
ARTIFICIAL NEURAL NETWORKS

DEEP LEARNING

Deep Learning

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

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.

Deep Learning

Auto Feature-Engineering

Deep Learning

Choosing optimal features can be very difficult and time consuming. Features often need to be crafted and then altered through dimensionality reduction or transformation in order to improve performance.

Examples in NLP, speech recognition, image recognition

Deep Learning

Deep learning tries to replace this time-consuming task of *feature engineering* via the introduction of efficient network architectures for unsupervised *feature learning*.

Deep Learning

A full list of deep learning modules can be found [here](#).

The major players are [Tensorflow](#), [Torch](#), [Theano](#), [Caffe](#).

ARTIFICIAL NEURAL NETWORKS

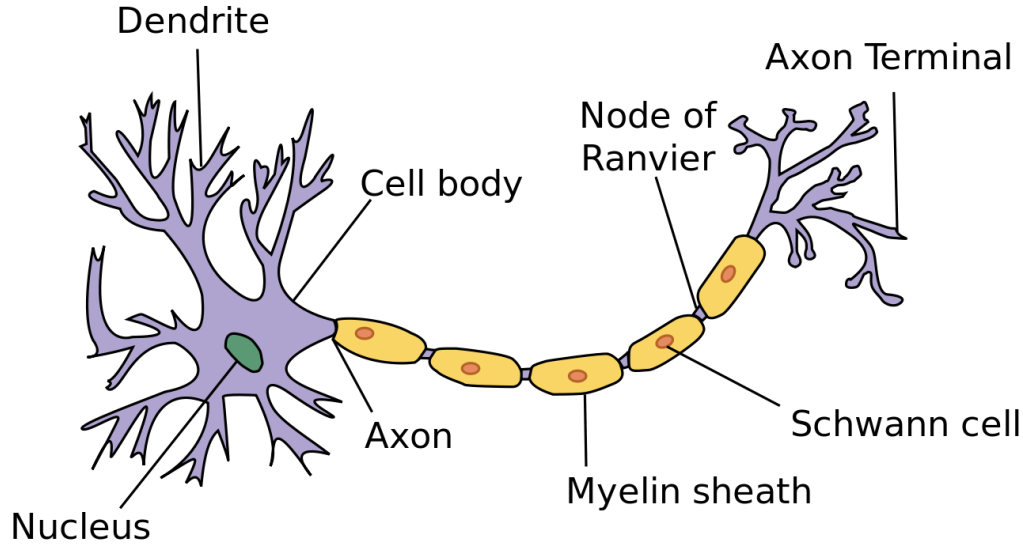
ARTIFICIAL NEURAL NETWORKS

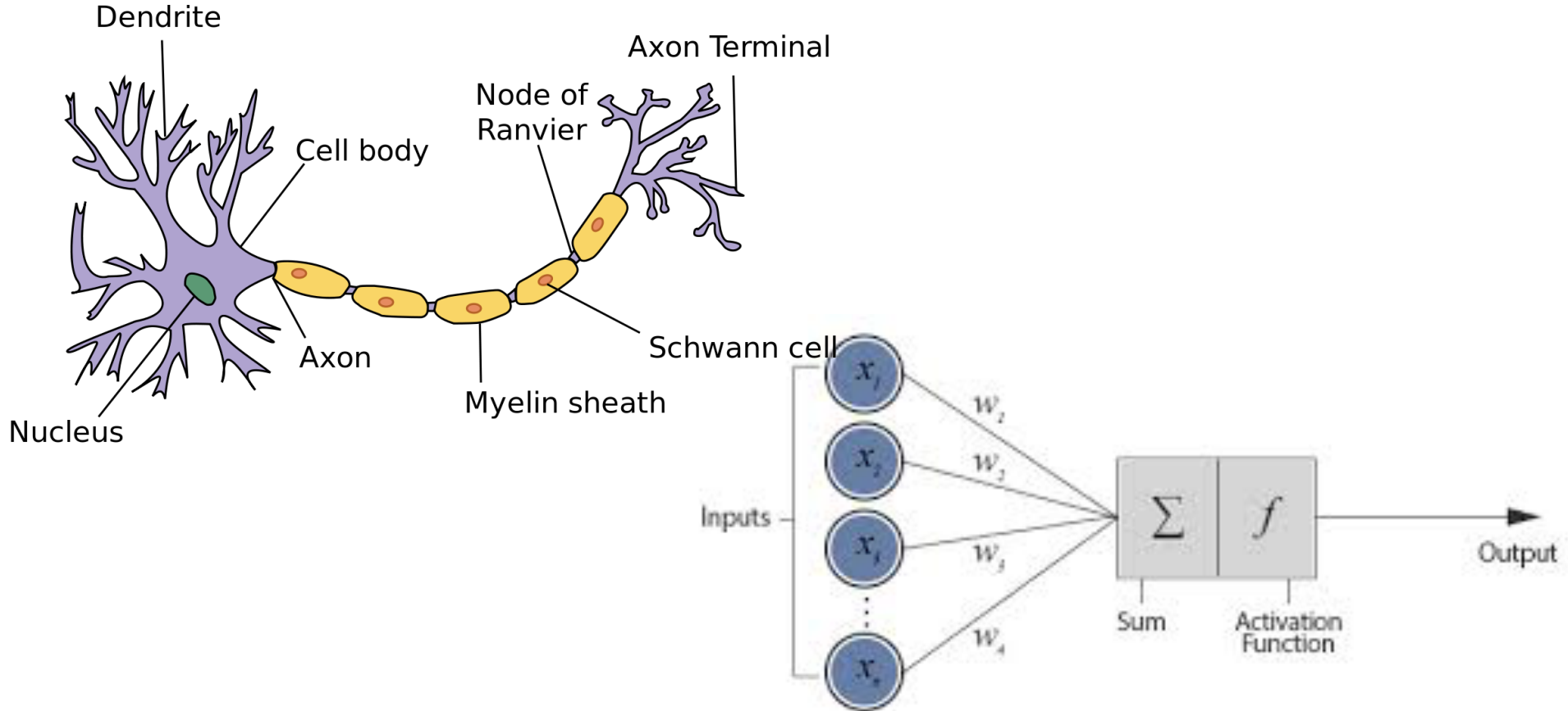
Artificial Neural Networks

A computational system
comprised of layers and each layer
is built of interconnected perceptrons

Artificial Neural Networks

Built to model the animal nervous system





Artificial Neural Networks

A computational system
comprised of layers and each layer
is built of interconnected perceptrons

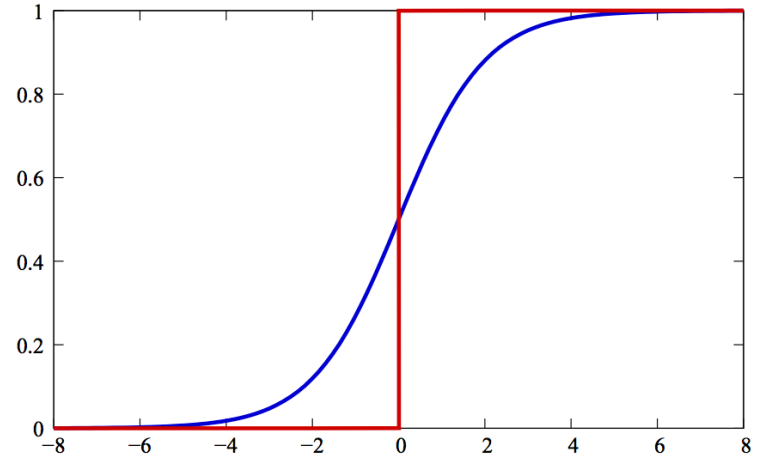
Single Perceptron

Takes in input and uses an activation function in order to output

Single Perceptron

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

f_{log} is called **logistic function**

**NOTE:**

A single perception can be like a logistic regression in and of itself!

Takes in input and uses an activation function in order to output

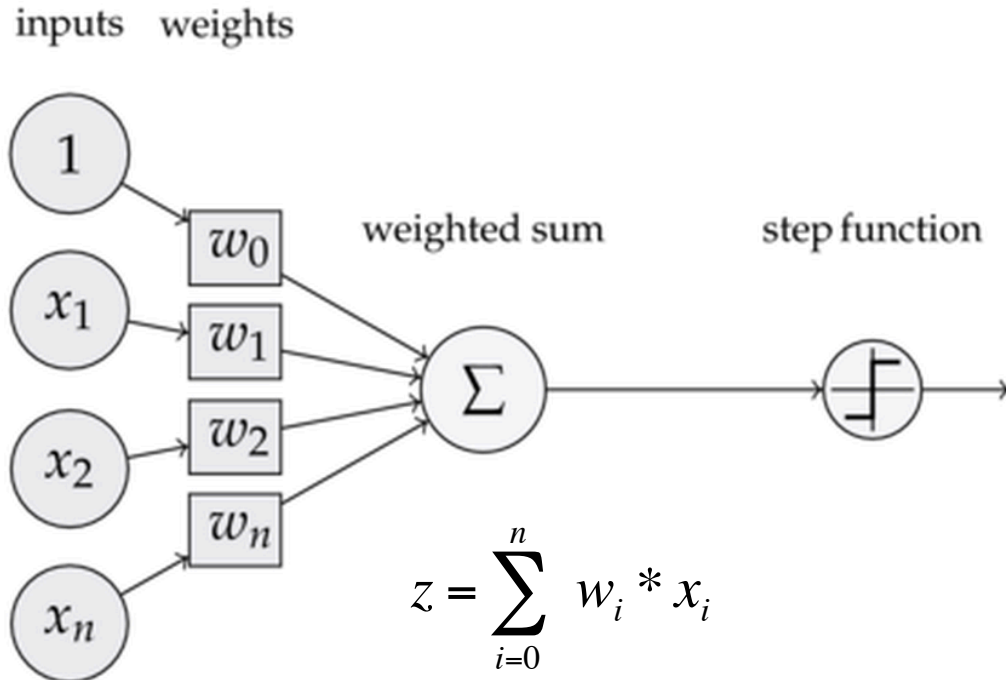
Single Perceptron

But what is z ? A weighted sum on the inputs!

$$z = \sum_{i=0}^n w_i * x_i$$

Where w is the weight on input x

Single Perceptron

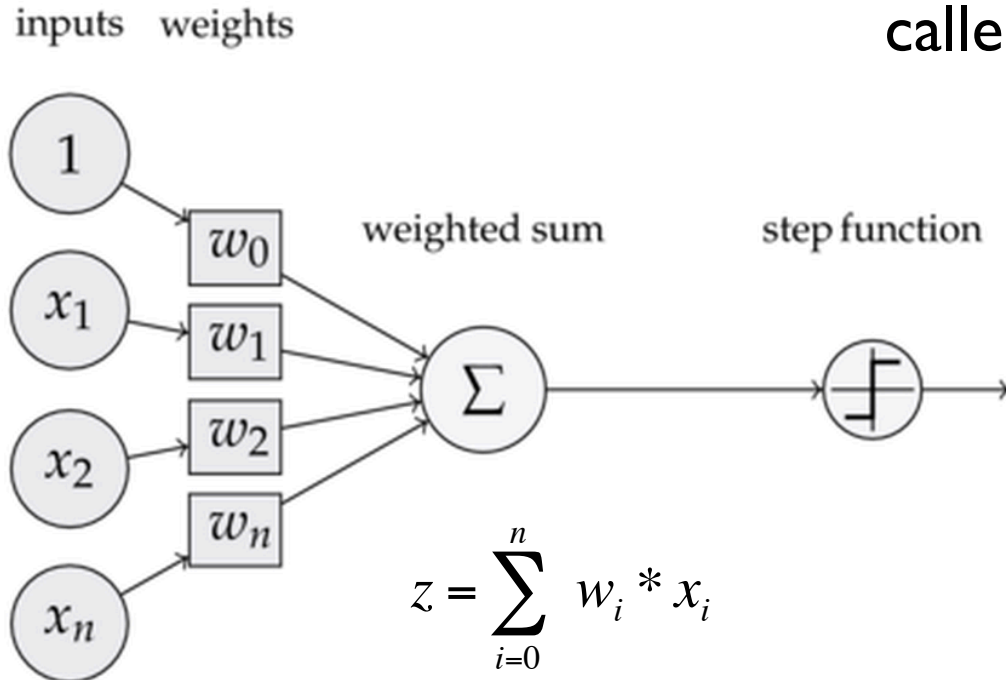


$$f_{log}(z) = \frac{1}{1 + e^{-z}}$$

f_{log} is called **logistic function**

Single Perceptron

If $f(z)$ is above a threshold, generally called theta, then the neuron “fires”

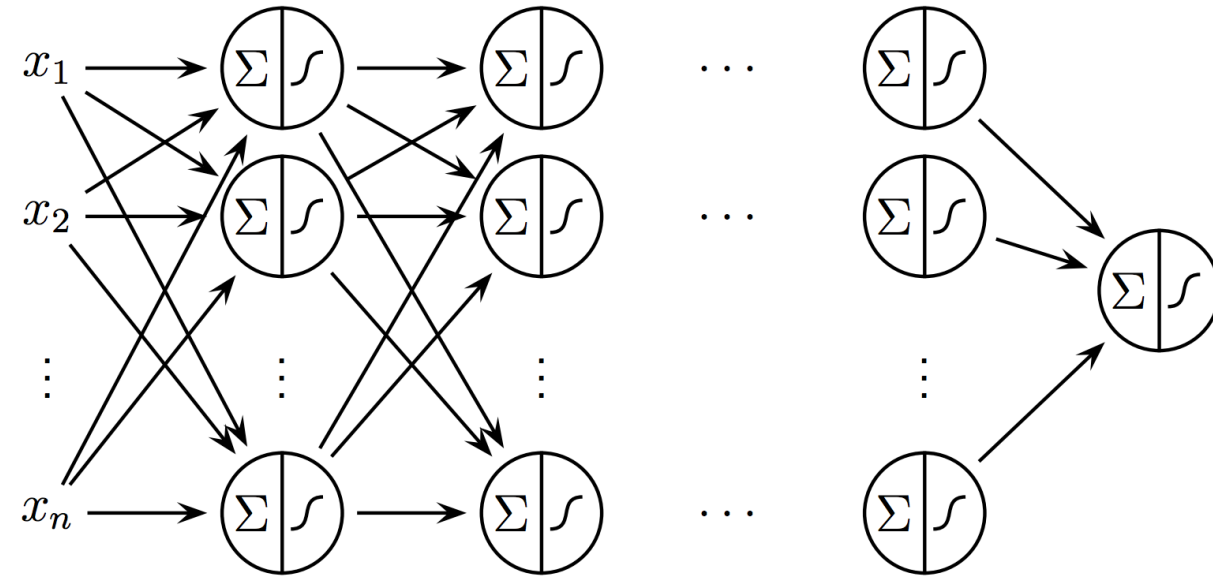


$$f_{log}(z) = \frac{1}{1 + e^{-z}}$$

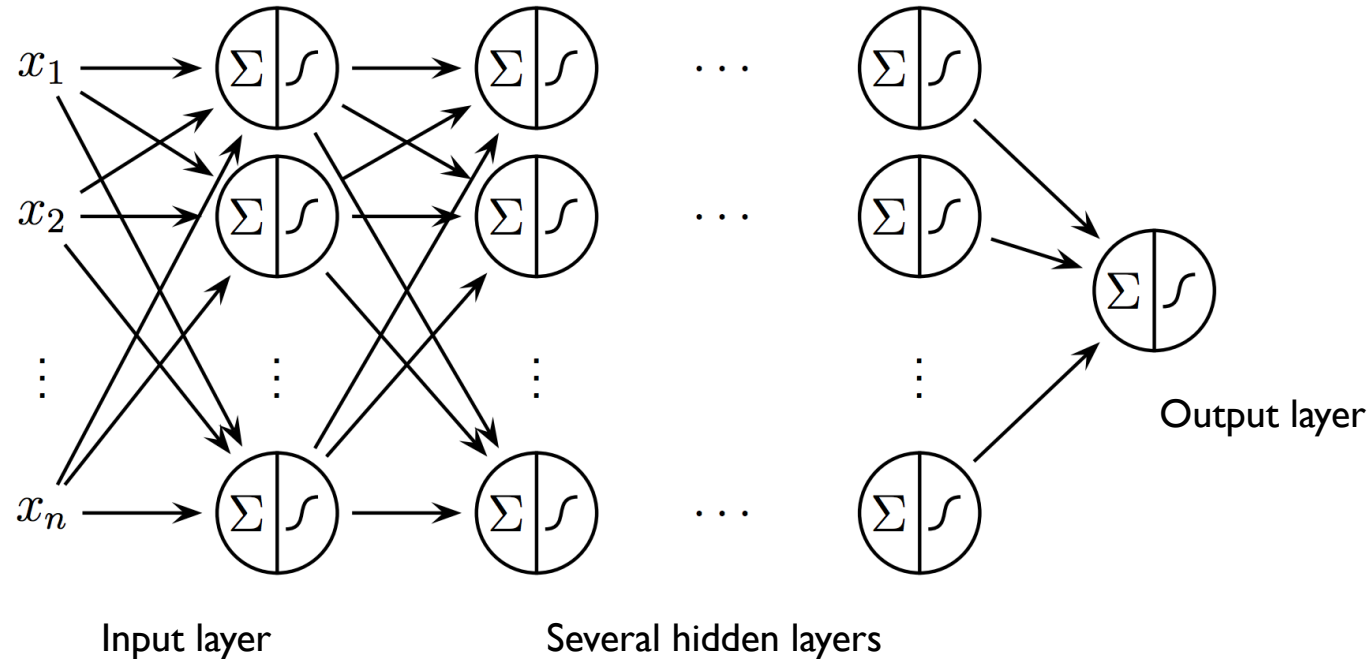
f_{log} is called **logistic function**

Artificial Neural Networks are
also known as multi layer perceptrons

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



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



But how does it learn?!

Back-Propagation

As we train the model we update the sigmoid function weights in order to get the best predictions possible.

This update is generally done via Stochastic Gradient Descent but has other options (batch gradient descent for example)

Pros

- Online model (updates as you go)
 - Doesn't need to be fit all of the time
- Very fast predictions
- Can approximate almost any type of function
- Can be used in a supervised and unsupervised manner
- Super cool

Cons

- Requires many training samples to be considered good
- Hard to describe what is happening
- Requires a lot of hardware / computation power
- Slow to train
- Sklearn only has unsupervised version
- Other versions are difficult to use

The most advanced ANN's use thousand's of neurons which is a lot right?

The most advanced ANN's use thousand's of neurons which is a lot right?

Sure but my dog has billions.....



<http://deepdreamgenerator.com/>

Google uses a supervised neural network to recognize content in photos.

<http://deepdreamgenerator.com/>

Turns out if you input an image you can ask the neural network to try and “re-create” the image as well

So what is Tensorflow?

Tensorflow is an open source machine learning module out of google labs

<https://github.com/tensorflow/tensorflow>

It is their version of a scikit-learn

It can perform not only deep learning techniques, but other more simpler machine learning functions (like linear and logistic regression)

We will be utilizing a a training module of tensorflow and seeing how they use SGD in order to optimize a neural network