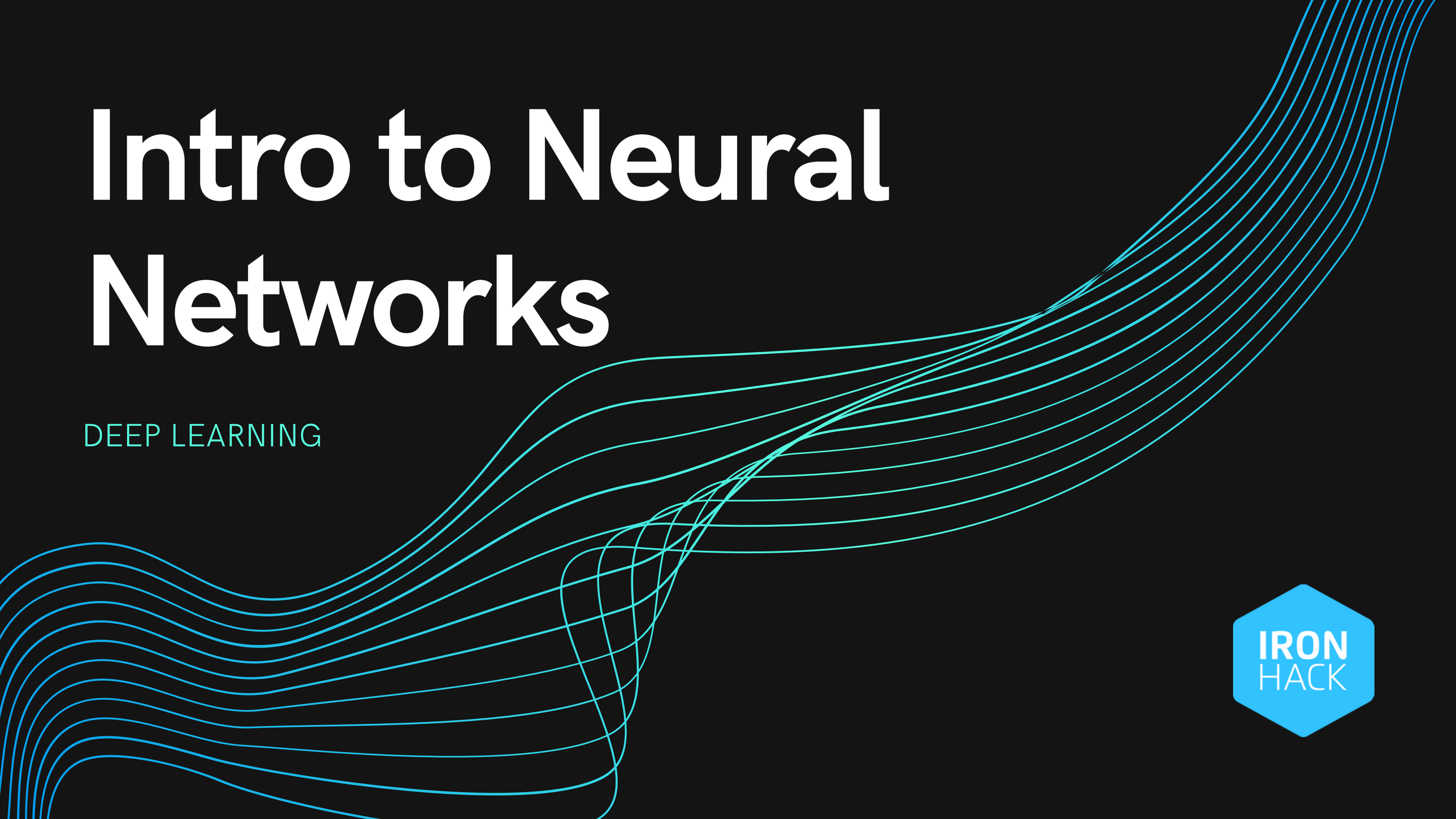
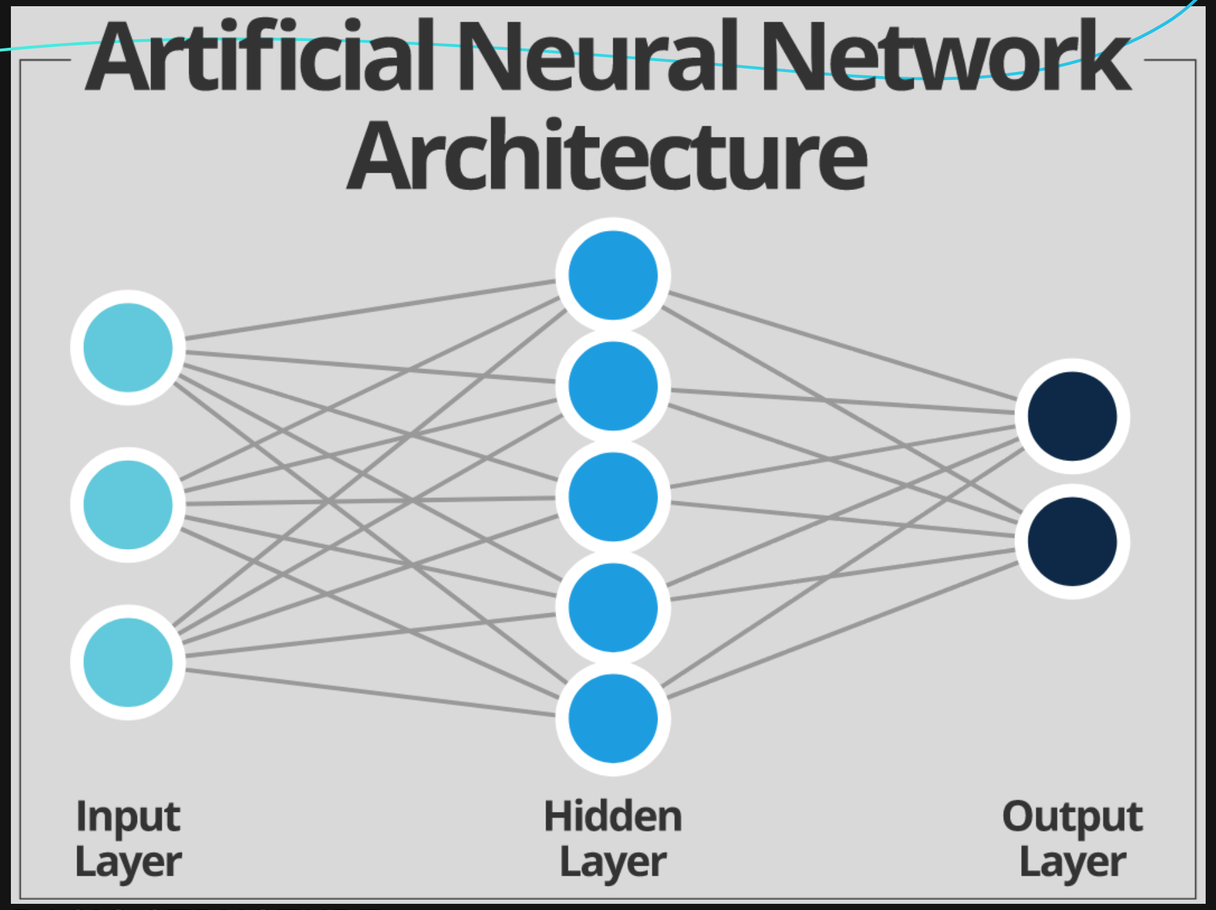
Intro to Neural Networks

DEEP LEARNING





BASIC STRUCTURE

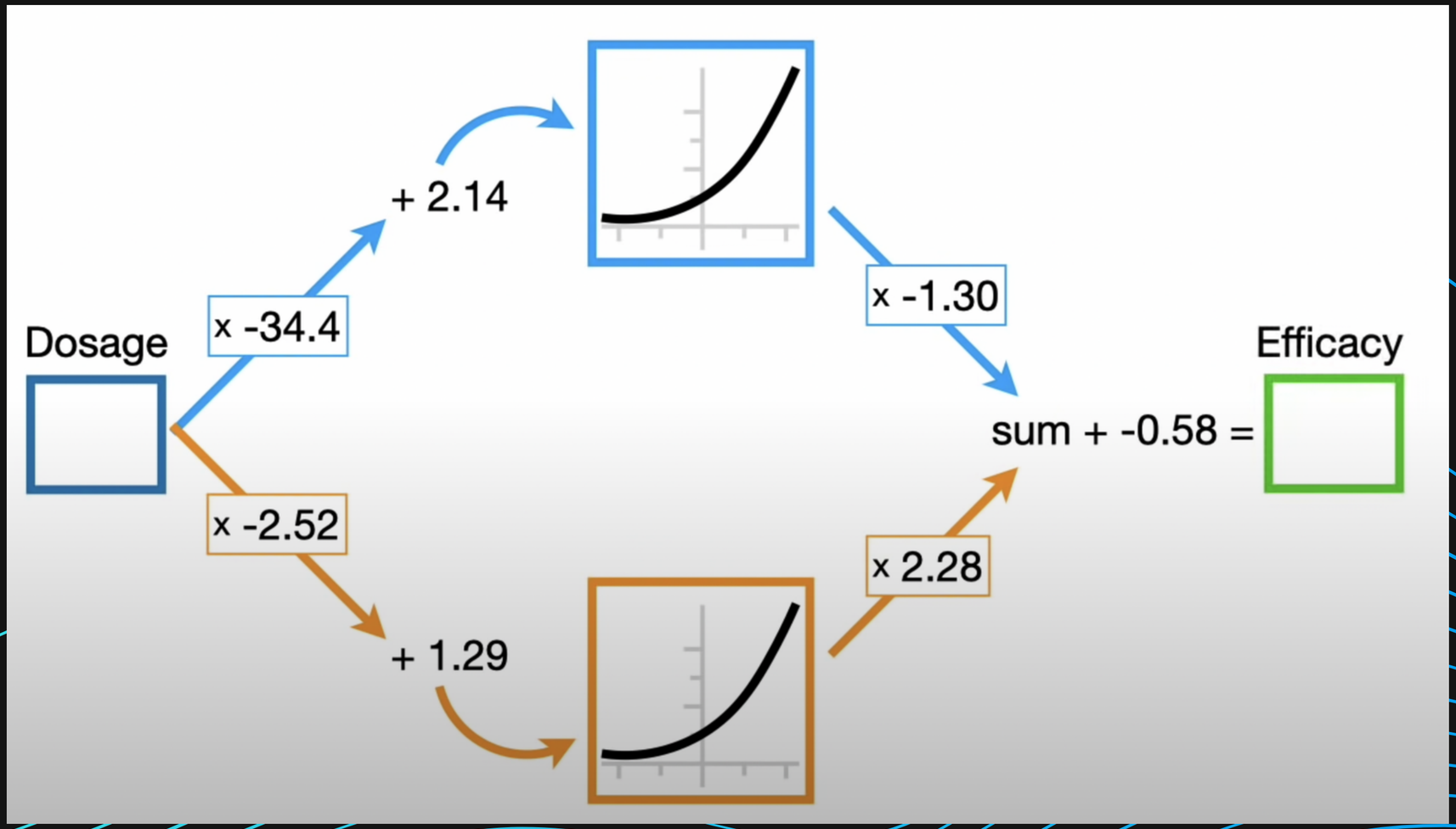
- Input, Output and Hidden layers
- Connections (Synapsis)
  - Weights
  - Bias
- Nodes (Neurons)
  - Activation Function

Weight, bias, activation function

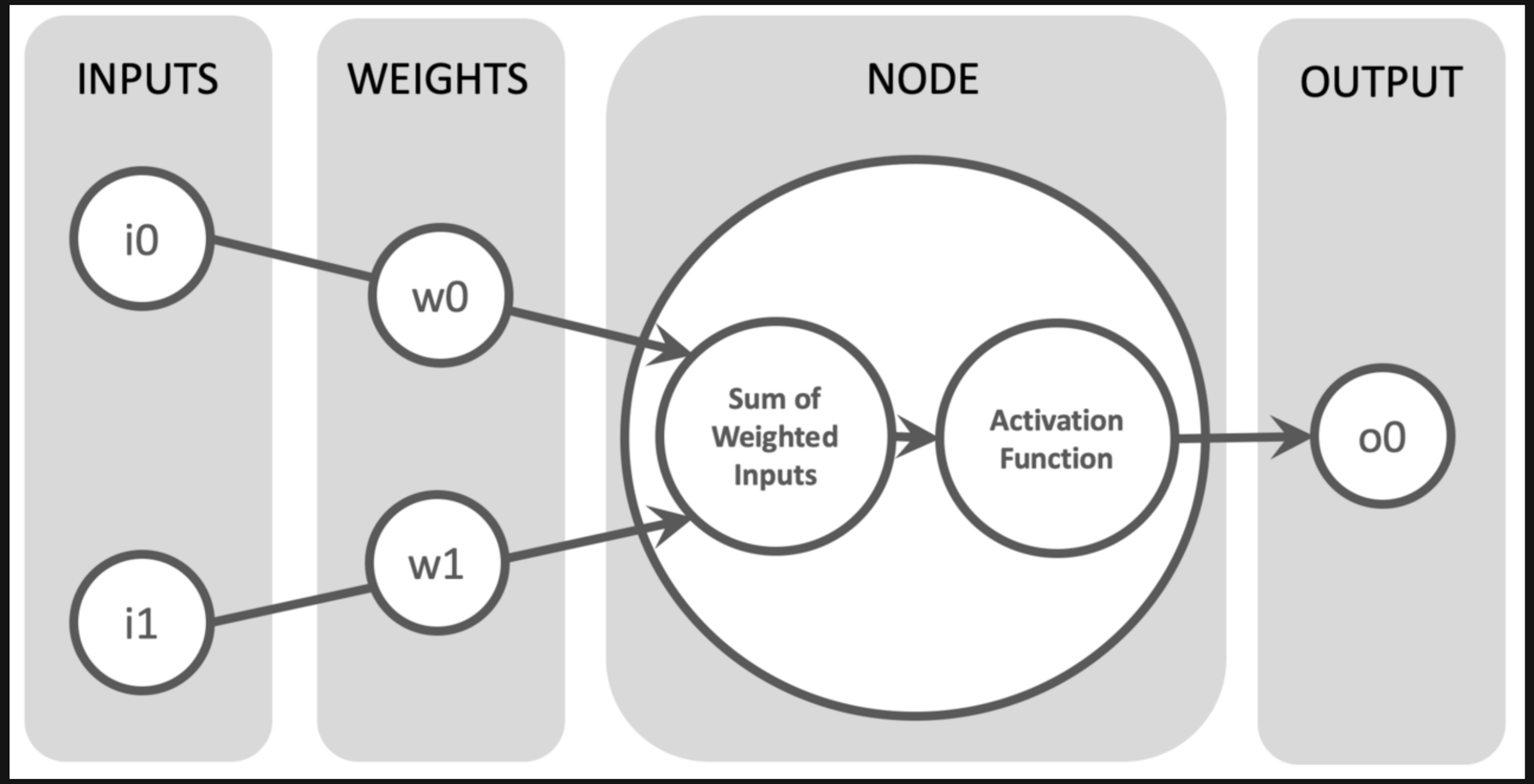
HOW MATH FITS THE STRUCTURE

Example from StatQuest awesome 3 part series on NN.

- pt. I
- pt. II
- pt. III



Each Node


INPUTS
WEIGHTS
NODE
OUTPUT
i0
i1
w0
w1
Sum of Weighted Inputs
Activation Function
o0

HOW DOES IT FIT THE DATA?

Backpropagation

ADJUSTING WEIGHTS AND BIASES

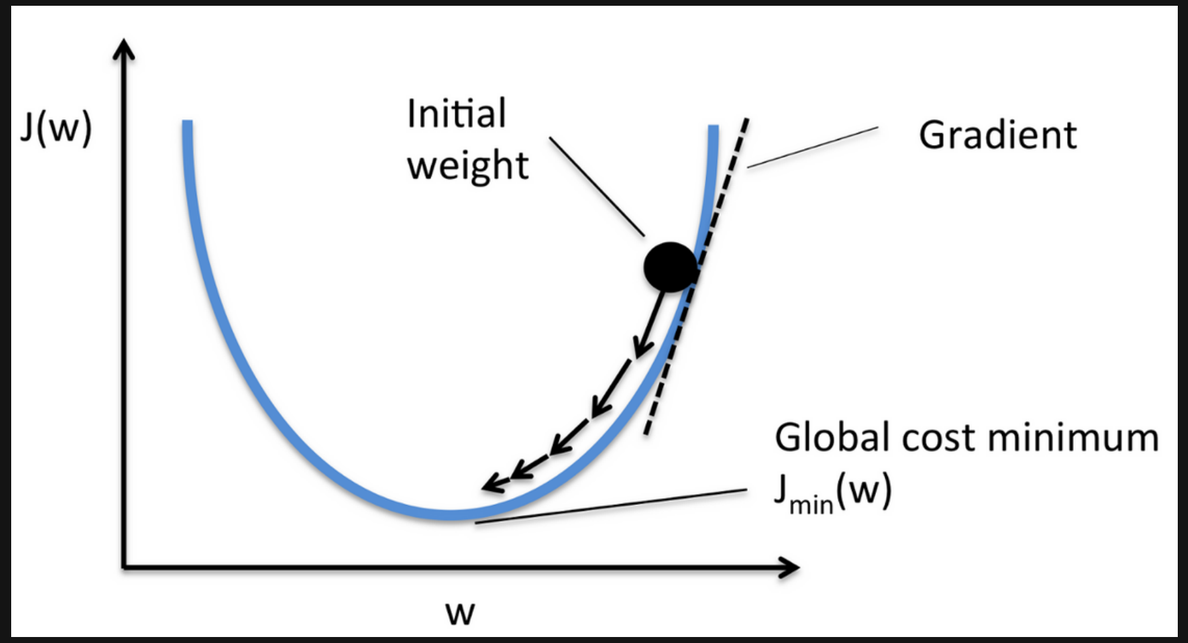
From the end to the beginning for each example

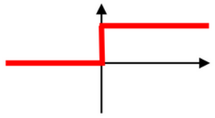
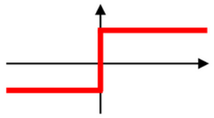
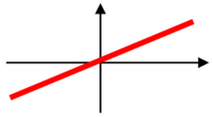
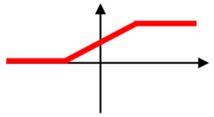
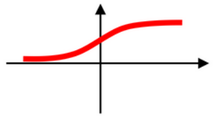
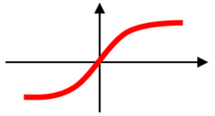
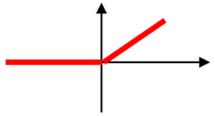
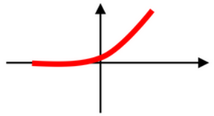
GRADIENT DESCENT

Finding the point of minimal error on cost function.

STOCHASTIC GRADIENT DESCENT

Using mini-batches of data



Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^z)$	Multi-layer Neural Networks	

Activation functions

Without activation functions, a Neural Network Model is basically a Linear Regression Model. They allow for the different shapes (non-linear) in fitting the model to the data.

LEARN YOUR ACTIVATION MOVES

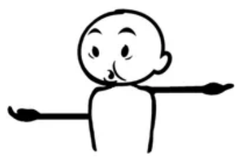
Sigmoid

$$y = \frac{1}{1+e^{-x}}$$

Tanh

$$y = \tanh(x)$$

Step Function



$$y = \begin{cases} 0, & x<n \\ 1, & x \geq n \end{cases}$$

Softplus

$$y = \ln(1+e^{x})$$

ReLU

$$y = \begin{cases} 0, & x<0 \\ x, & x \geq 0 \end{cases}$$

Softsign

$$y = \frac{x}{(1+|x|)}$$

ELU

$$y = \begin{cases} \alpha(e^{x}-1), & x<0 \\ x, & x \geq 0 \end{cases}$$

Log of Sigmoid

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

Swish

$$y = \frac{x}{1+e^{-x}}$$

Sinc

$$y = \frac{\sin(x)}{x}$$

Leaky ReLU

$$y = \max(0.1x, x)$$

Mish

$$y = x(\tanh(\text{softplus}(x)))$$