Artificial Neural Networks

Mass Street University

PER EDUCATIONEM PROGRESSUS

What we will cover

- History of them
- How they are different
- How Neural Networks work

Mass Street University

PER EDUCATIONEM PROGRESSUS

Foundations

- It came from neuroscience in the early 40’s.
- Neural Networks use tendrils and neurons to process information.
- First form was Hebbian learning.

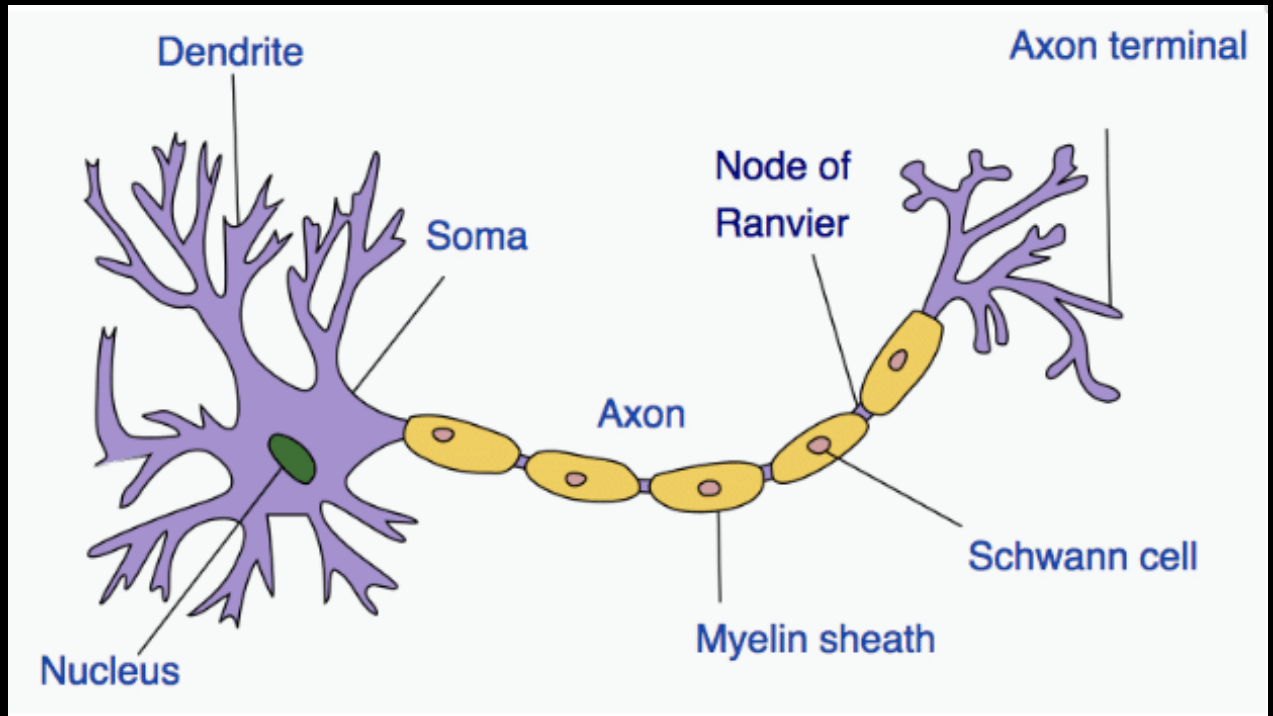
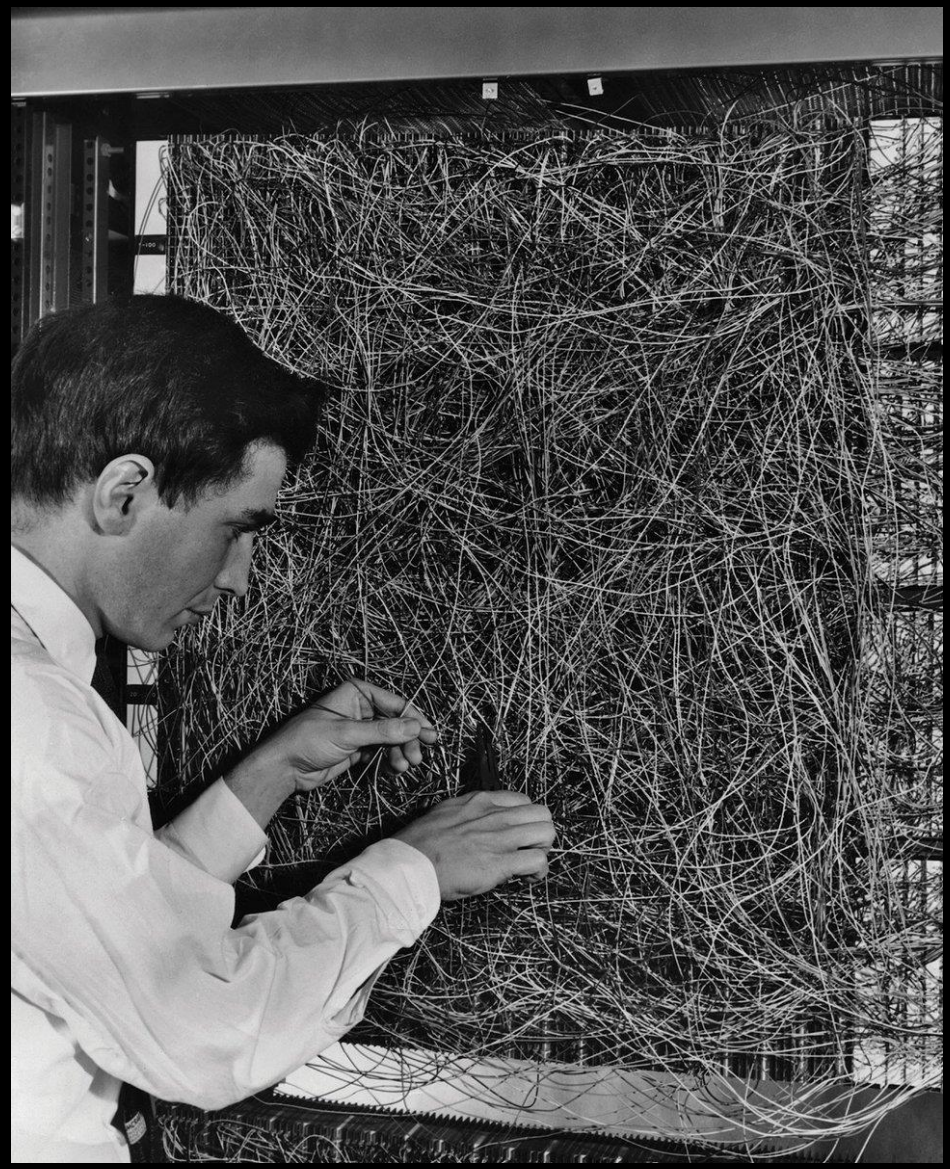


Image Source: https://en.wikipedia.org/wiki/Neuron

First Perceptron

- Perceptrons were first created in 1958
- It was made from copper wires.
- No mass progress made for 40 years.



Mass Street University
PER EDUCATIONEM PROGRESSUS

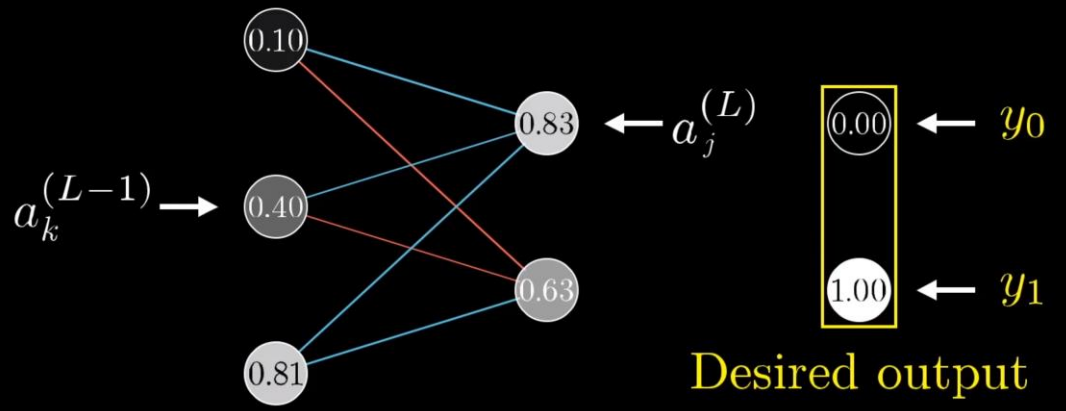
Modern Perceptrons / Neural Networks

- They use numerical representations.
- Made up of 6 parts.
- Considered ‘Black Box’ models.
- Uses Hidden Layers.

Output Layers

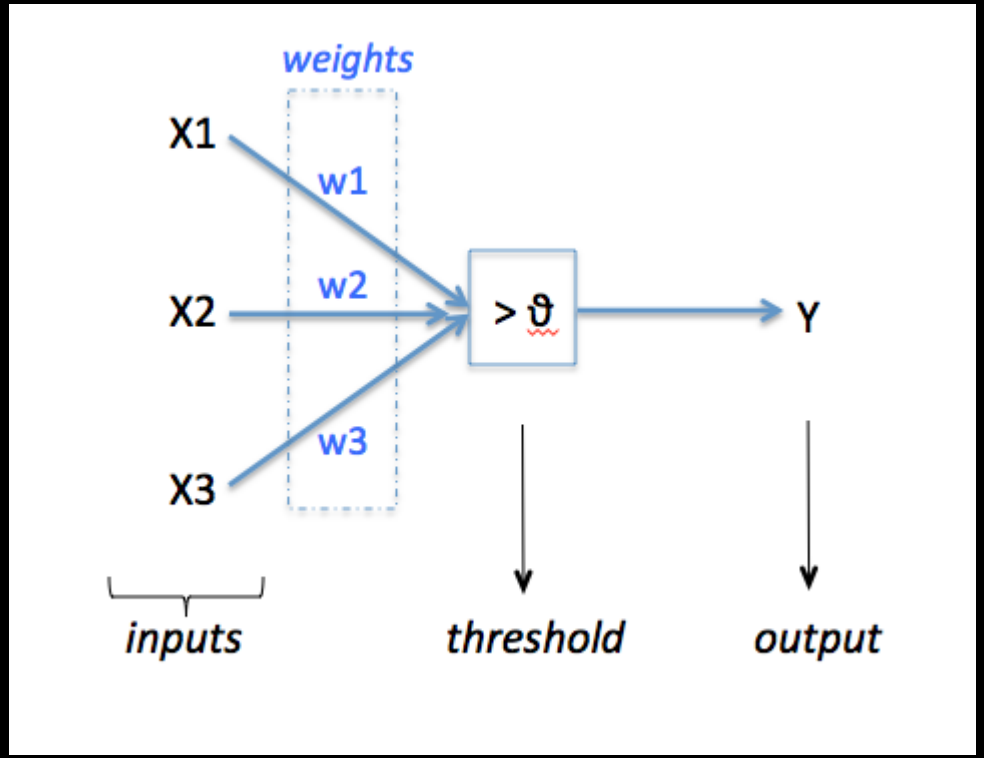
$$C_0 = \sum_{j=0}^{n_L - 1} (a_j^{(L)} - y_j)^2$$

- Target variables.
- The number of output neurons is based on the number of target variables or ‘classes’ in out data.
- Probabilities are between 0 and 1.



Weights and Biases

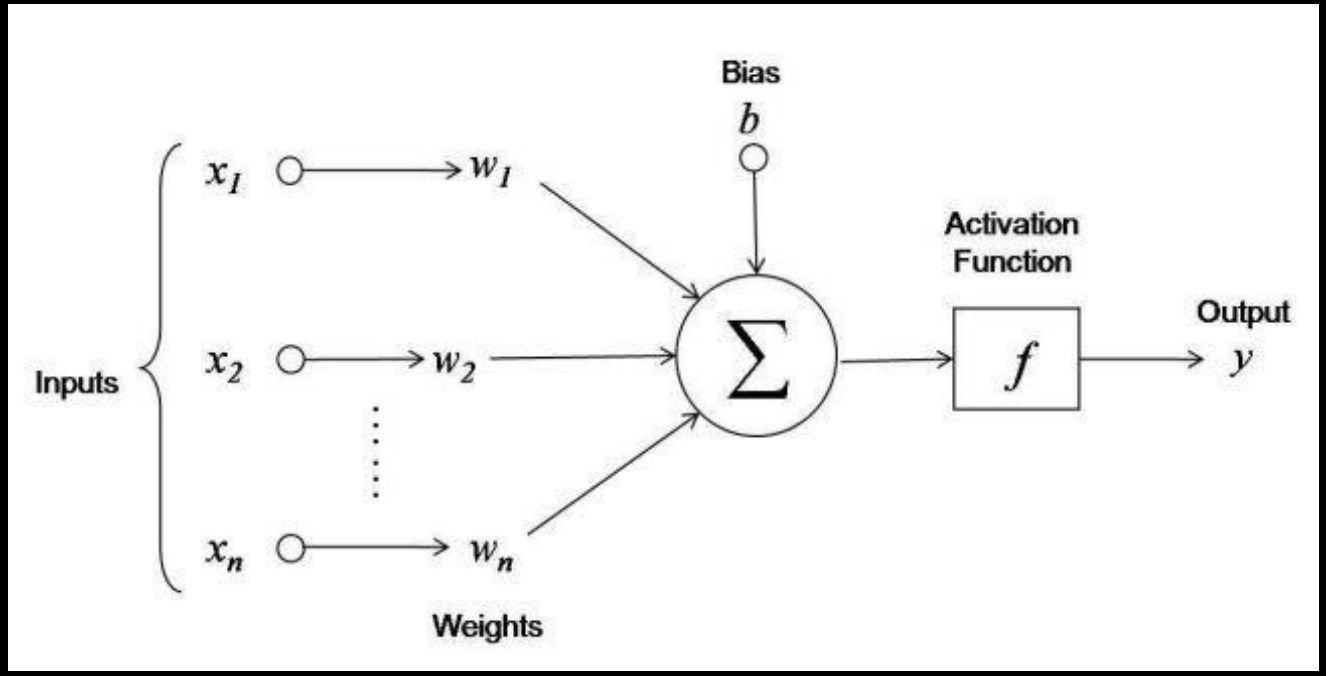
- It is how we connect neurons and control values in our network.
- Updating weight values is how neural networks learn.
- Biases adjust how the model fits the data.



Mass Street University
PER EDUCATIONEM PROGRESSUS

Activation Functions

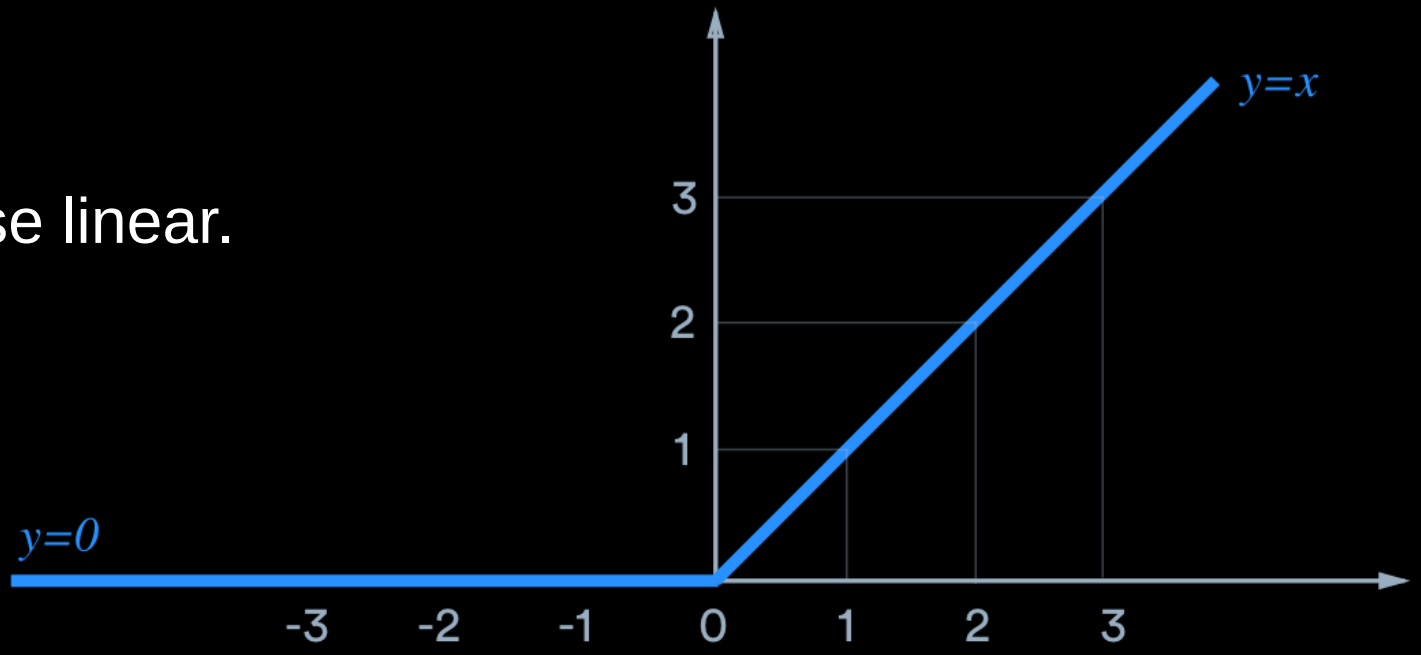
- Determine how the neurons fire and function.
- They are applied after every layer.
- Biases affect the activation value inside activation functions.



Mass Street University
PER EDUCATIONEM PROGRESSUS

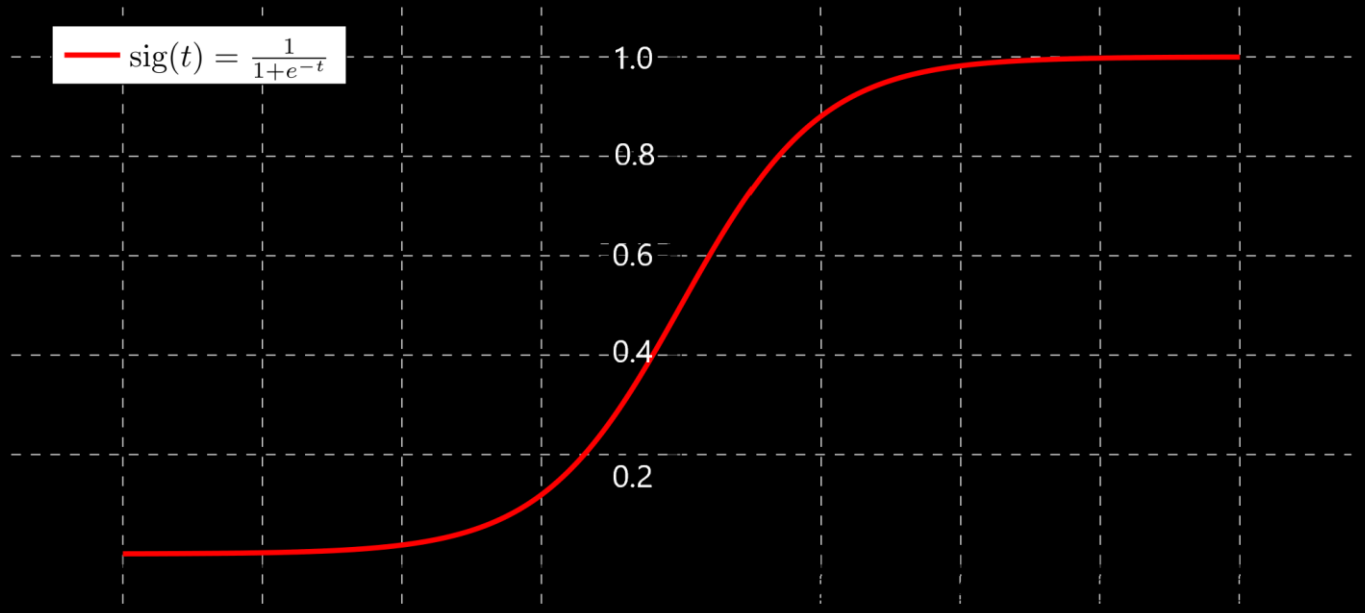
ReLU

- Clipping function.
- Zero when negative, otherwise linear.



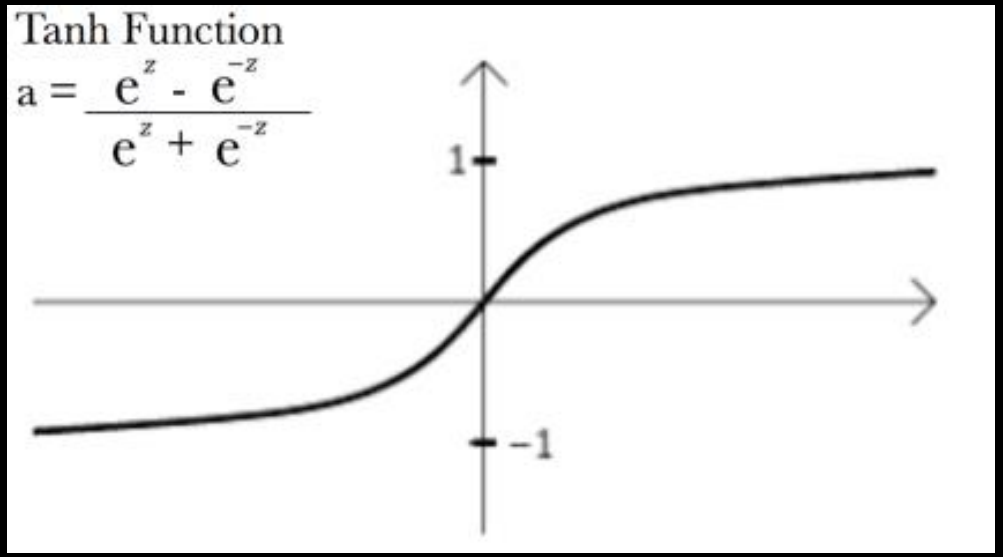
Sigmoid

- Squishes values between zero and one.
- Causes Exploding and Vanishing Gradients.



TanH

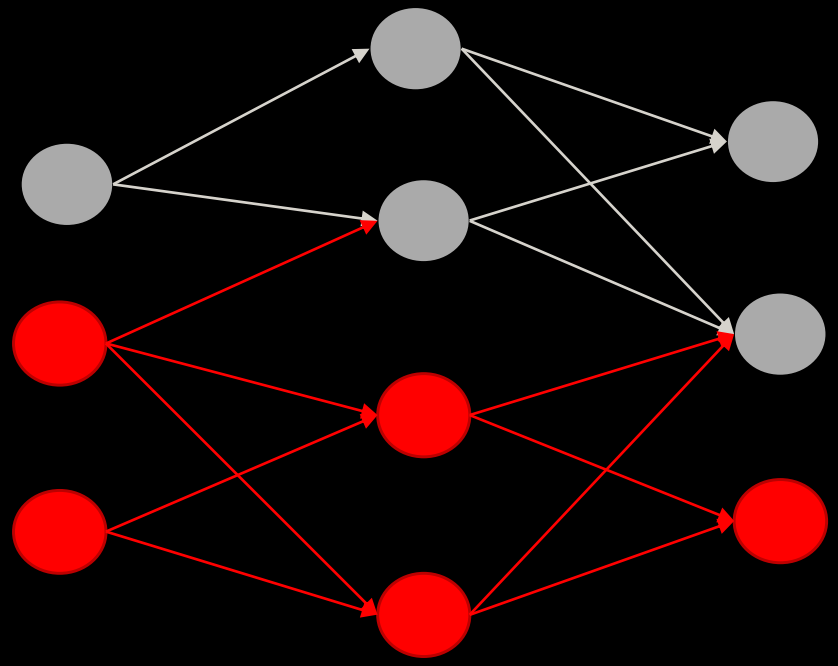
- Zero Centered.
- Less prone to exploding and vanishing gradients.
- Squishes numbers from -1 to +1.



Mass Street University
PER EDUCATIONEM PROGRESSUS

Dead Neurons

- Dead neurons kill other neurons later on in the network even if live neurons are connected to them.
- Once a neuron dies it stops the training process for itself and every neuron after it.



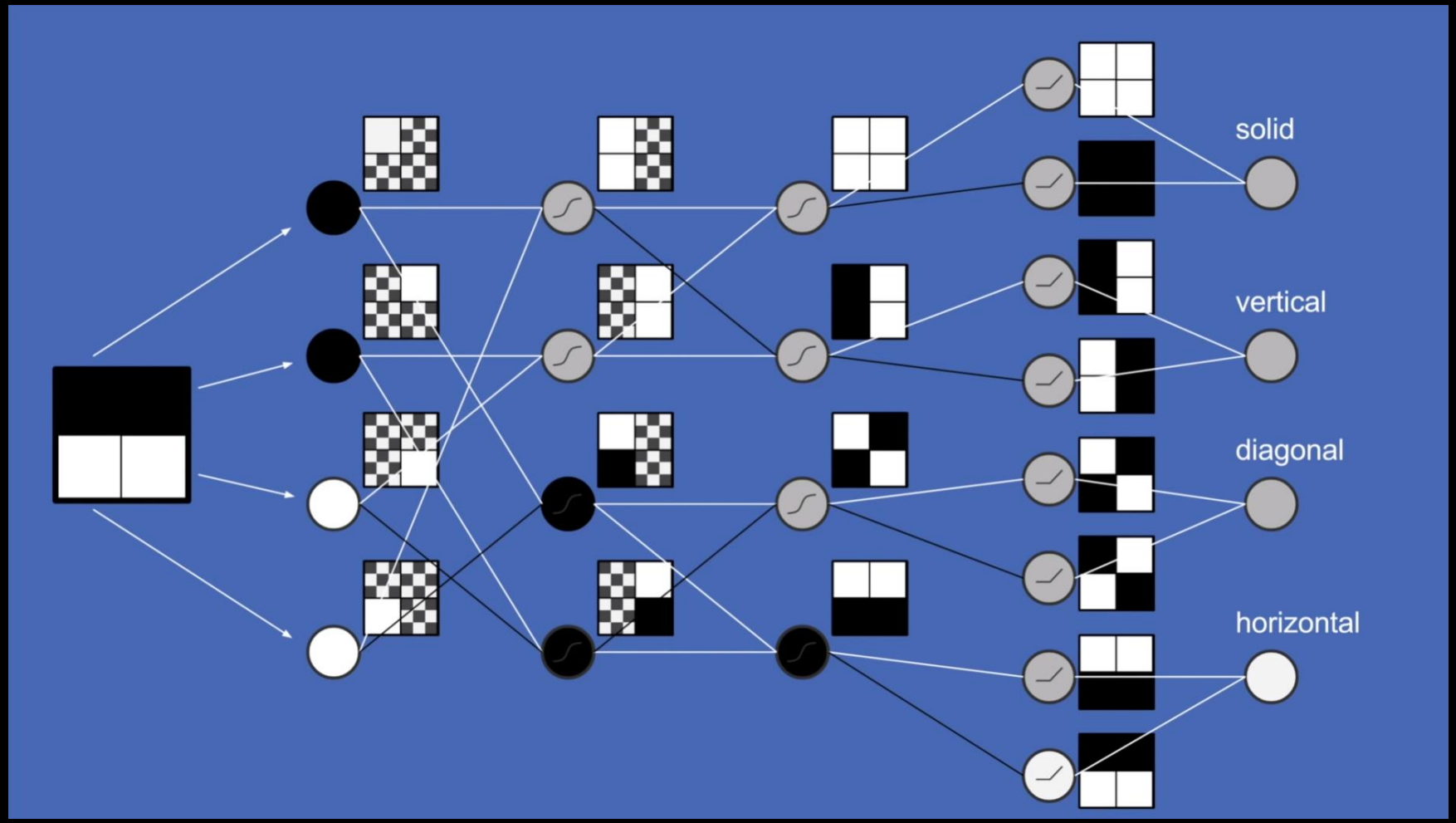
Output layer

- Uses a softmax or linear output.
- Softmax is used for discrete classes.
- Linear is used for continuous variables and regression.

Neural Network Example

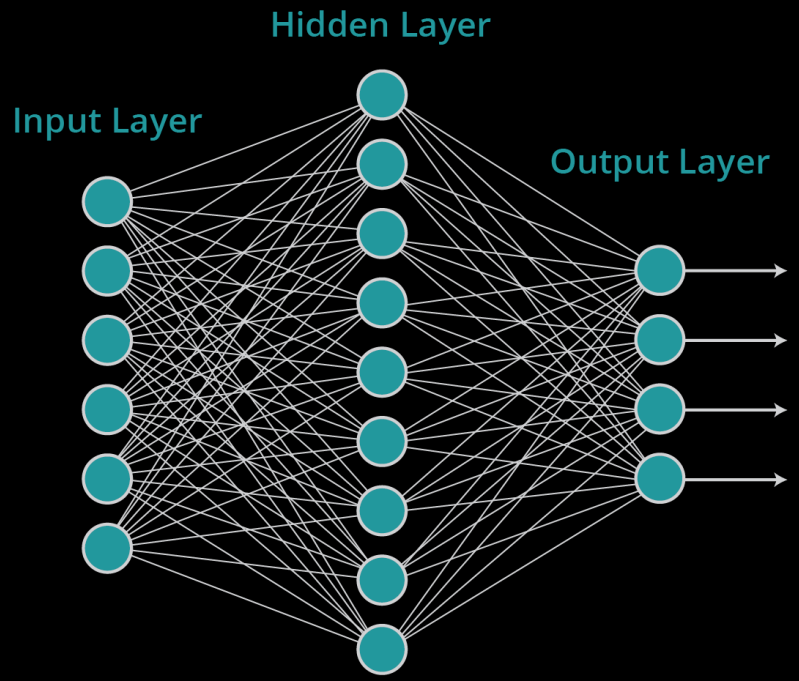
- Black = -1
- White = 1
- Gray = 0
- Classification

  should be class 8

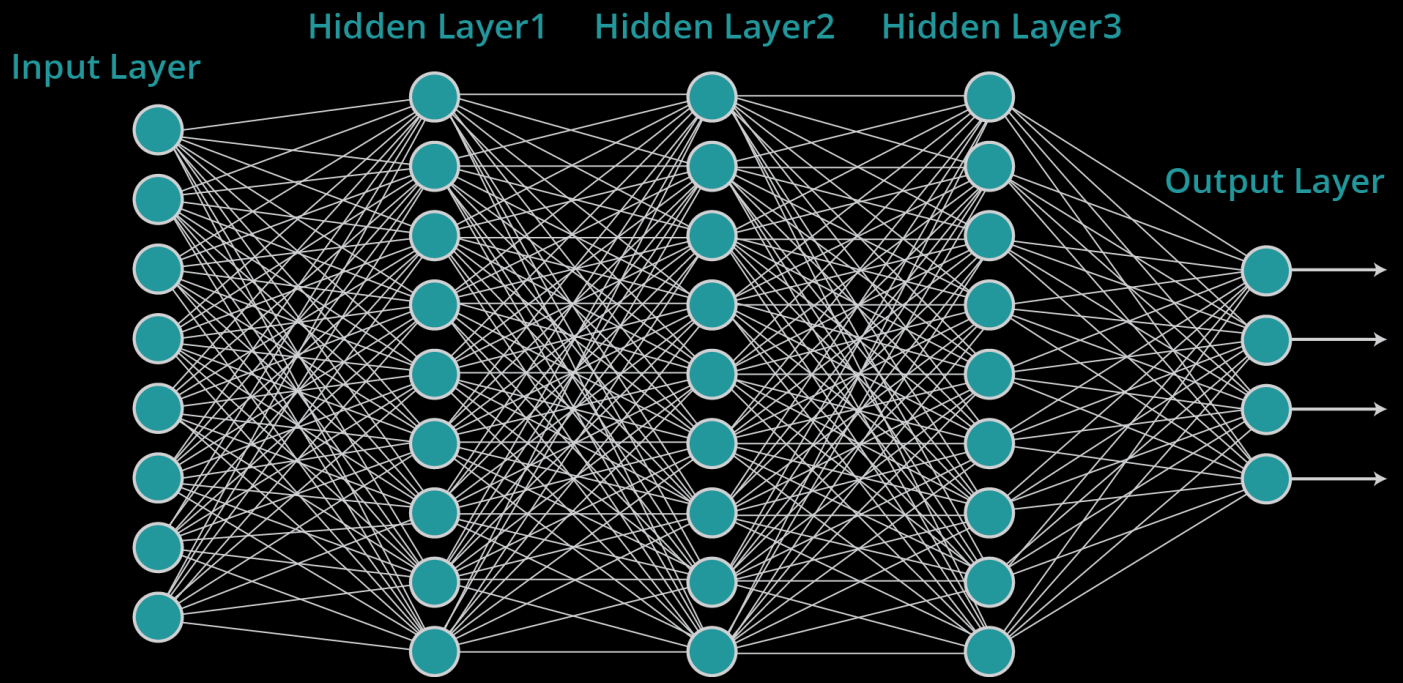


Actual Neural Networks

“Non-Deep” Feedforward Neural Network



Deep Neural Network



Mass Street University
PER EDUCATIONEM PROGRESSUS

Equations used in Deep Learning

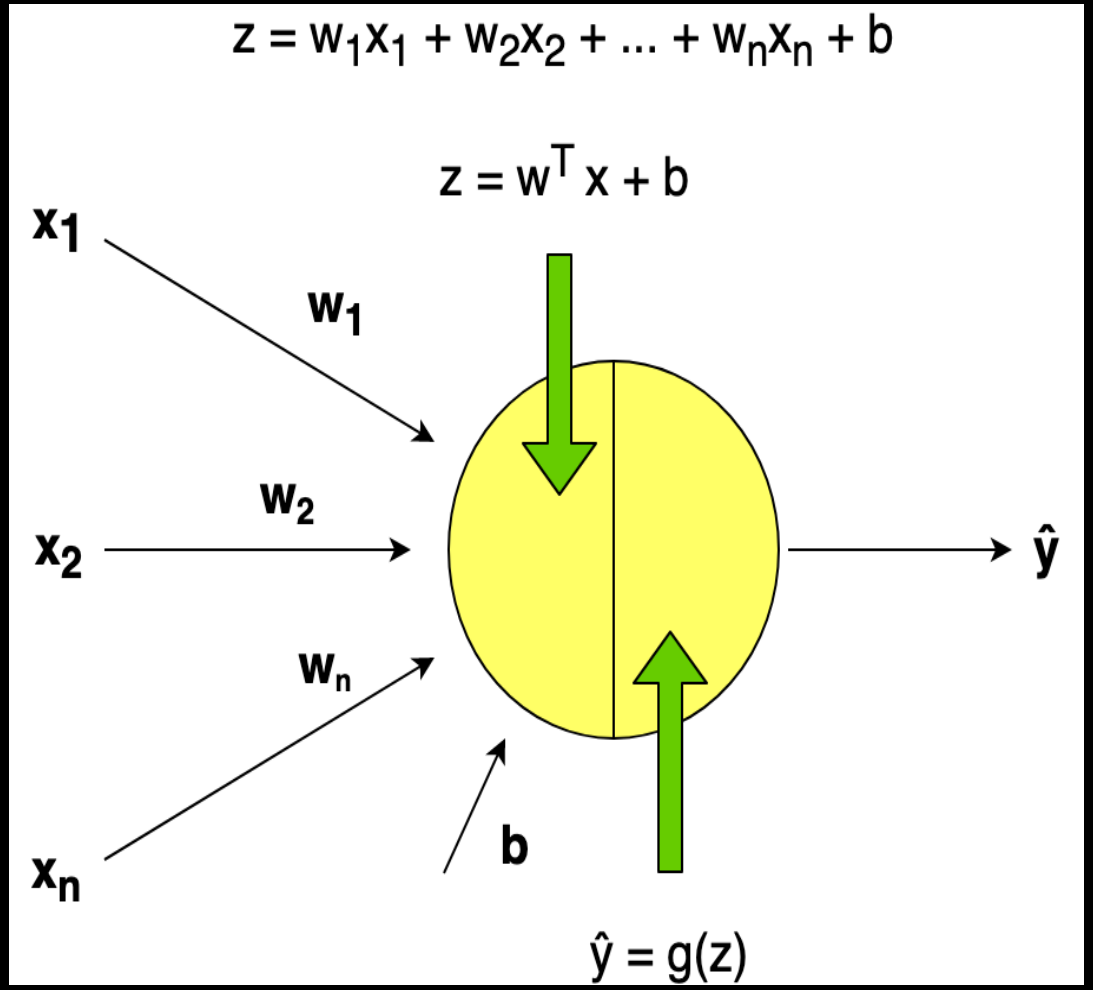
- Neuron Weight function.
- Activation Function.
- Output Function.
- Back Propagation.
- Loss Function.
- Dropout.
- Normalization.
- Regularization.
- Optimizers.

Mass Street University

PER EDUCATIONEM PROGRESSUS

Neuron Weight Function

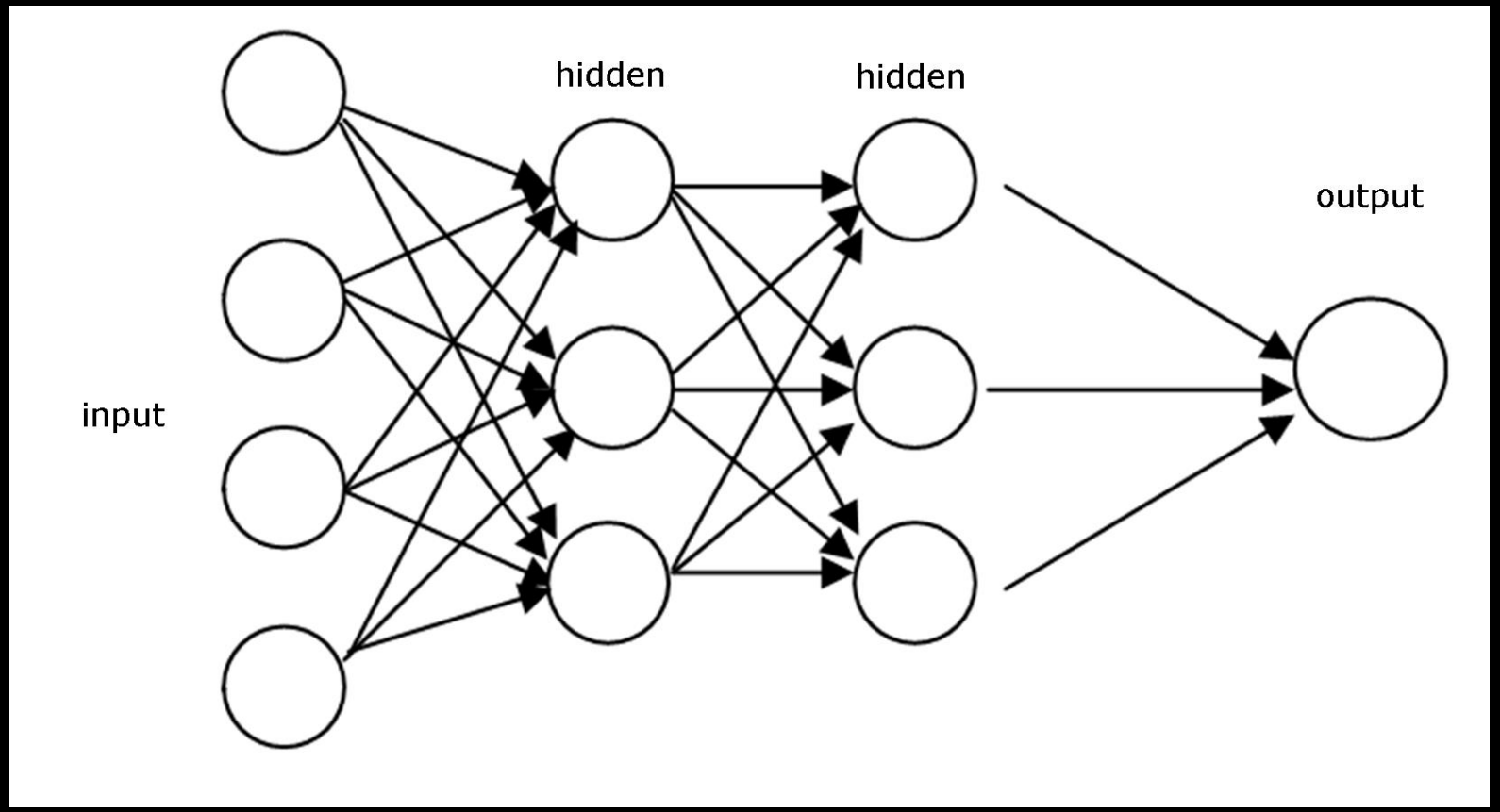
- Where $\hat{y}$ =A_n and z = A_{in}
- $A_{in} = A_n * W_n + B$
- $A_{in} = (A_1 * W_1) + (A_2 * W_2) + (A_3 * W_3) + B$



Mass Street University
PER EDUCATIONEM PROGRESSUS

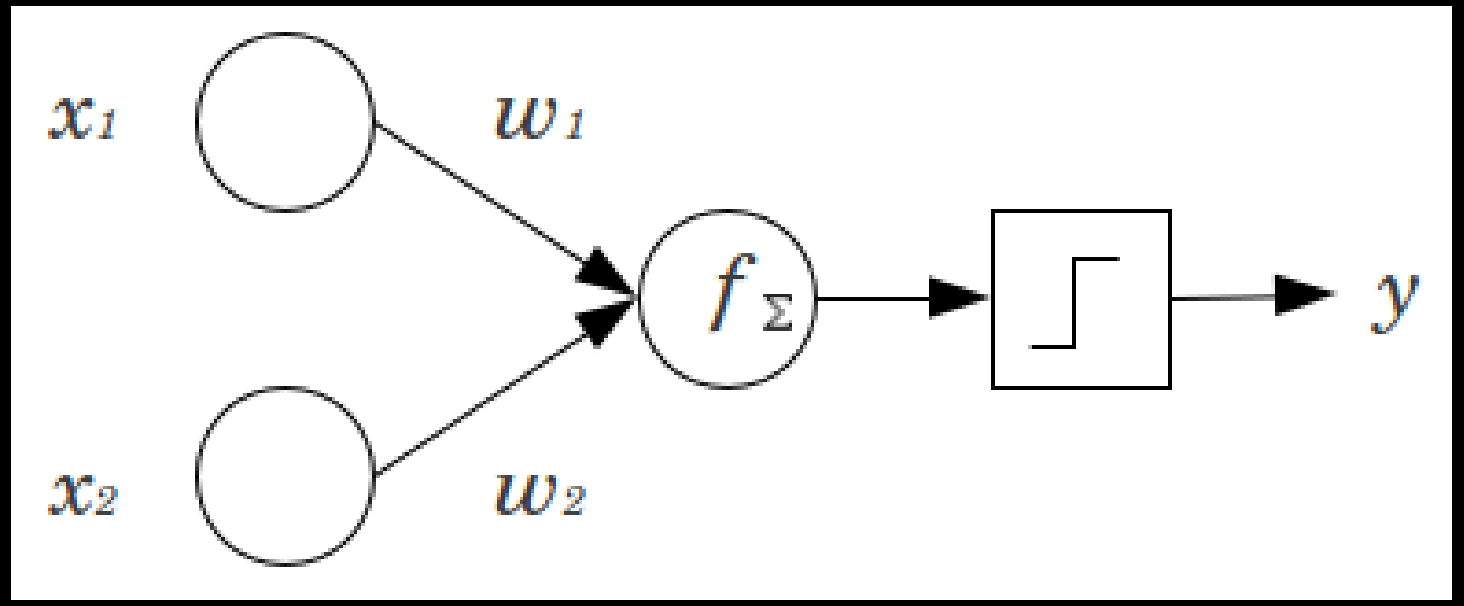
Feed Forward Network

- Only travels in one direction.
- Weights are updated by hand.



Perceptrons

- Most basic form of neural networks.
- Doesn’t use biases.
- Many perceptrons added together create modern day neural networks.



Mass Street University
PER EDUCATIONEM PROGRESSUS