ML

NIRMA
UNIVERSITY

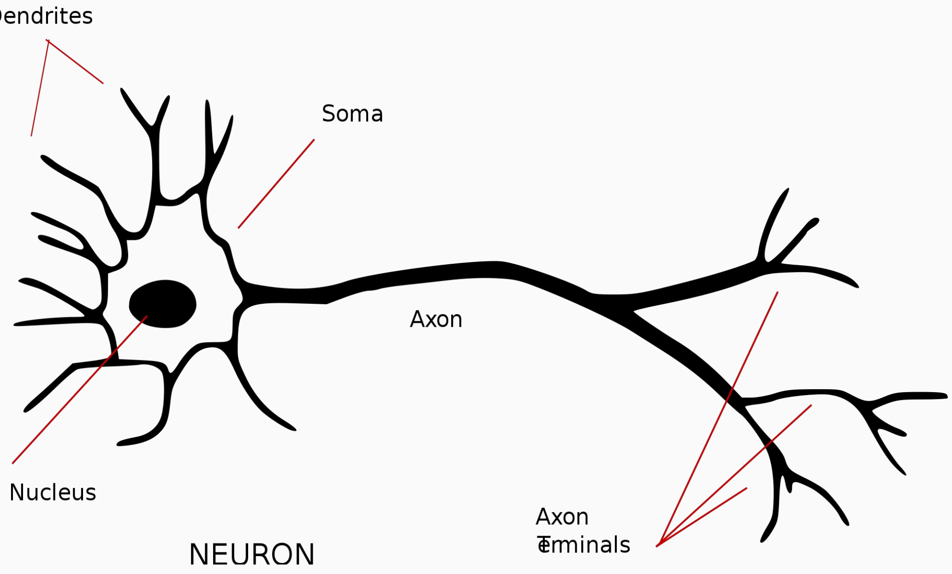
# Artificial Neural Networks

## Machine Learning Lecture Series: 3

# Neurons

# Biological Neuron

In biology, neurons are the building blocks of the very robust neural network that connects every corner of your body with the brain and allows brain to operate one’s body.



# Biological Neuron

Dendrites act as input for the neuron and axon terminals as the output of the neuron.

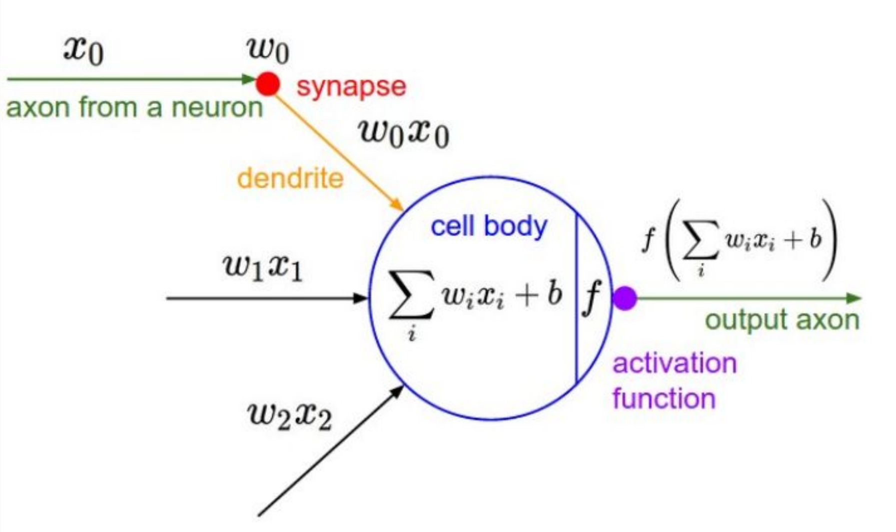
Axon Terminals connect with next neuron's dendrites.

Nucleus/ Body performs necessary computations and converts data into electrical signals which propagate through the axon.

# Artificial Neuron

Similarly, computer scientists thought of a mathematical function that could resemble biological neuron and function in similar fashion.

So, the artificial neuron has a similar structure. It contains a nucleus (processing unit), several dendrites (inputs), and one axon (output).

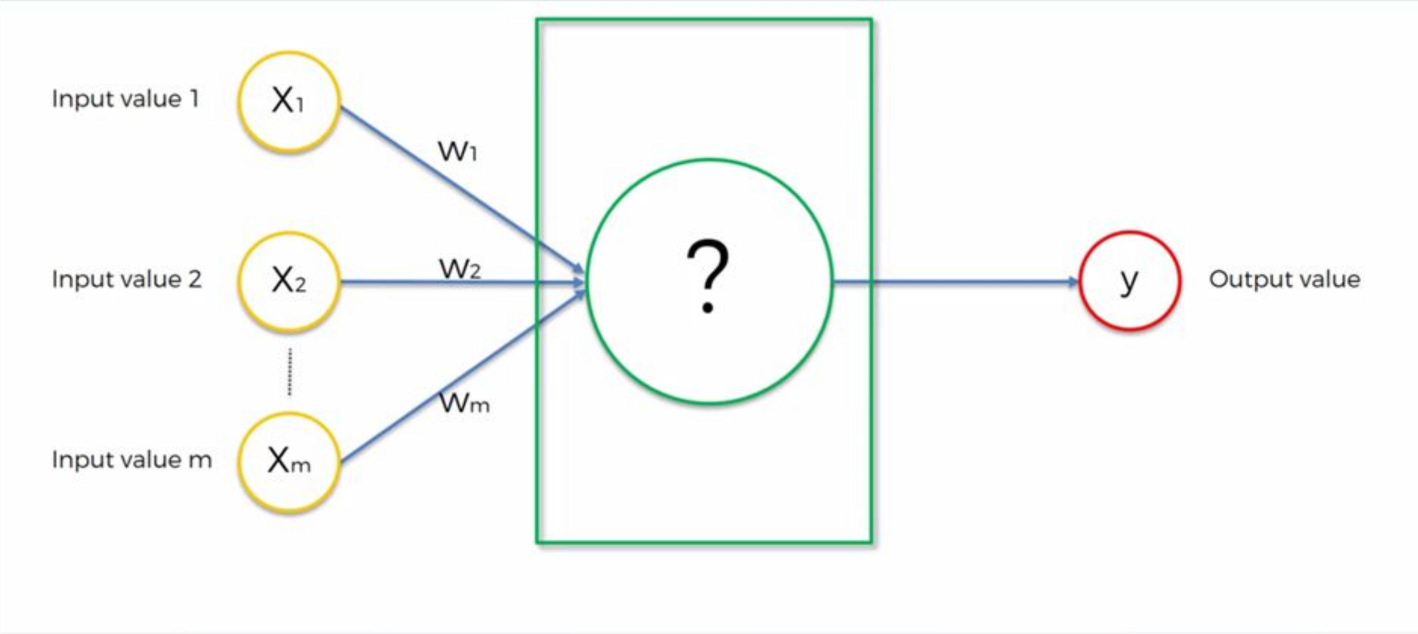


# More about Neuron

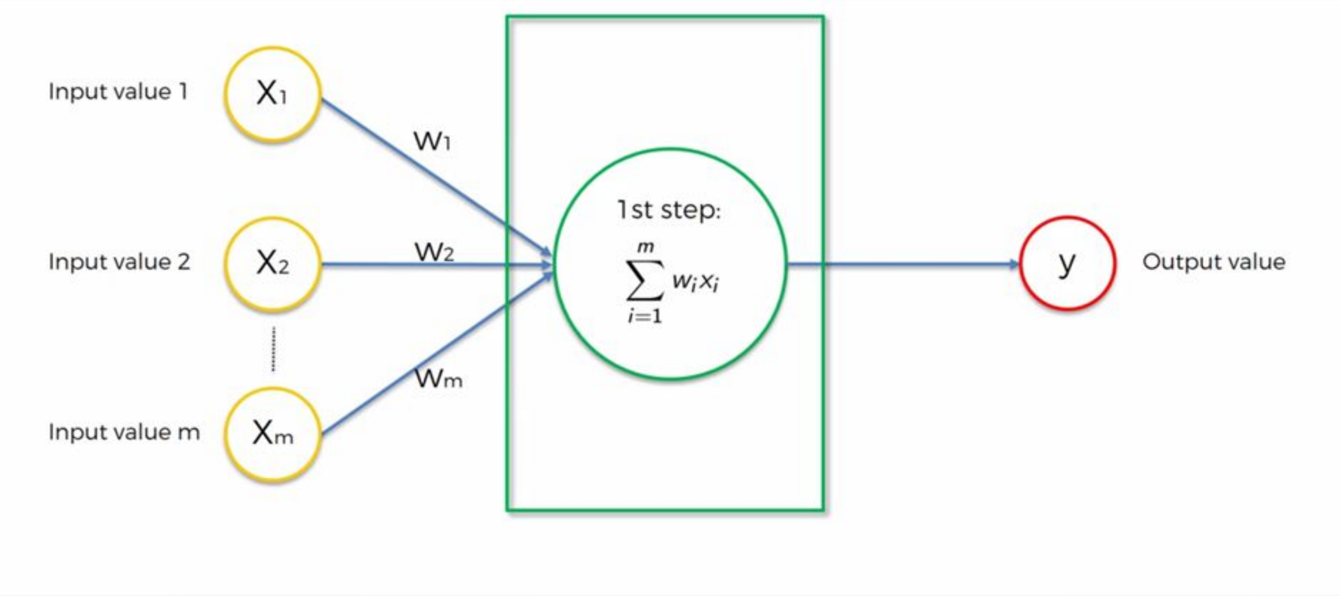
In practice, a neuron is a mathematical function which takes input(s), processes the input(s) by implementing the mathematical function and gives output(s).

When an input comes in, it gets multiplied by a weight value that is assigned to that particular input. That is, if a neuron has three inputs, then it has three weights that can be adjusted individually. The beauty of a neural net is that it can learn these weights on its own during the training phase by backpropagation algorithm.

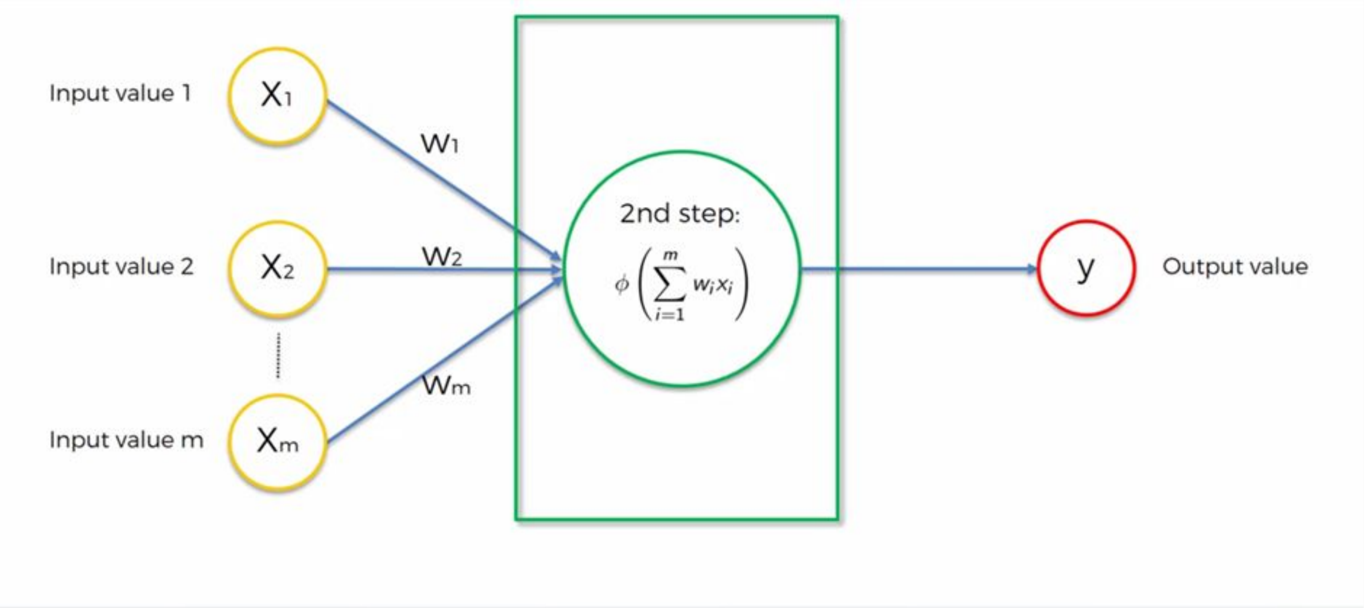
# Let’s say hello to our neuron



# Baby step 1(Weighted Sum)



# Baby step 2(Activation Function)



# Activation Functions

# What is an activation function?

Also known as transfer function, it transfers the knowledge gathered by the neuron into some meaningful quantity(mostly a single quantity) which is further propagated to other neurons in the network.

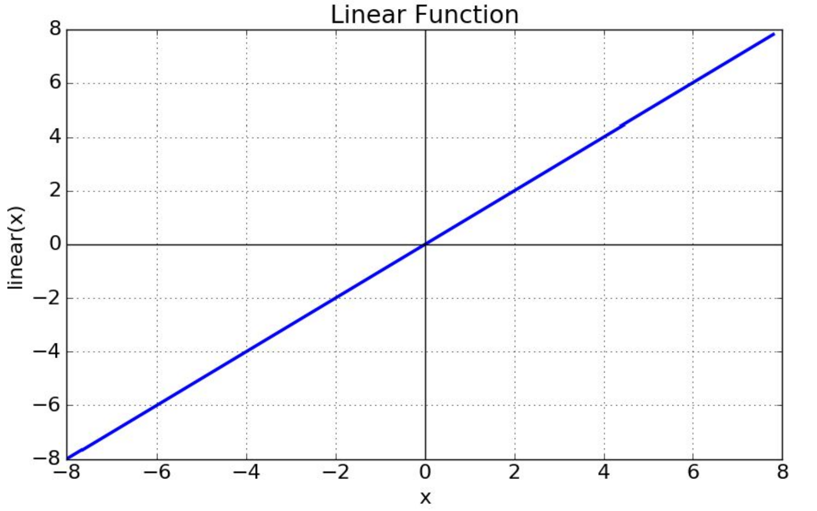
In practice, they are mathematical functions which take the net sum as input and outputs a value by applying a simple function over it.

There are many such functions which are used in practice viz. Sigmoid, ReLU, Linear, etc.

# Linear

Mathematically,

$$f(x) = x$$

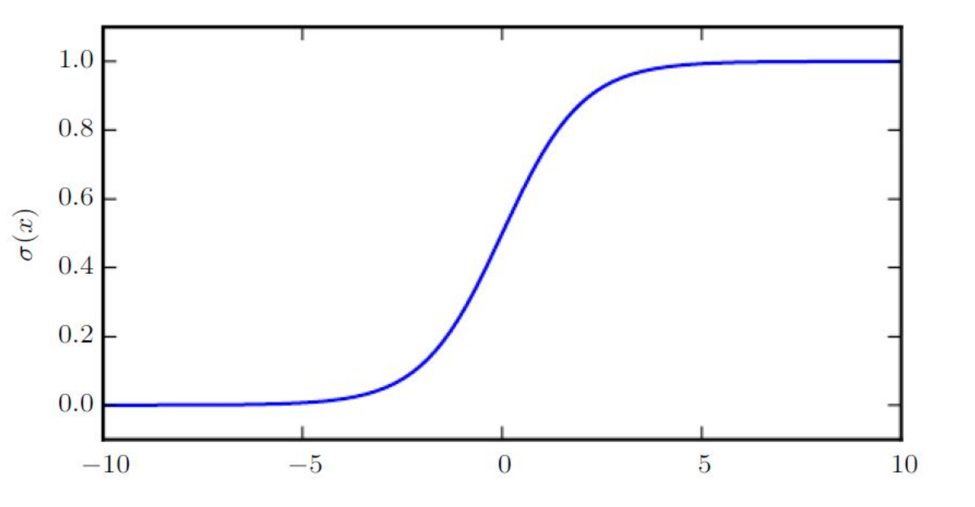


# Sigmoid

Mathematically,

$$\frac{1}{1+\exp(-x)}$$

Visually,

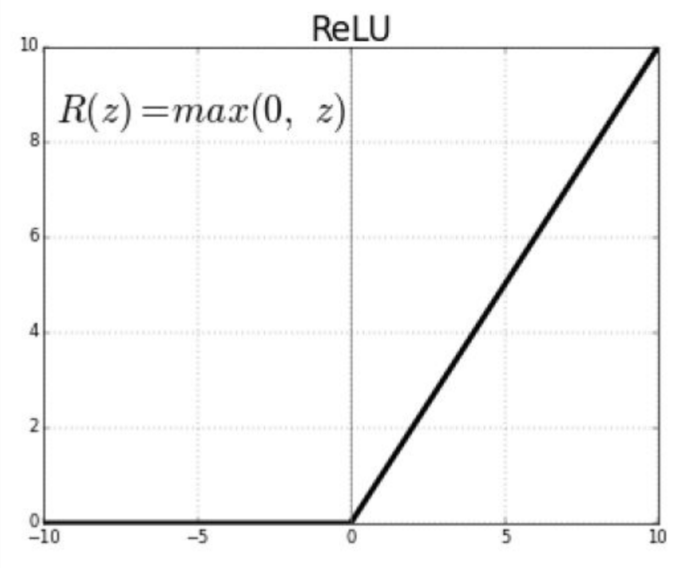


# Rectified Linear Unit(ReLU)

Mathematically,

$$\max(0, x)$$

Visually,

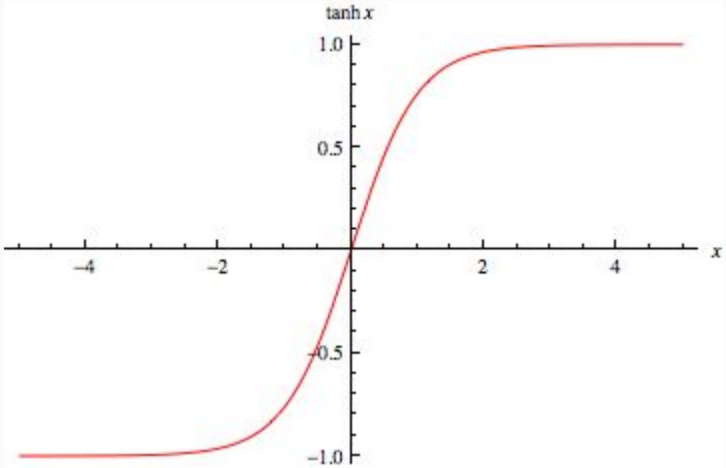


# Hyperbolic Tangent(tanh)

Mathematically,

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

Visually,



# Neural Networks

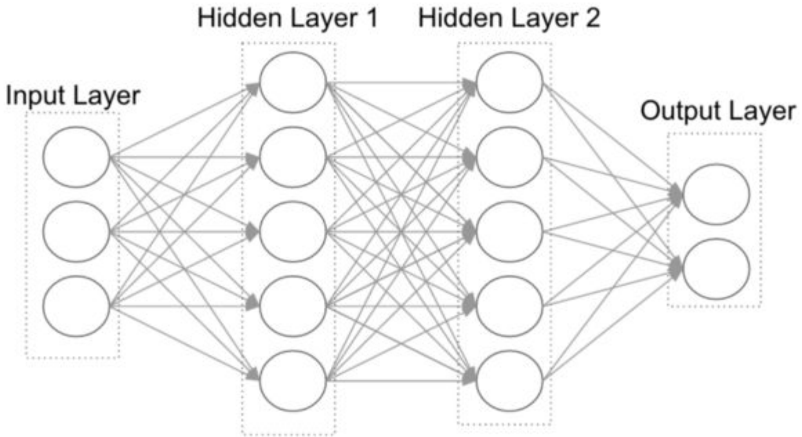
# Feed Forward Neural Net

Input -> Hidden* -> Output

Shallow Network: <= 2 Hidden Layers.

Deep Network: > 2 Hidden Layers.

# Feed Forward Neural Net




**Feedforward neural network with 2 hidden layers**

# Feed Forward Neural Net

The leftmost layer in this network is called the input layer, and the neurons within the layer are called input neurons.

The rightmost or output layer contains the output neurons, or, as in this case, a single output neuron.

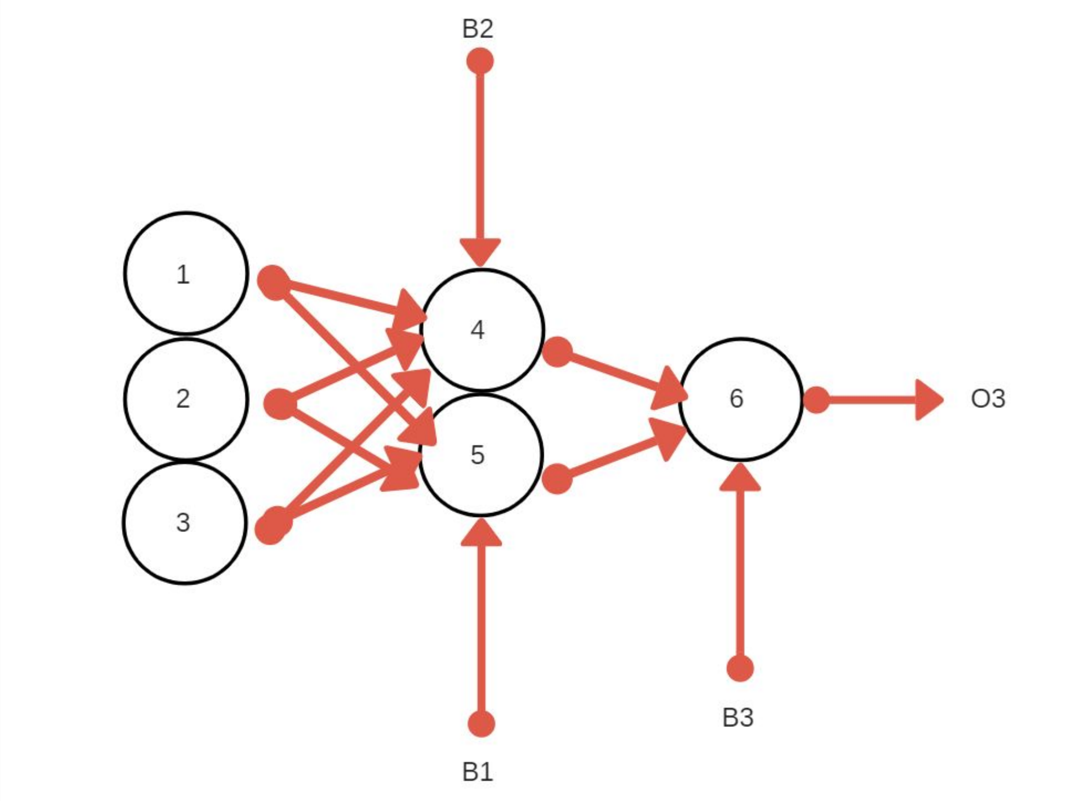
The middle layers are called a hidden layers, since the neurons in this layer are neither inputs or outputs.

# Feed Forward Neural Net

As we see here that the output of the input layer neurons go to every neuron in the next layer ( first hidden layer ) as inputs, such a connection is called **Densely Connected** or **Fully connected layer**.

We let the network decide which inputs are necessary and which are not. As and when it will process the dataset, the network will decide upon which inputs make an impact on deciding output per se which features are useful and which are redundant. It will eventually assign a negligible or 0 weight to the redundant input features.(That’s the beauty of Machine Learning)

# Forward Pass




I1 = 5
I2 = 3
I3 = 1

W14 = -1
W15 = 2
W24 = 3
W25 = 0
W34 = 1
W35 = 6
B1 = 1
B2 = -2
W46 = 3
W56 = 4
B3 = -6

T = 1

Activation Function is Sigmoid everywhere.

Simple task of binary classification with 1 hidden layer and 3 independent input variables.

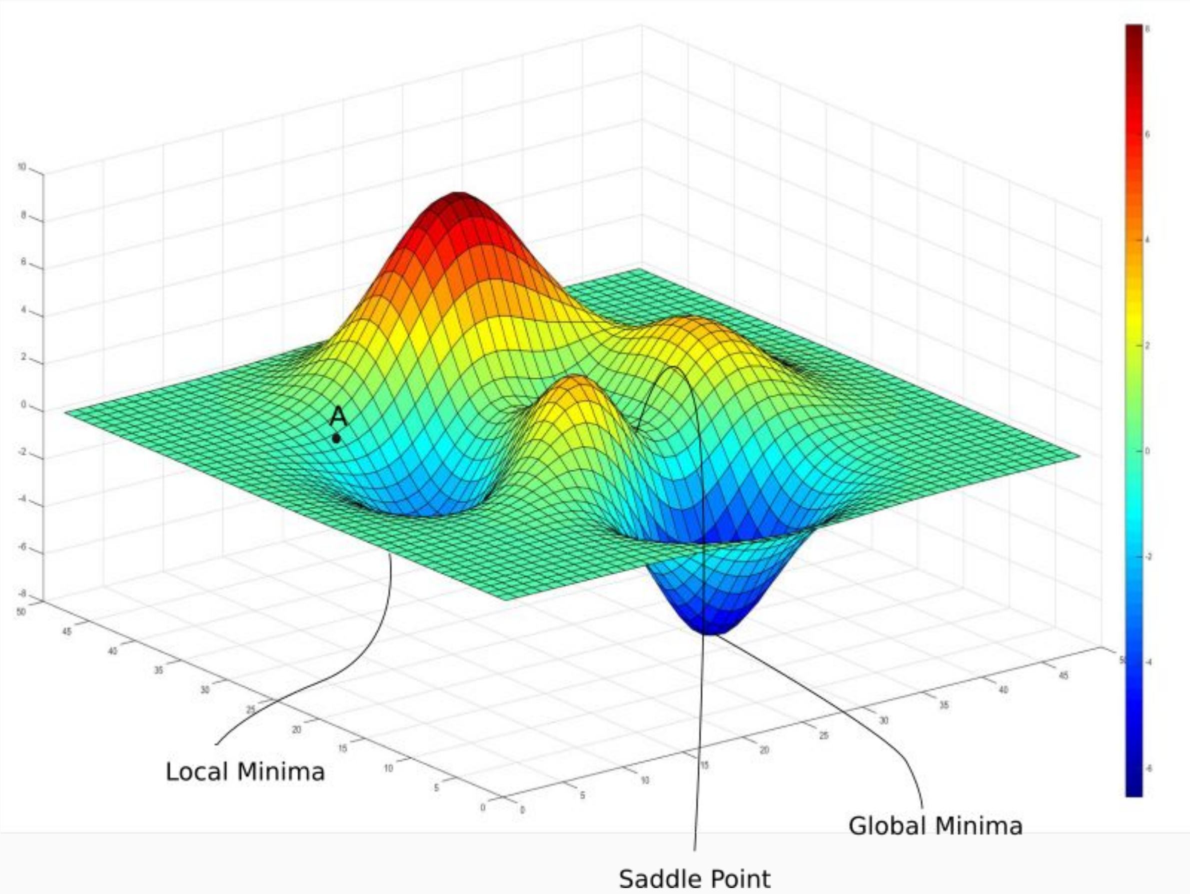
Solve a forward pass.

Error = $\frac{1}{2}(T - O_3)^2$

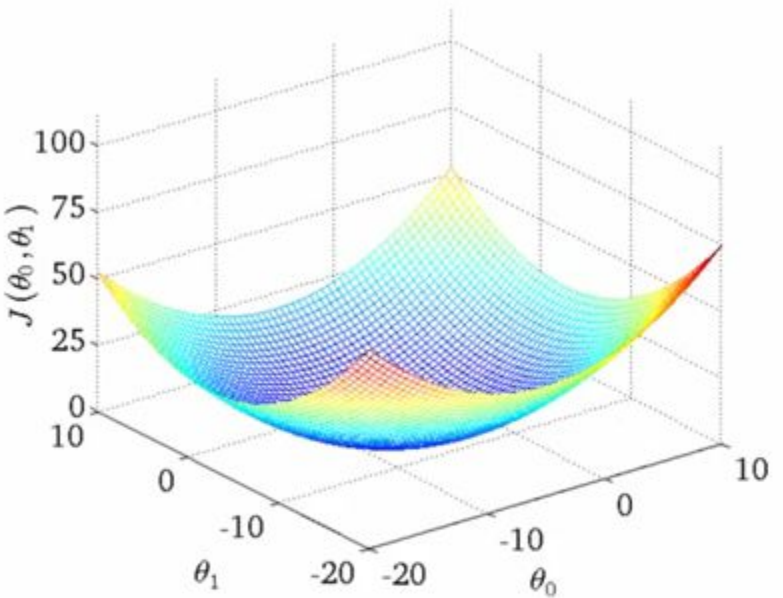
# Why Such Error Function? I.e Why Are We Using Squared Error?

# CONVEX OPTIMIZATION

# Non - Convex Optimization: Multiple Minima

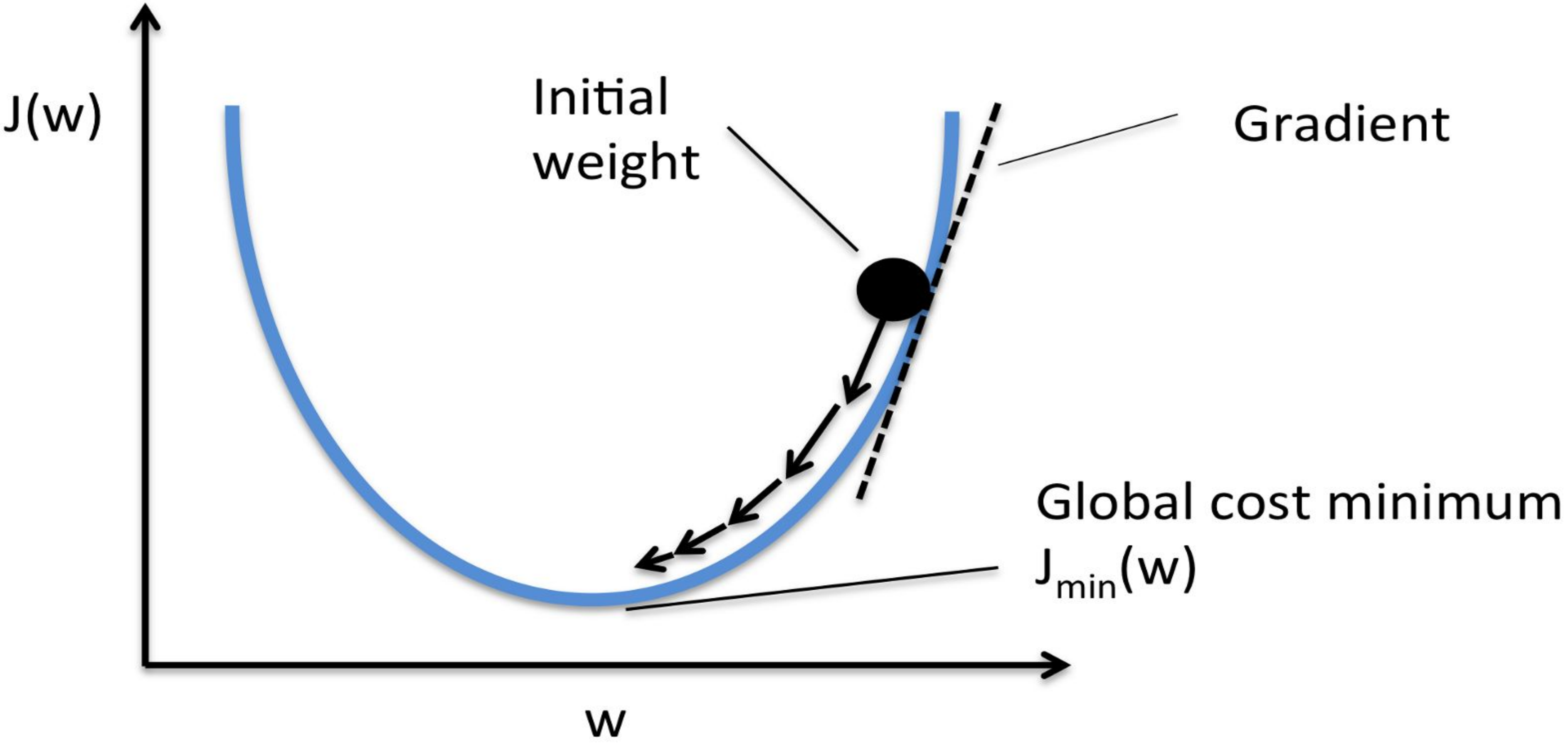


# Convex Optimization: Single Minima



A Neural Network learns through Back Propagation by using Gradient Descent Approach.

# Gradient Descent



That’ll be all for ANN. Now we’ll move onto CNN and see why ANN would not be our best choice for images

Prepared By: Param Popat
https://github.com/parampopat/machinelearning_lecture