Intro to Deep Learning

General Assembly Data Science Cohort 40

What do you know about deep learning?
What terms come to mind?

Deep Learning: Definition

- Subset of machine learning that attempts to model high level abstractions in data in order to vastly improve performance in both supervised and unsupervised learning.
- It achieves this by using multiple layers of "processors", each of which contains a set of non-linear transformation functions that learn representations within the data.

Deep Learning: Definition(Cont)

- This approach is mainly motivated by how we believe that the brain works.

- We learn simple ideas and then use these simple ideas to form hierarchies of more complex ideas.

- Deep learning is about applying this approach to machine learning tasks.

Examples of Deep Learning Applications

- Face filters on Snapchat and Instagram stories.

- Google's Deep Dream project
- Generating text descriptions from photos
- Language translation

- Self-driving cars

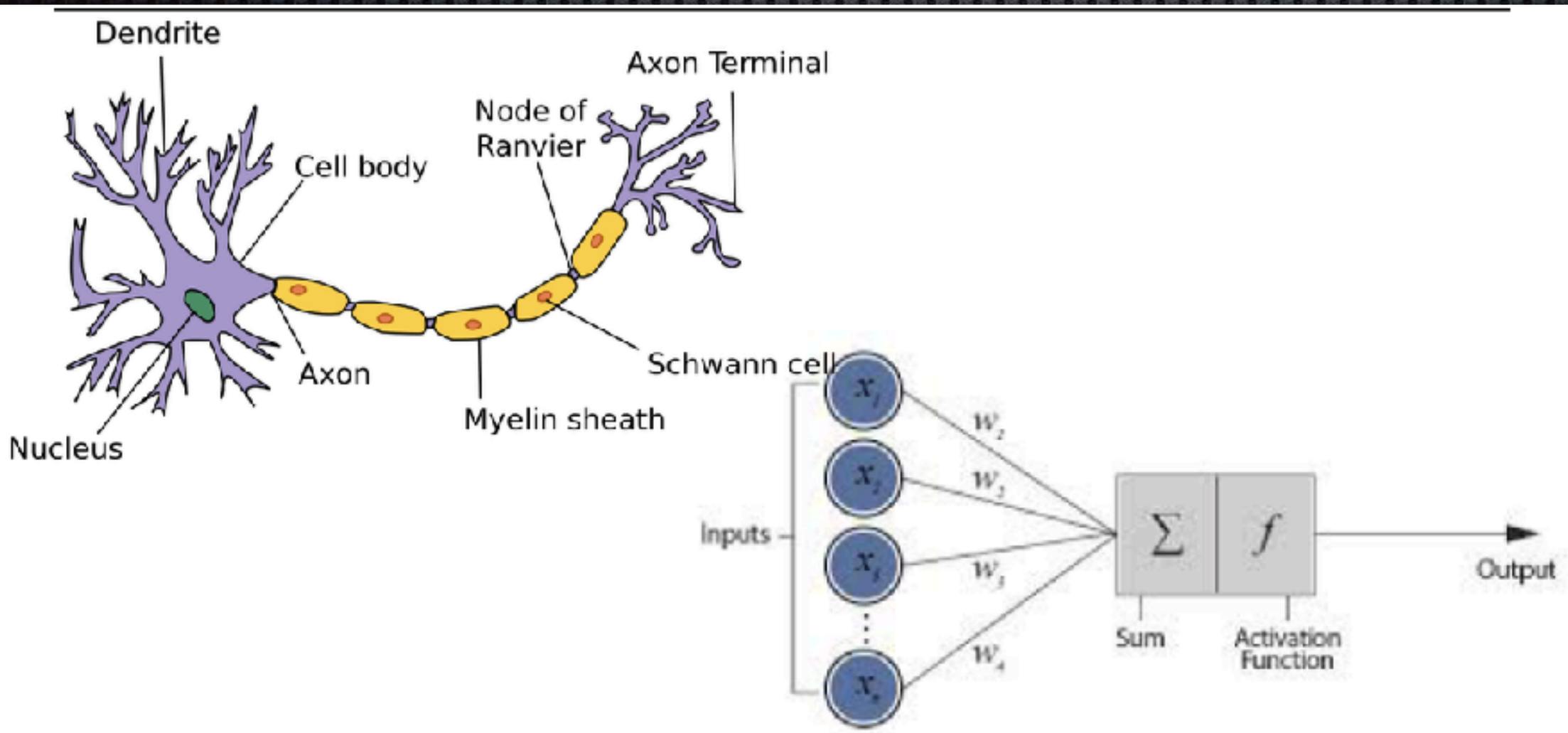
- Sound, image generation (music generated by AI)

Deep Learning Libraries in Python

- TensorFlow
- Keras
- PyTorch
- Theano
- Caffe

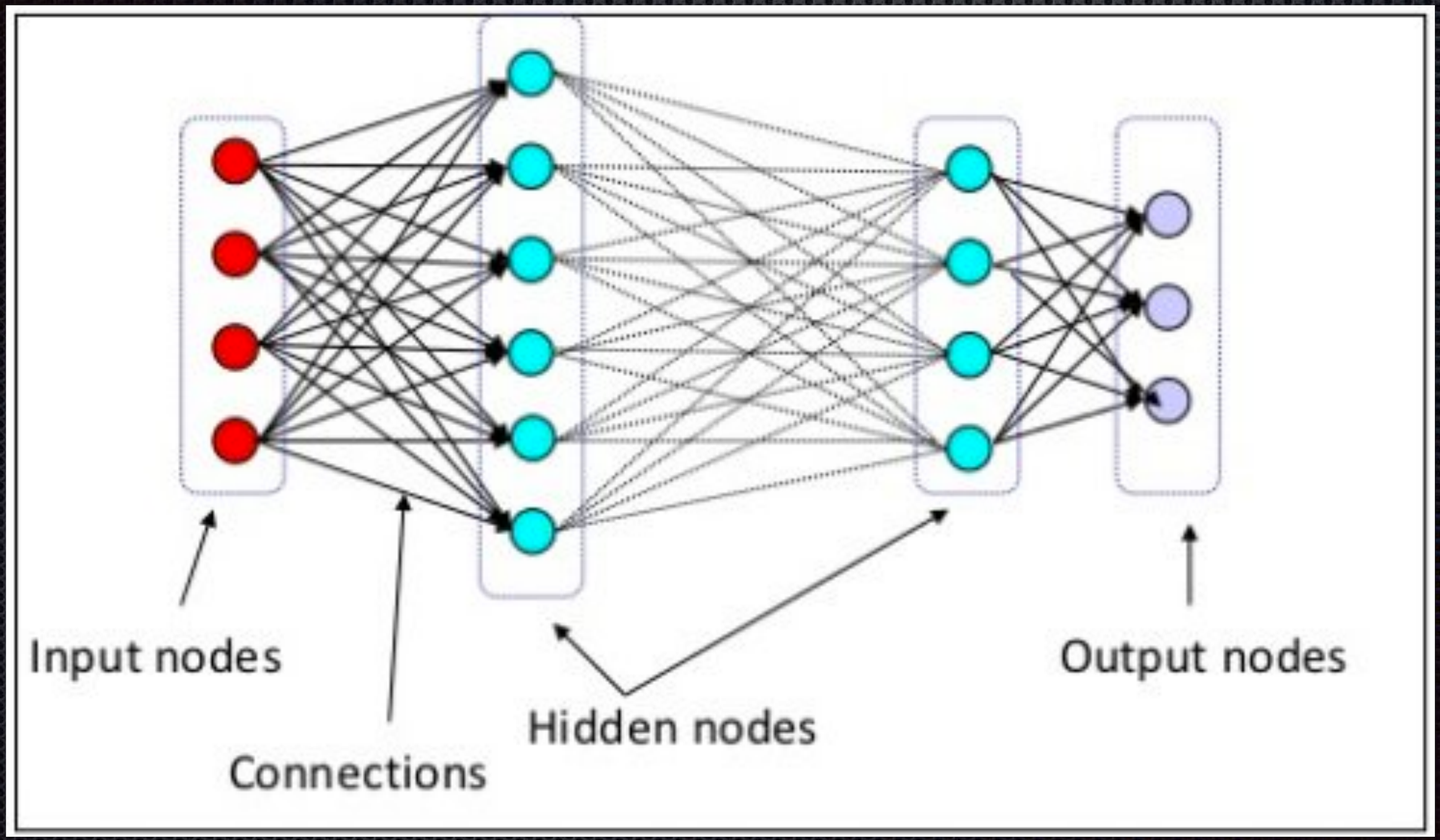
Artificial Neural Networks

- A computational system comprised of layers and each layer is built of interconnected/multilayer perceptrons.
- Built to model an animal’s nervous system .


Dendrite
Axon Terminal
Node of
Ranvier
Cell body
Axon
Schwann cell
Myelin sheath
Nucleus
Inputs
Sum
Activation
Function
Output

How Deep Learning Works

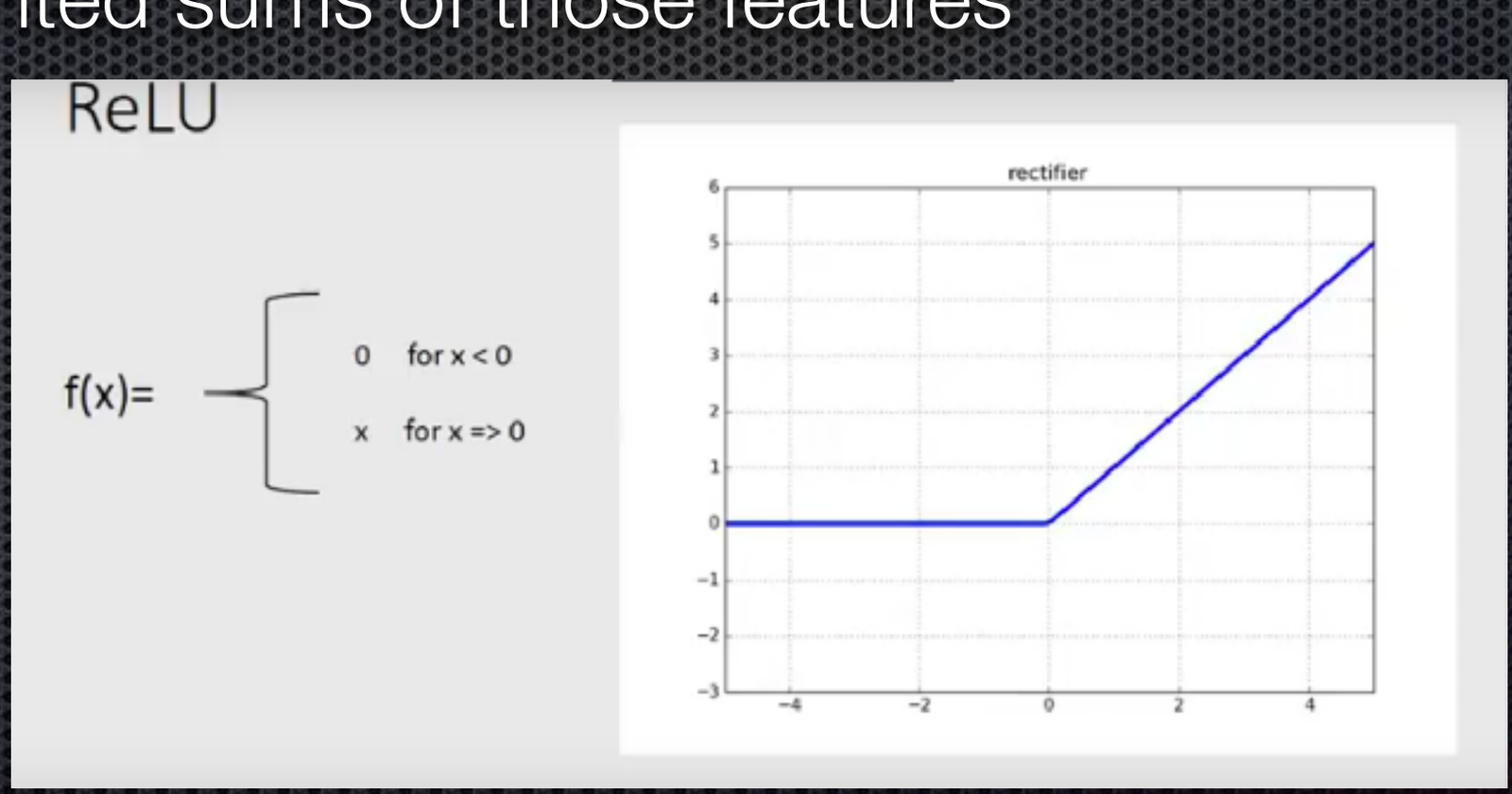
Layers



Source: KDNuggets

Activation Functions: ReLu

- Takes in input and pass it through the layer's preassigned activation function.
- That input could be either the features themselves or weighted sums of those features

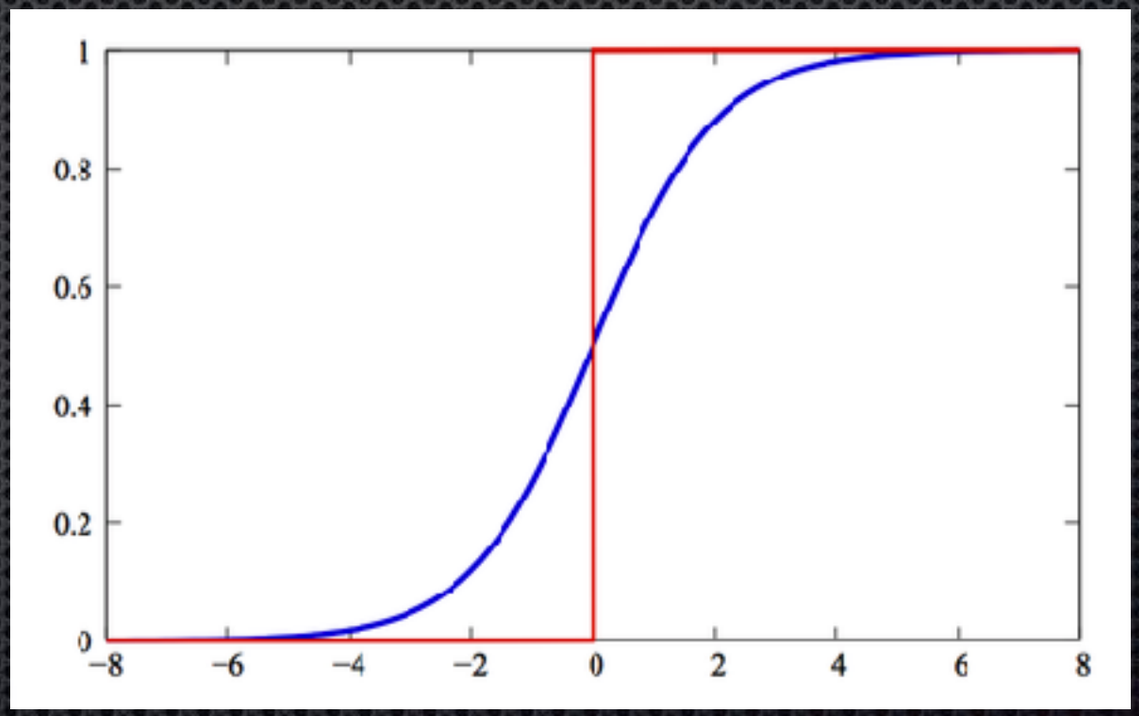


Source: Towards Data Science

Sigmoid Function

$$f_{log}(z) = \frac{1}{1 + e^{-z}}$$

f_{log} is called logistic function



Softmax Function

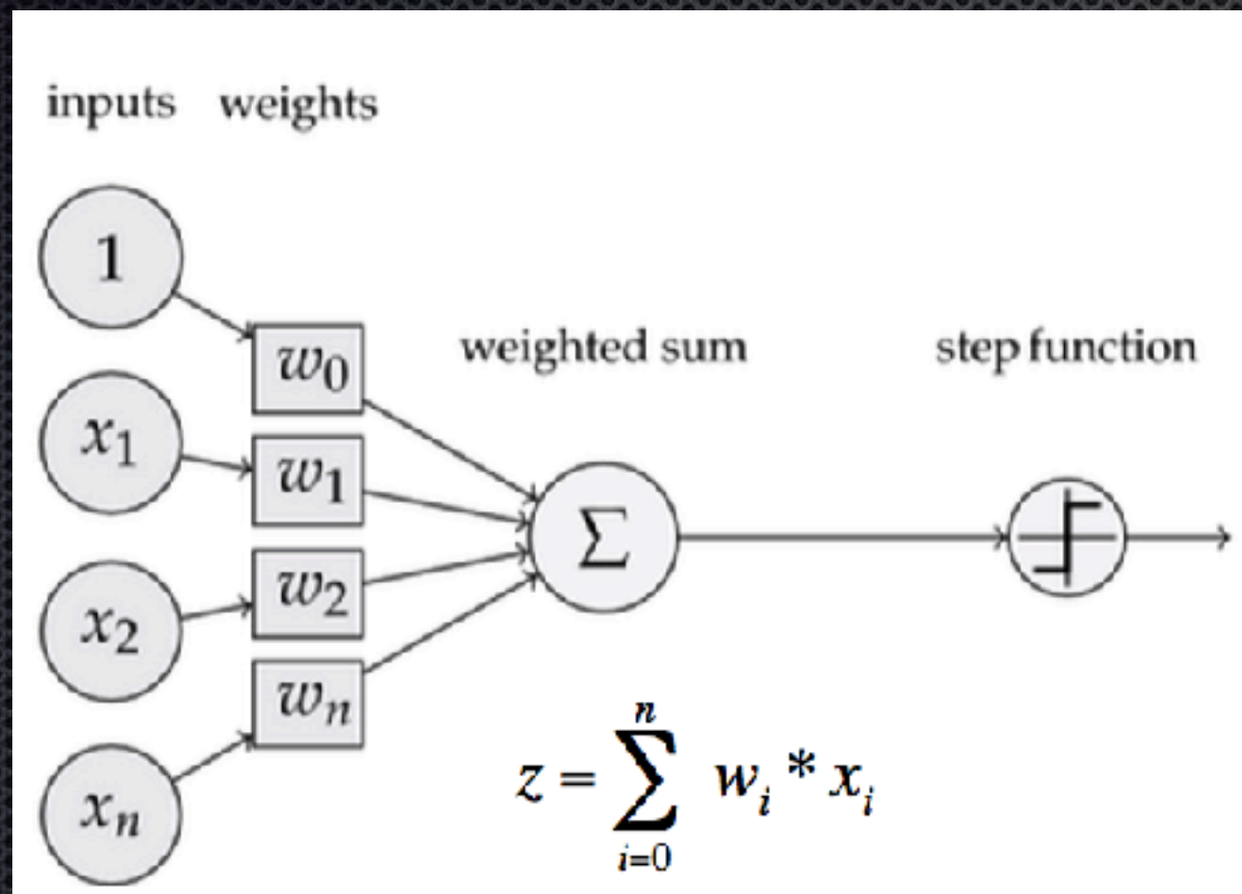
Softmax Function

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$$

- A generalization of the logistic function.
- the output of the softmax function is used to represent a categorical distribution, a probability distribution over K different possible outcomes(categories)

Source: Hyunsung Lee

Simple Neural Network with Sigmoid Activation Function

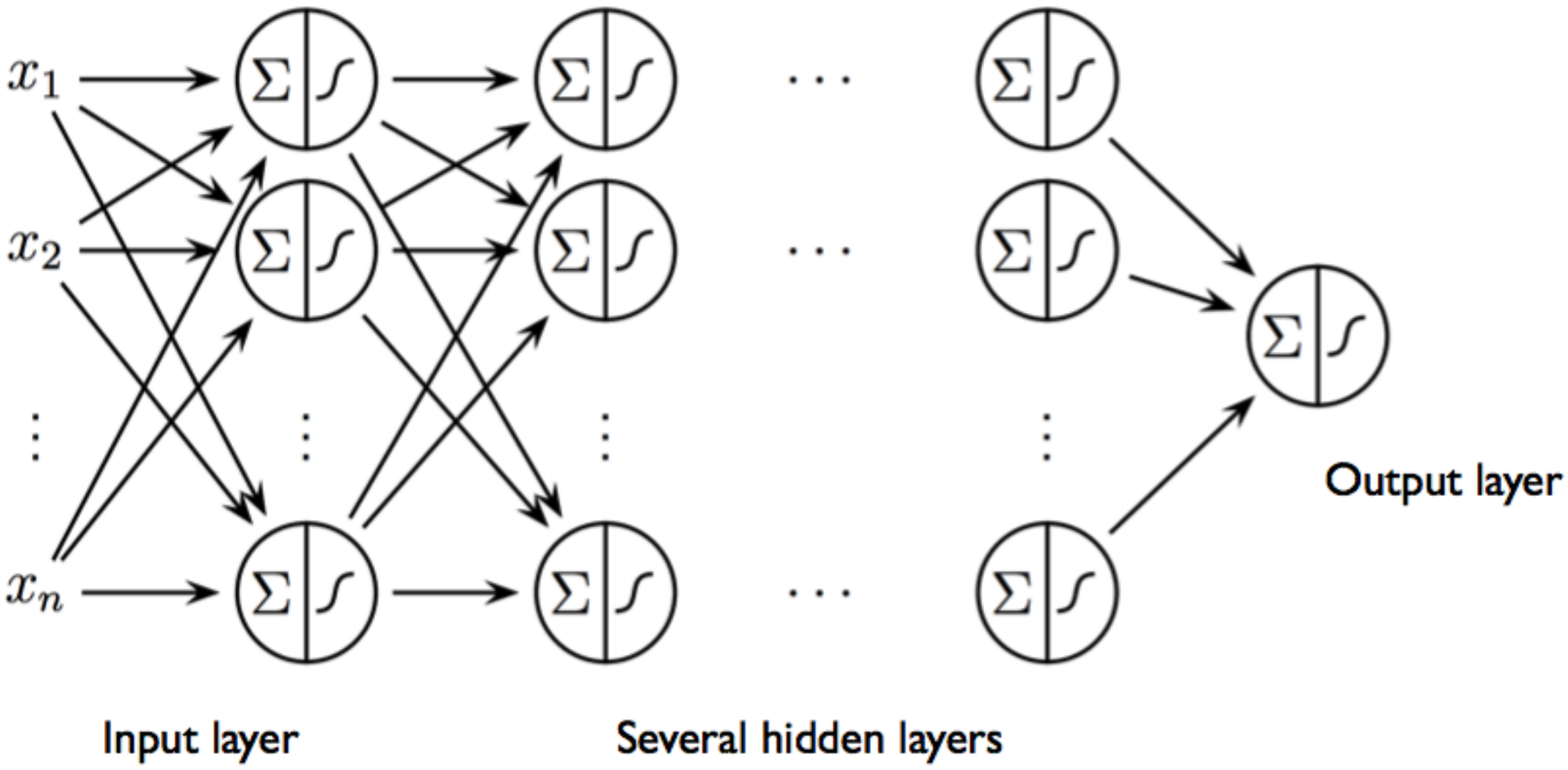


$$f_{log}(z) = \frac{1}{1 + e^{-z}}$$

f_{log} is called logistic function

Multilayer Perceptrons

A multi layer perceptrons (MLP) is a finite acyclic graph. The nodes are neurons with logistic activation.



But How Does it Learn???

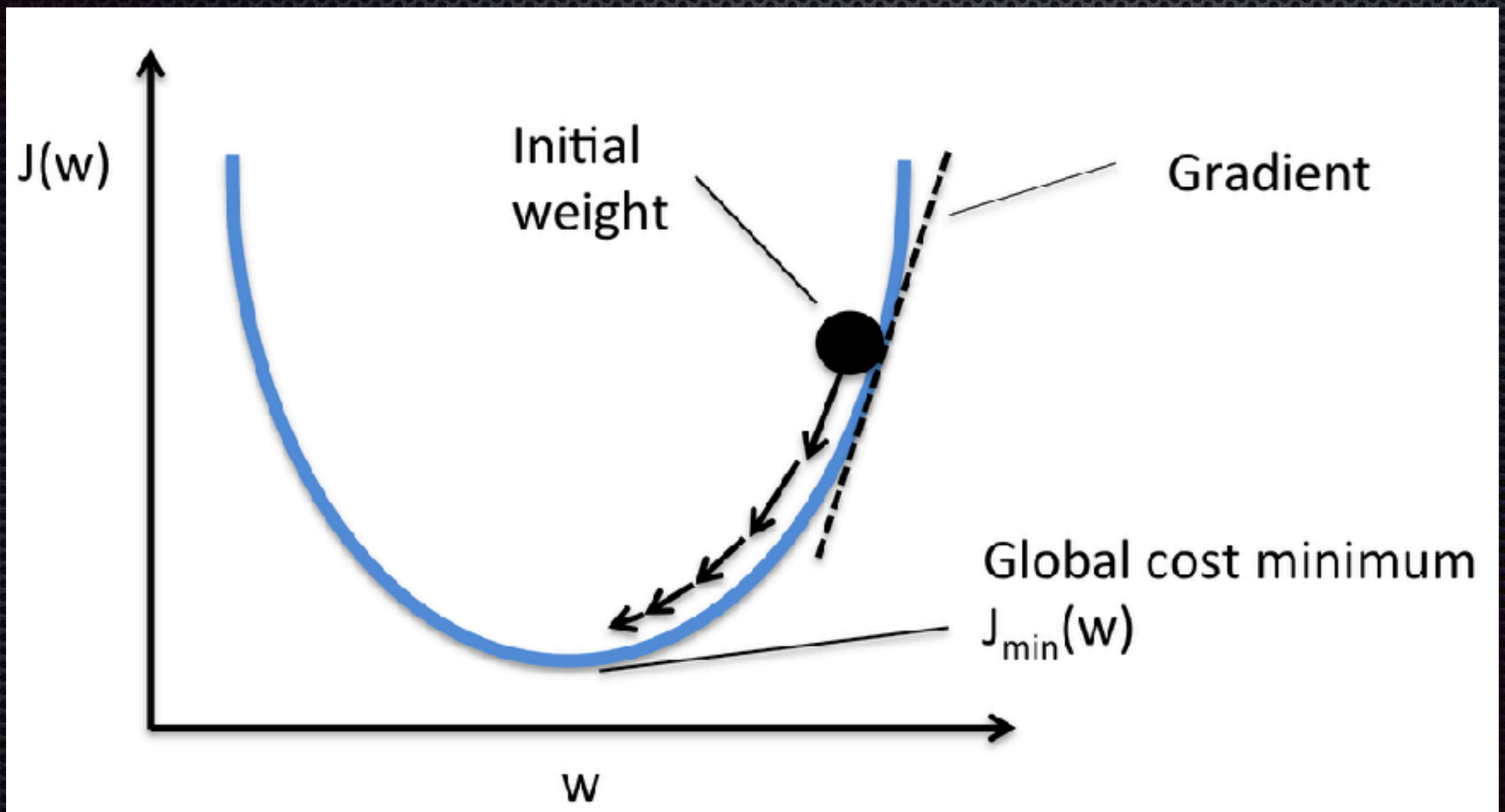
Forward Propagation

- In our example of the sigmoid function, we witnessed an example of forward propagation in which we took inputs multiplied them by weights, summed those products, and passed them through an activation function (sigmoid.)

Backpropagation

- How an algorithm learns from it mistakes by optimizing the weights.
- Forward propagate with a neural network.
- Make predictions with the algorithm and use that to calculate the error.
- The values of all the weights all have an influence on the cost, which means there's a relationship. This means there's a gradient (slope) which requires calculus to find the minimum cost. Gradient descent finds the derivative of the loss function for each weight.
- Backpropagating goes back to updating the weights using gradient descent.

Gradient Descent



Source: Sebastian Raschka

Various Types of Deep Learning

Supervised Learning

- **General:** Deep Belief Networks, Multilayer Perceptrons
- **Text:** Recurrent Neural Net
- **Image processing/object recognition:** Convolutional Neural Net, Deep belief net
- **Audio/Speech recognition:** Recurrent Neural Net
- **Time Series:** Recurrent Neural

Unsupervised Learning

- Used for feature engineering, pattern detection, and dimensionality reduction
- **Tools:** Autoencoders,Restricted Boltzman Machine

???Questions???