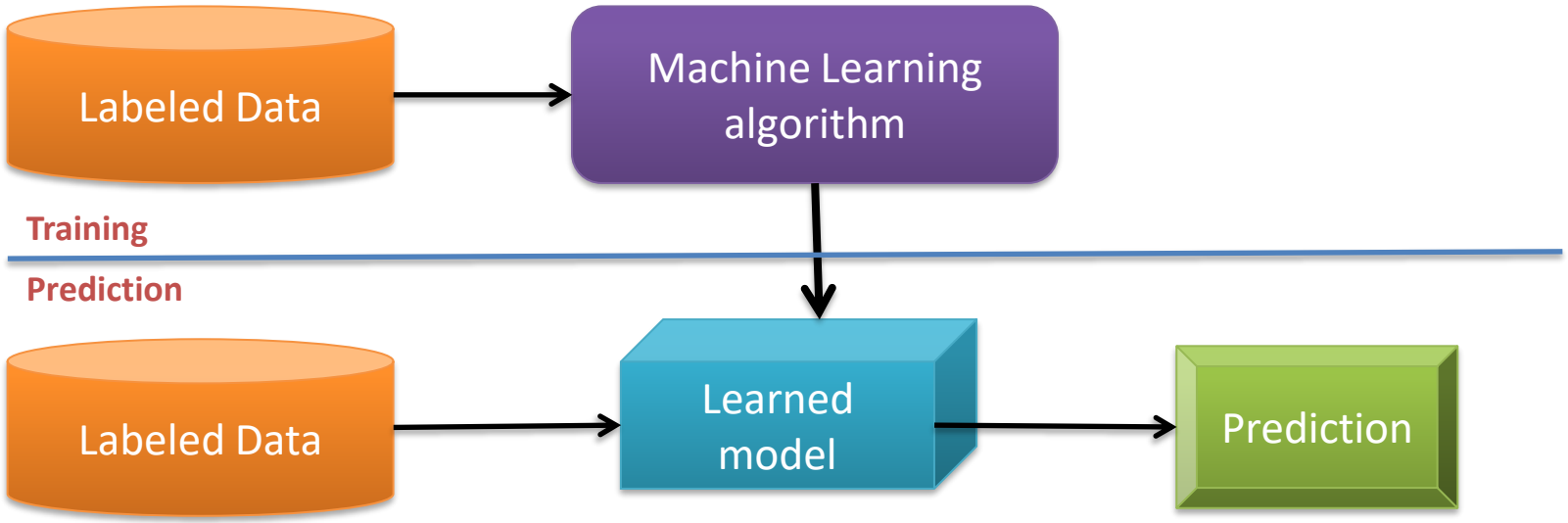
Deep Learning

Machine Learning Basics

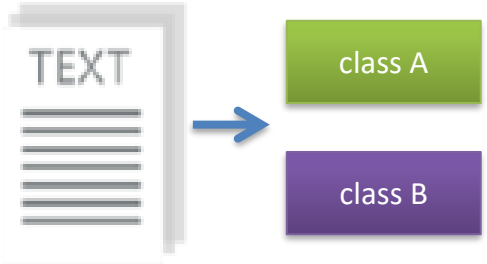
- ***Artificial Intelligence*** is a scientific field concerned with the development of algorithms that allow computers to learn without being explicitly programmed
- ***Machine Learning*** is a branch of Artificial Intelligence, which focuses on methods that learn from data and make predictions on unseen data



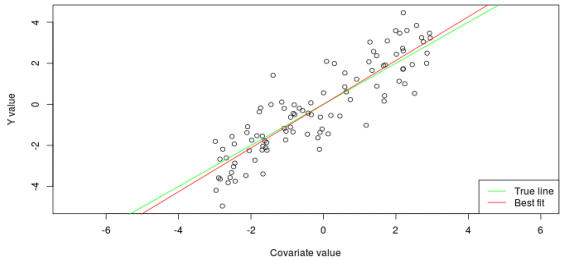
Picture from: Ismini Lourentzou – Introduction to Deep Learning

Machine Learning Types

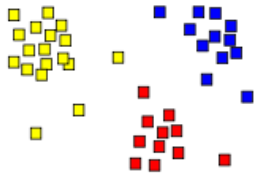
- ***Supervised***: learning with labeled data
  - Example: email classification, image classification
  - Example: regression for predicting real-valued outputs
- ***Unsupervised***: discover patterns in unlabeled data
  - Example: cluster similar data points
- ***Reinforcement learning***: learn to act based on feedback/reward
  - Example: learn to play Go



Classification



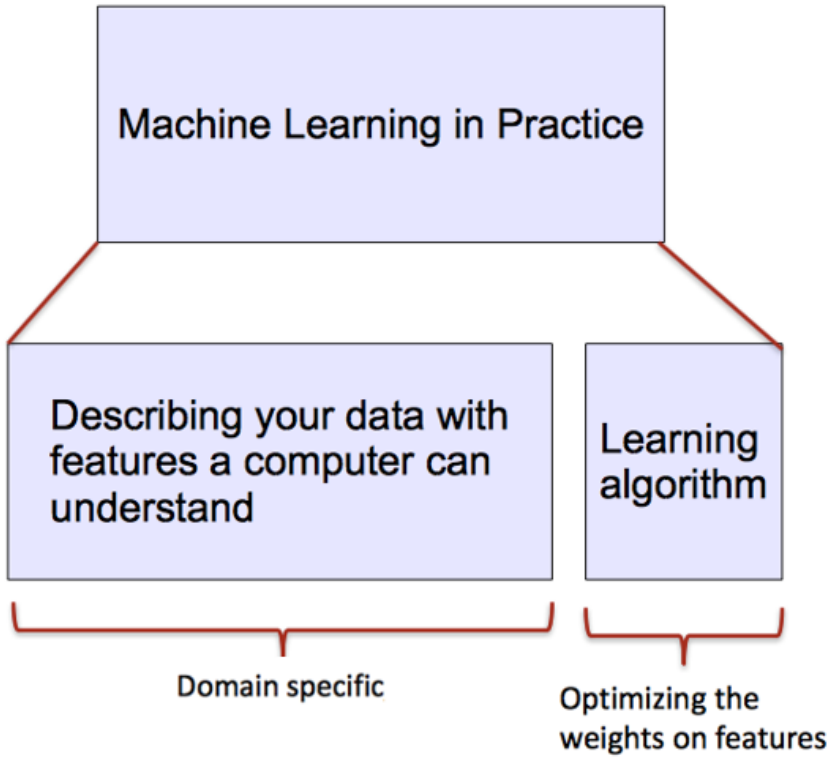
Regression



Clustering

ML vs. Deep Learning

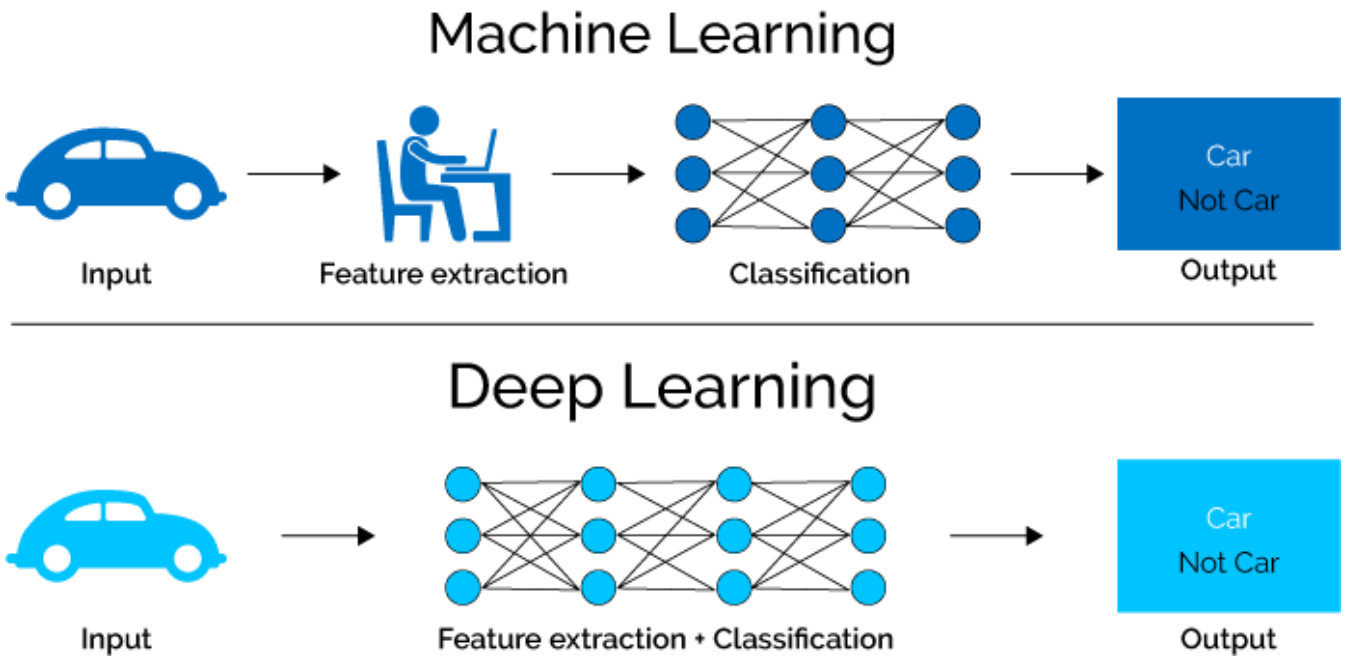
- Conventional machine learning methods rely on human-designed feature representations
  - ML becomes just optimizing weights to best make a final prediction



Picture from: Ismini Lourentzou – Introduction to Deep Learning

ML vs. Deep Learning

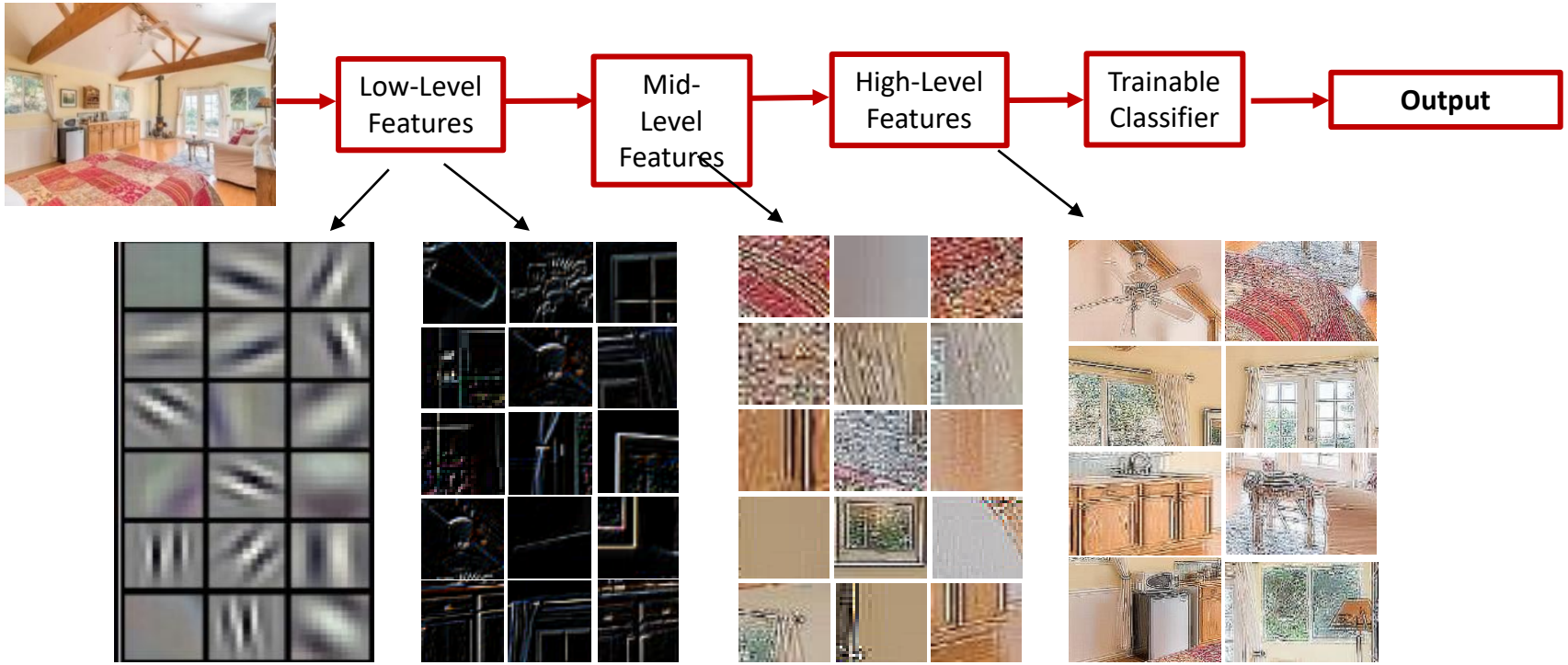
- ***Deep learning*** (DL) is a machine learning subfield that uses multiple layers for learning data representations
  - DL is exceptionally effective at learning patterns



Picture from: https://www.xenonstack.com/blog/static/public/uploads/media/machine-learning-vs-deep-learning.png

ML vs. Deep Learning

- DL applies a multi-layer process for learning rich hierarchical features (i.e., data representations)
  - Input image pixels → Edges → Textures → Parts → Objects



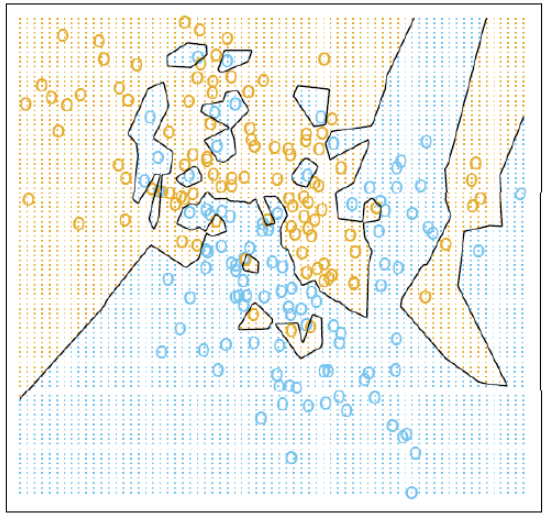
Slide credit: Param Vir Singh – Deep Learning

Why is DL Useful?

- DL provides a flexible, learnable framework for representing visual, text, linguistic information
  - Can learn in supervised and unsupervised manner
- DL represents an effective end-to-end learning system
- Requires large amounts of training data
- Since about 2010, DL has outperformed other ML techniques
  - First in vision and speech, then NLP, and other applications

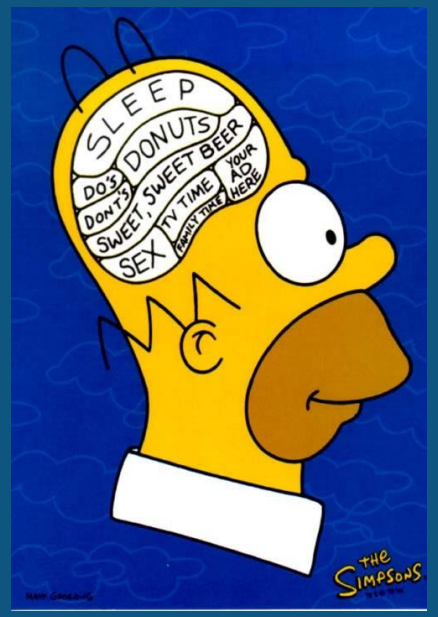
Representational Power

- NNs with at least one hidden layer are universal approximators
  - Given any continuous function $h(x)$ and some $\epsilon > 0$, there exists a NN with one hidden layer (and with a reasonable choice of non-linearity) described with the function $f(x)$, such that $\forall x, |h(x) - f(x)| < \epsilon$
  - I.e., NN can approximate any arbitrary complex continuous function
- NNs use nonlinear mapping of the inputs x to the outputs $f(x)$ to compute complex decision boundaries
- But then, why use deeper NNs?
  - The fact that deep NNs work better is an empirical observation
  - Mathematically, deep NNs have the same representational power as a one-layer NN



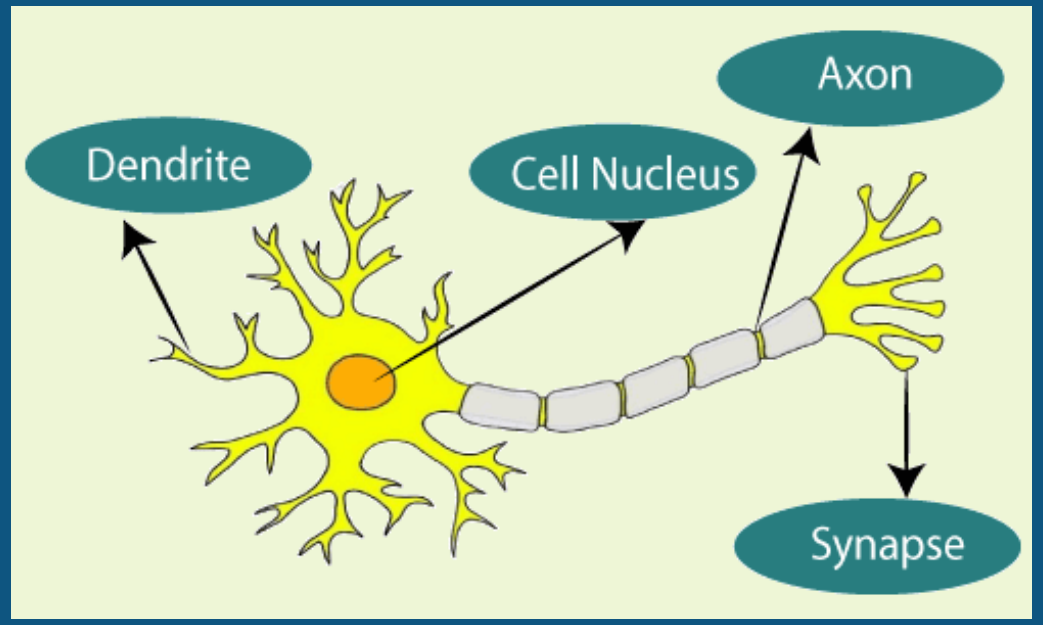
Artificial Neural Networks

Biological Motivation



- Can we simulate the human learning process?
- Biological learning system (brain)
  - complex network of neurons
  - ANN are loosely motivated by biological neural systems. However, many features of ANNs are inconsistent with biological systems

How Does our Brain Work?



Do you know which way the signal flows?

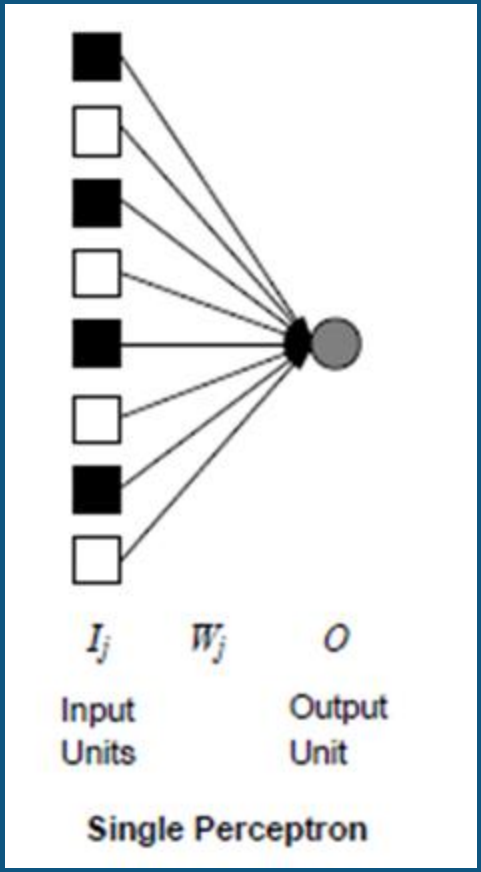
- A neuron is connected to other neurons via its input and output links
  - Dendrites receive the signal (input) and Axons transmit signal (output)
  - The neuron sums the incoming weighted values and this value is input to an activation function
  - The output of the activation function is the output from the neuron

Neural Network Learning

- Neural Network: Learning approach based on modeling adaptation in biological neural systems
- Perceptron: Initial algorithm for learning simple neural networks (single layer) developed in the 1950's.
- Multi-Layer ANNs: More complex algorithm for learning multi-layer neural networks developed in the 1980's.

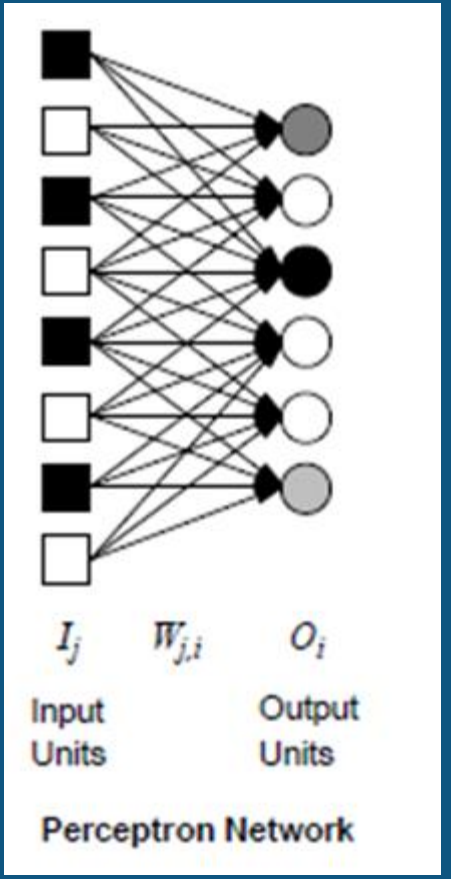
Perceptron

Perceptron is type of ANN that can be seen as the simplest kind of feedforward neural network: a linear classifier, introduced in the late 1950s.



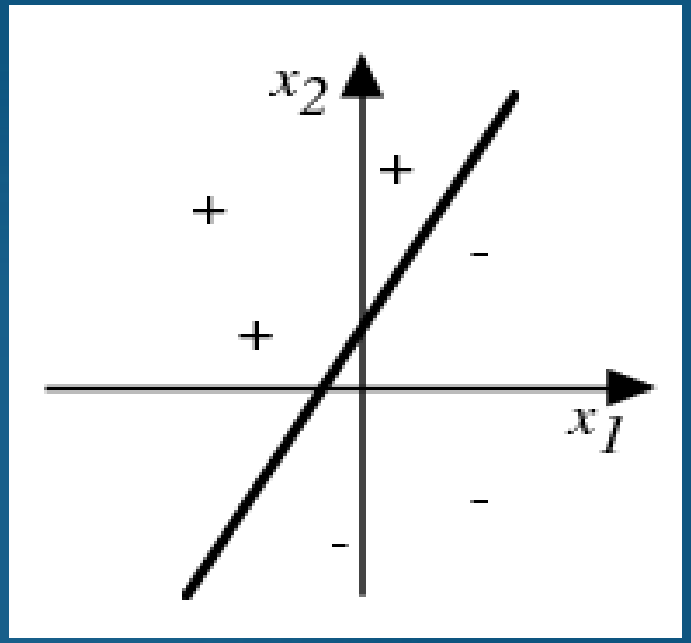
Perceptron network:

- single-layer
- feed-forward: data only travels in one direction

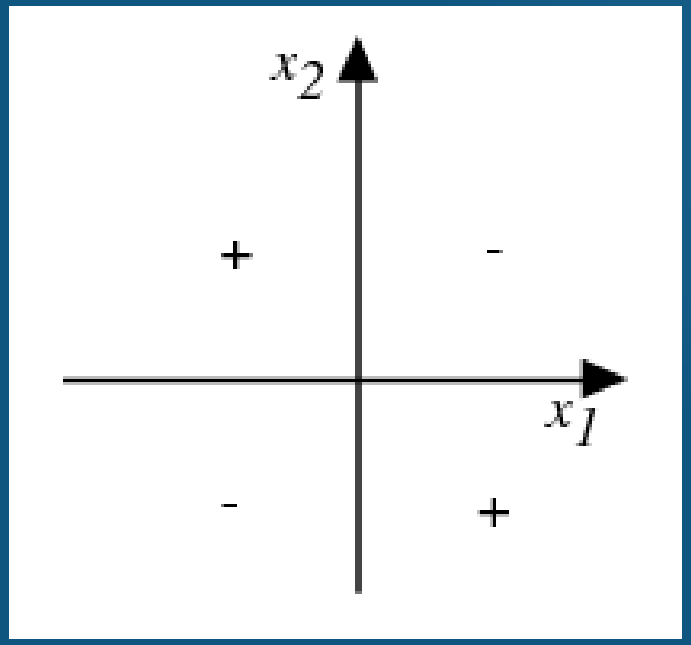


Perceptron

Perceptron convergence theorem (Rosenblatt 1962): Perceptron will learn to classify any linearly separable set of inputs.

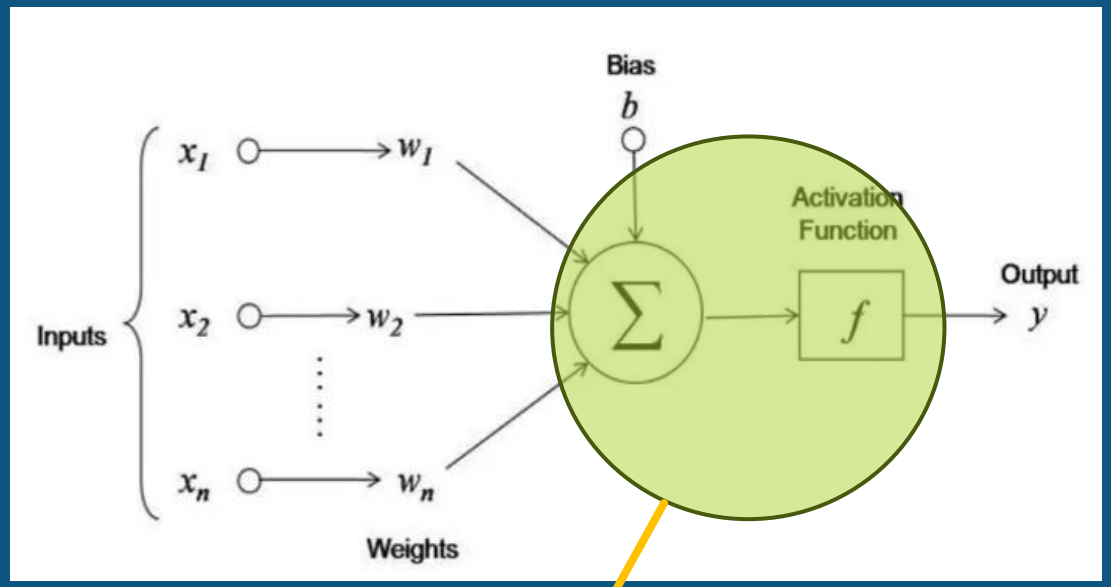


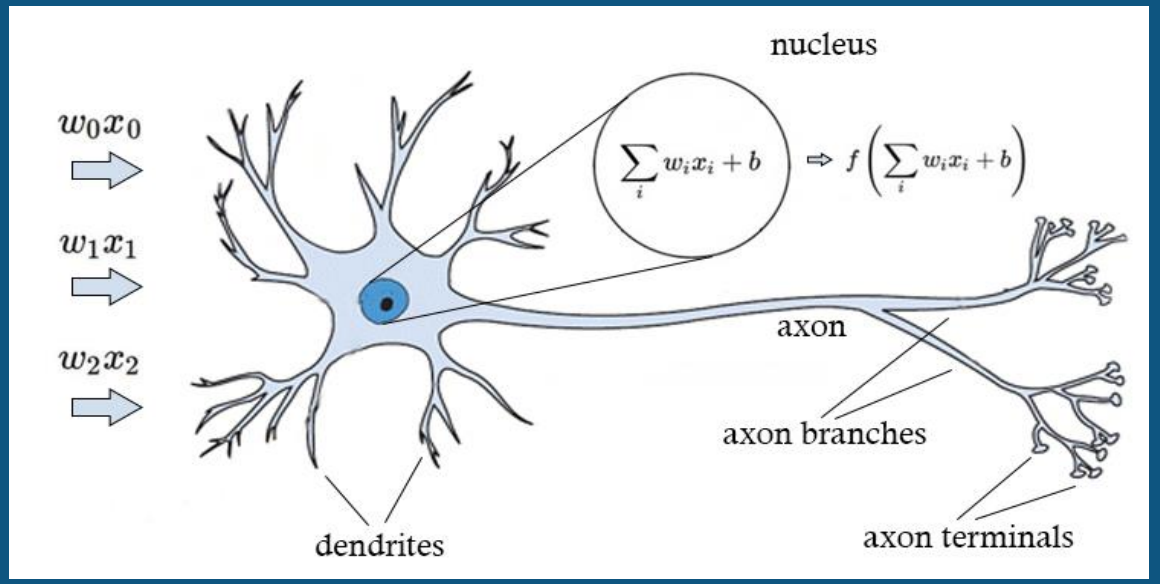
Linear Separation

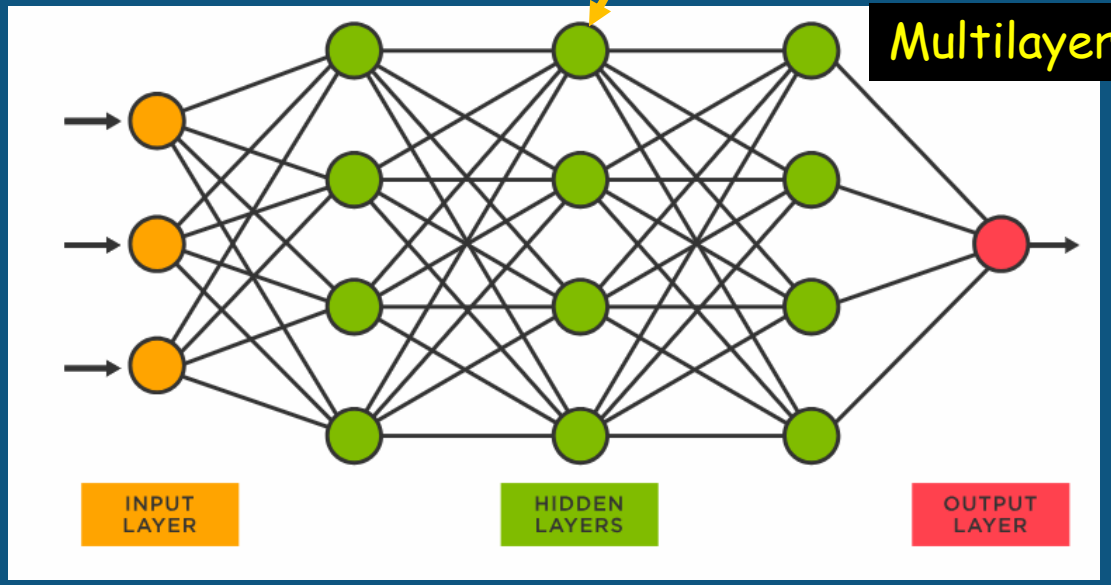


XOR function (no linear separation)

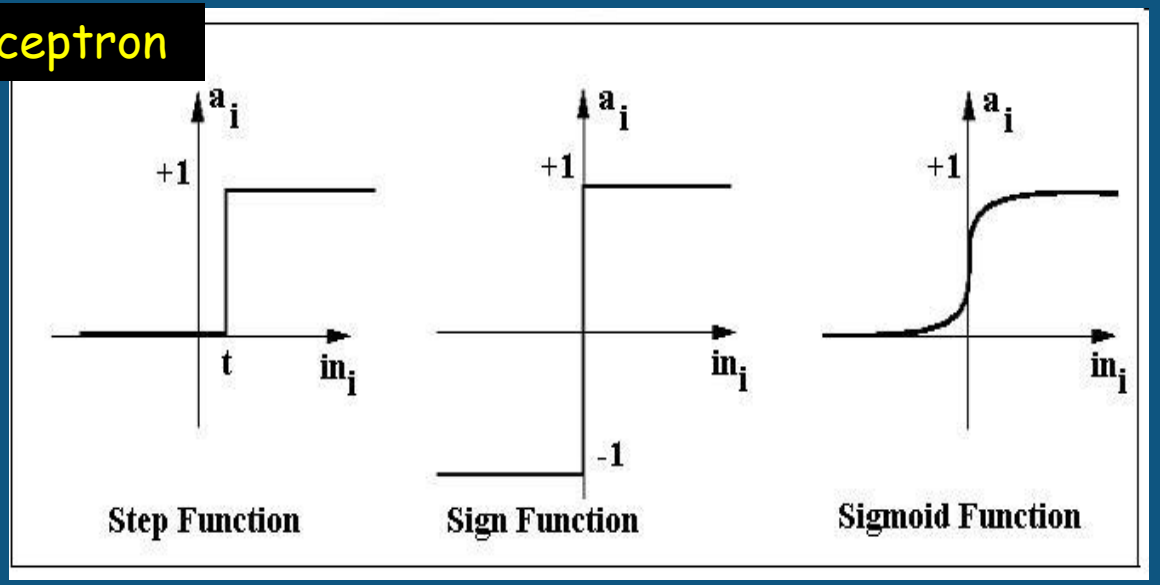
Perceptron



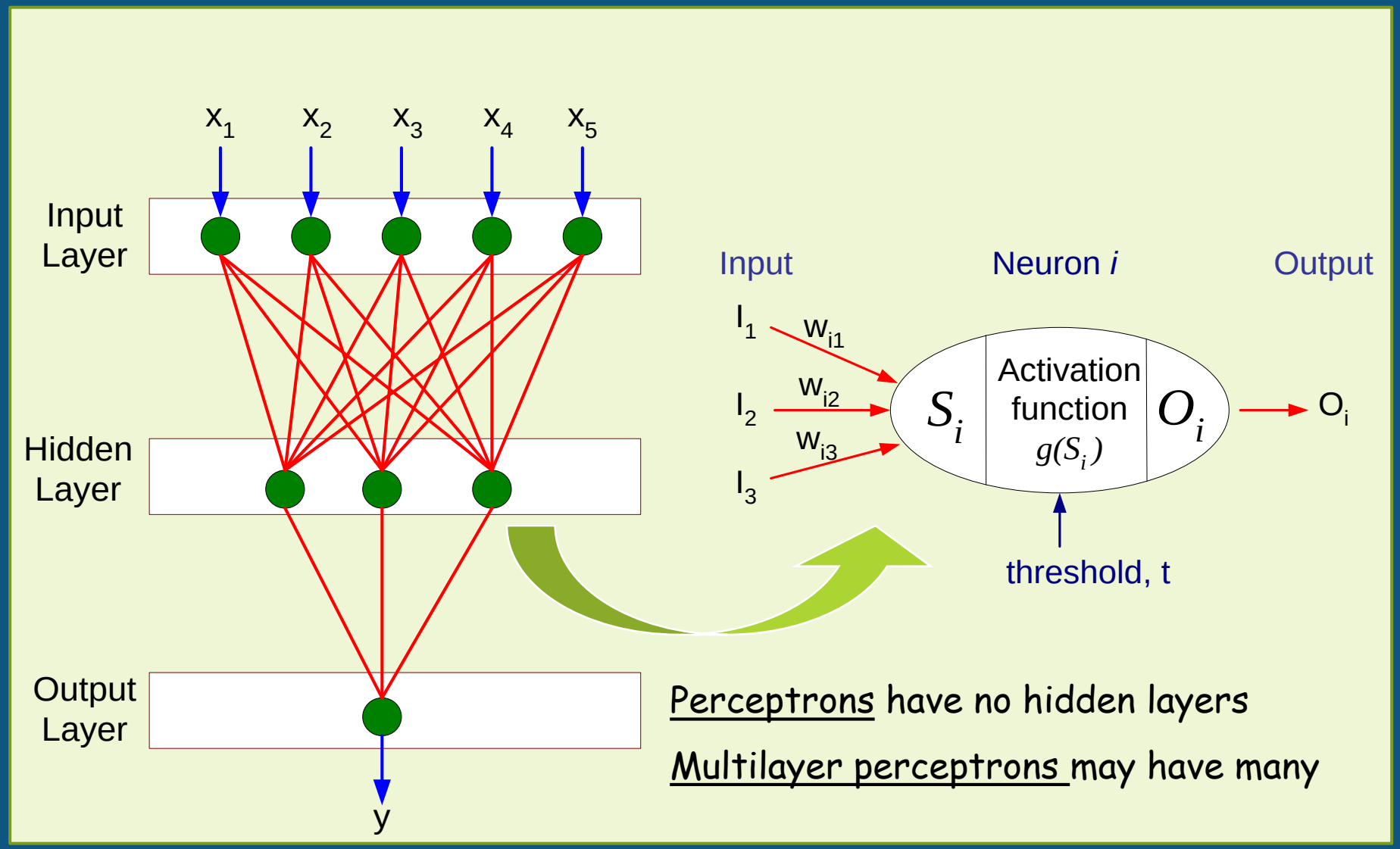




Multilayer perceptron



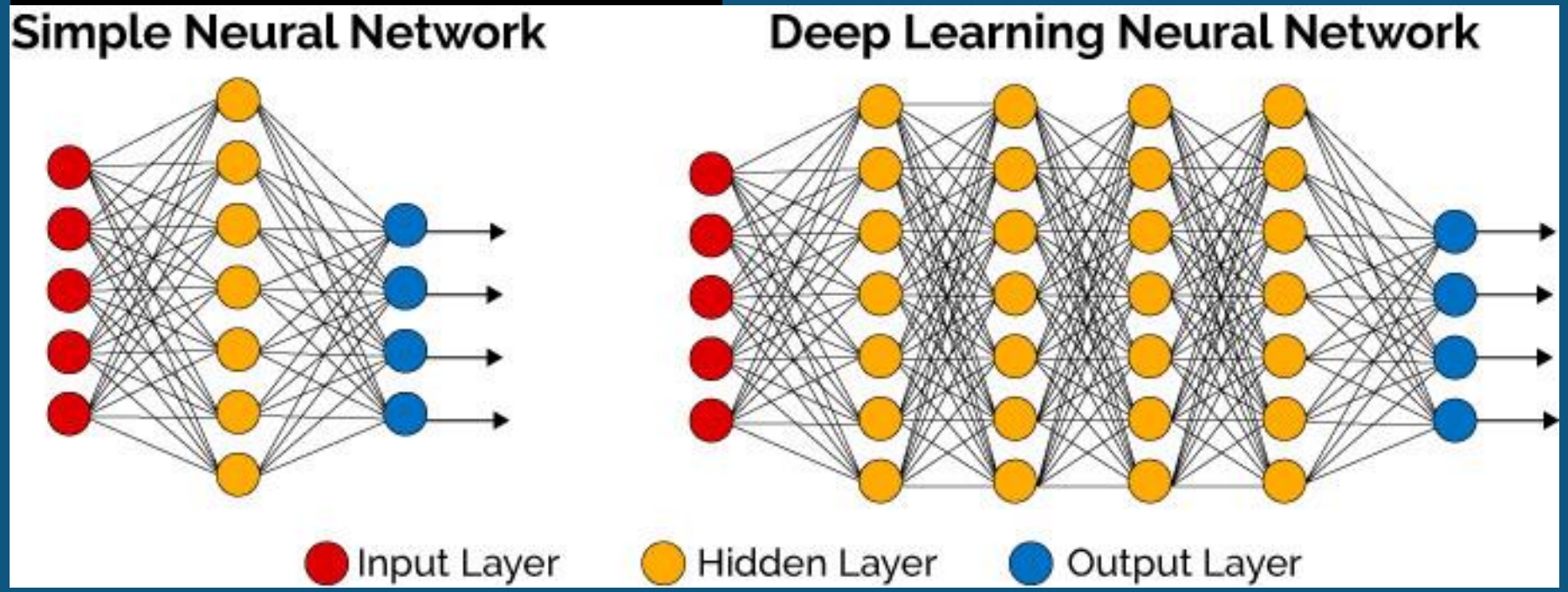
General Structure of an ANN

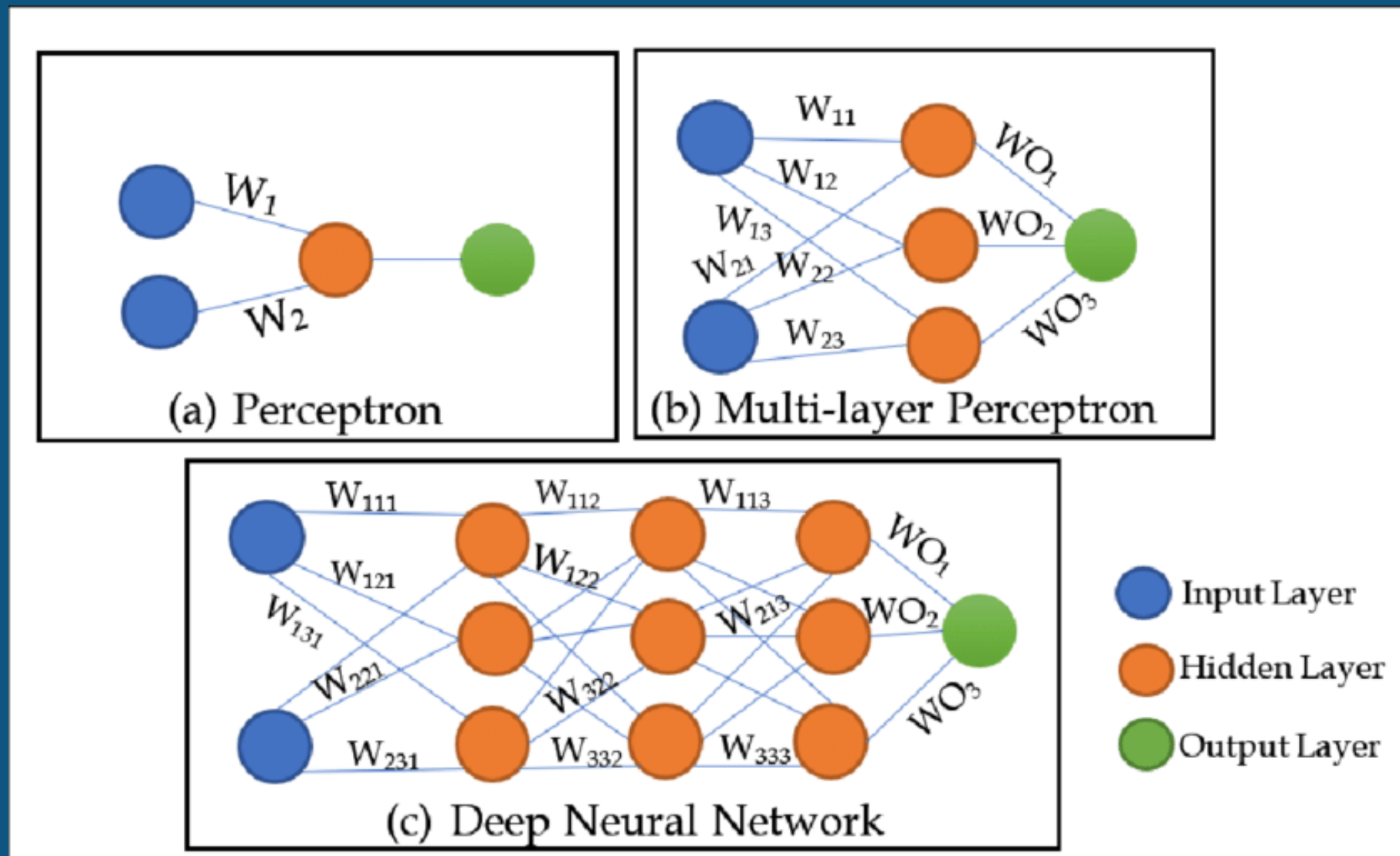


Perceptrons have no hidden layers

Multilayer perceptrons may have many

Artificial Neural Network (ANN)
Vanila Neural Network (VNN)



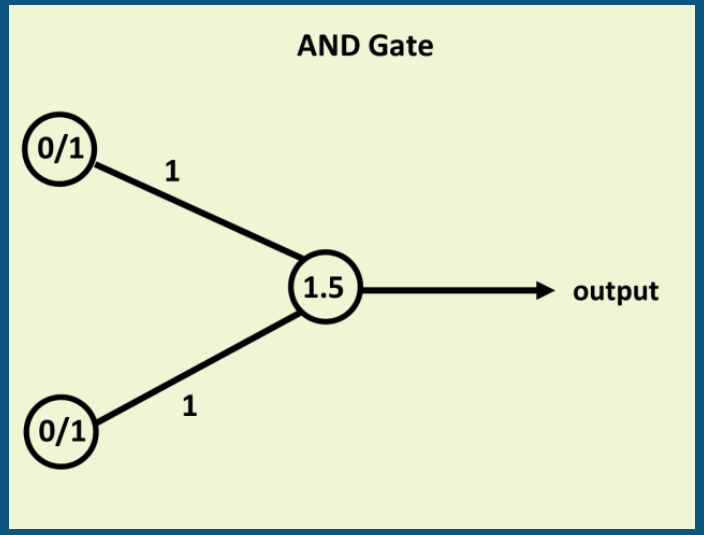

W1
W2
(a) Perceptron
W11
W12
W13
W21
W22
W23
WO1
WO2
WO3
(b) Multi-layer Perceptron
W111
W112
W113
W121
W122
W131
W213
W221
W322
W231
W332
W333
WO1
WO2
WO3
(c) Deep Neural Network
Input Layer
Hidden Layer
Output Layer

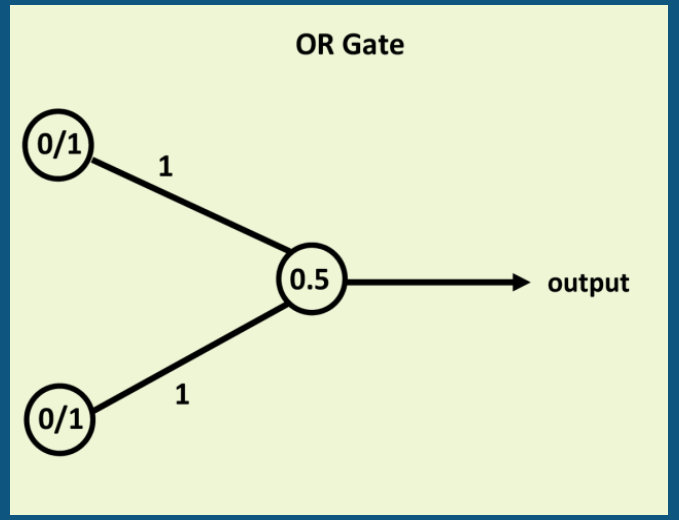
Simple Examples from Boolean Algebra

What logic functions do these table represent?

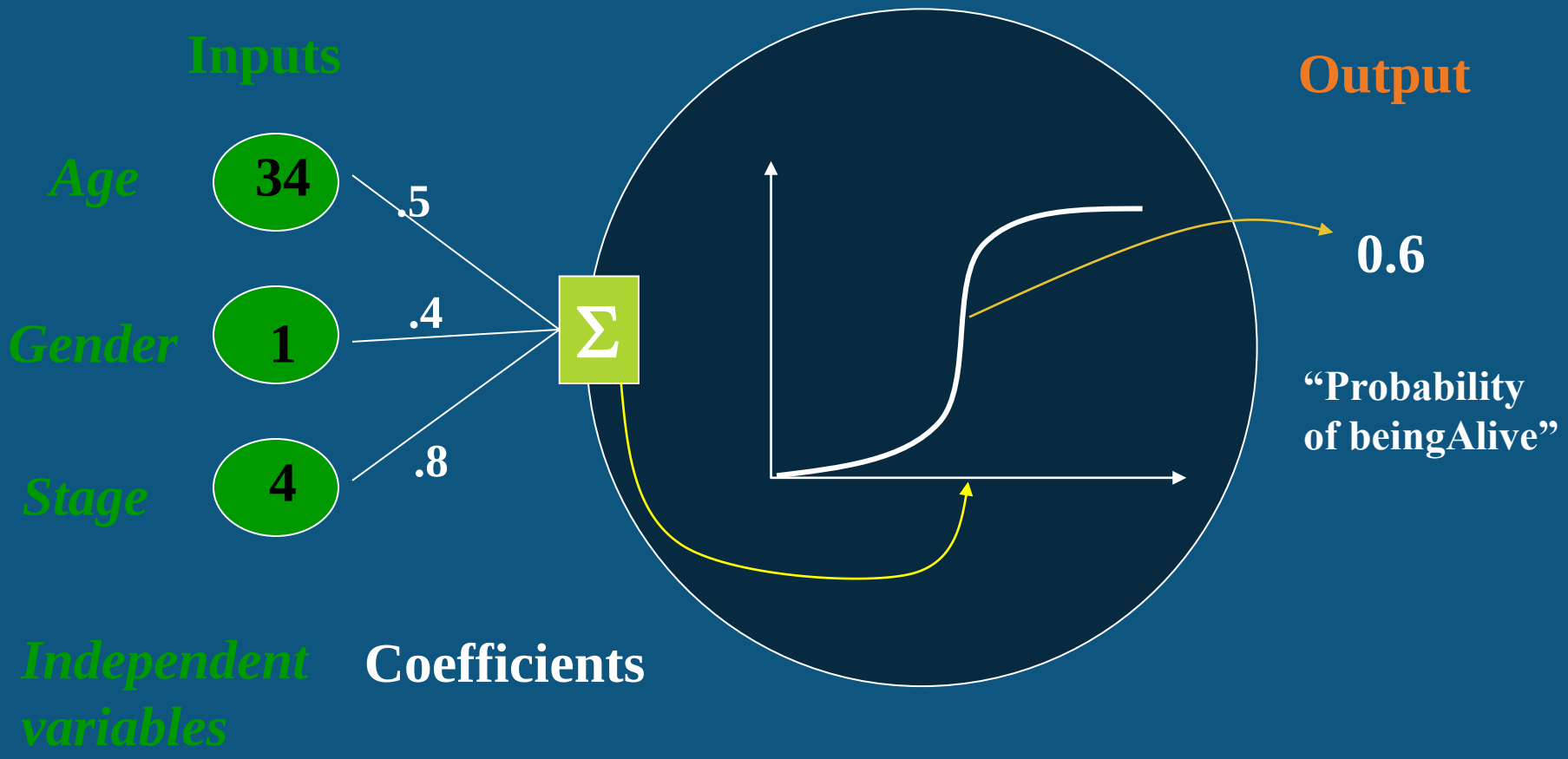
In_1	In_2	Out
0	0	0
0	1	0
1	0	0
1	1	1

In_1	In_2	Out
0	0	0
0	1	1
1	0	1
1	1	1





Logistic function



Simple Examples from Boolean Algebra

What logic functions do these table represent?

In_1	In_2	Out
0	0	0
0	1	1
1	0	1
1	1	0

XOR GATE

BOOLEAN EXPRESSION

$$A \cdot \overline{B} + \overline{A} \cdot B$$

$$(A + B) \cdot (\overline{A} + \overline{B})$$

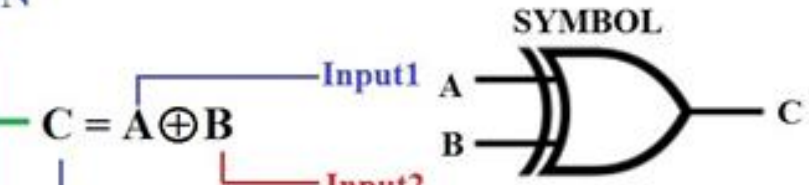
C = A⊕B

Input1

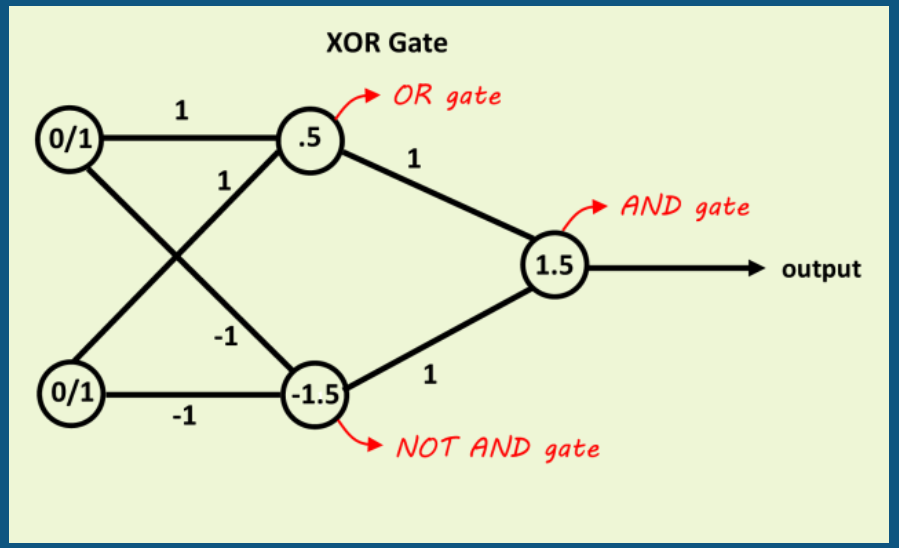
Input2

Output

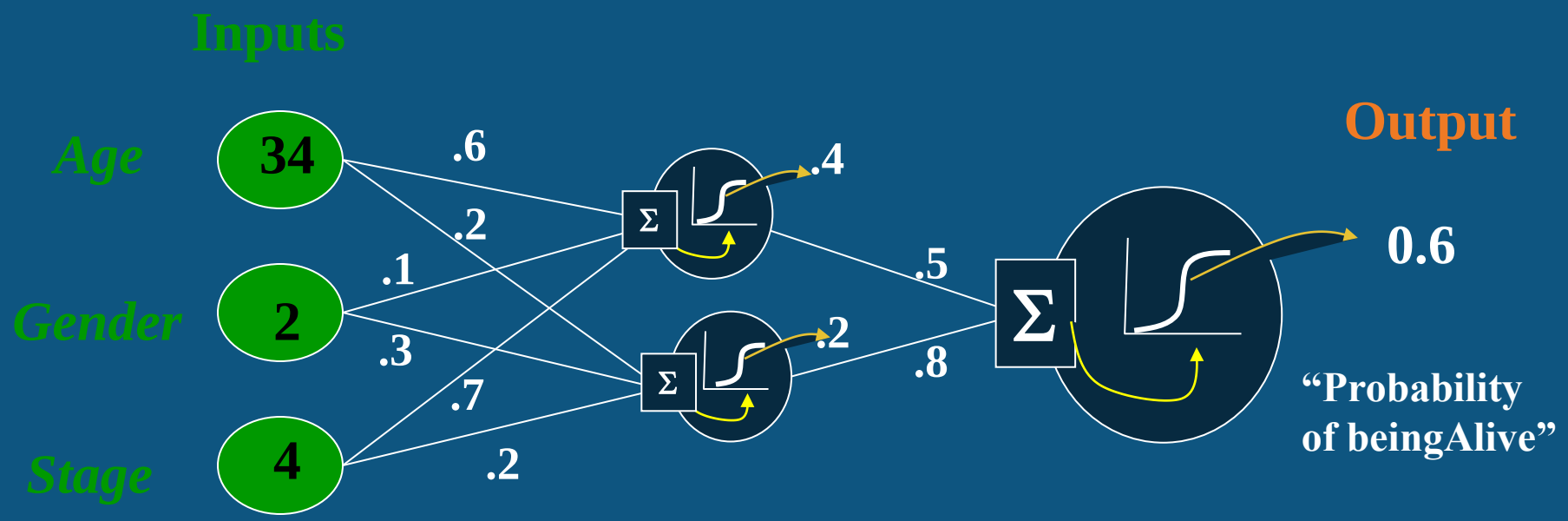
SYMBOL



INPUT		OUTPUT
A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0



Neural Network Model



Independent variables

Weights

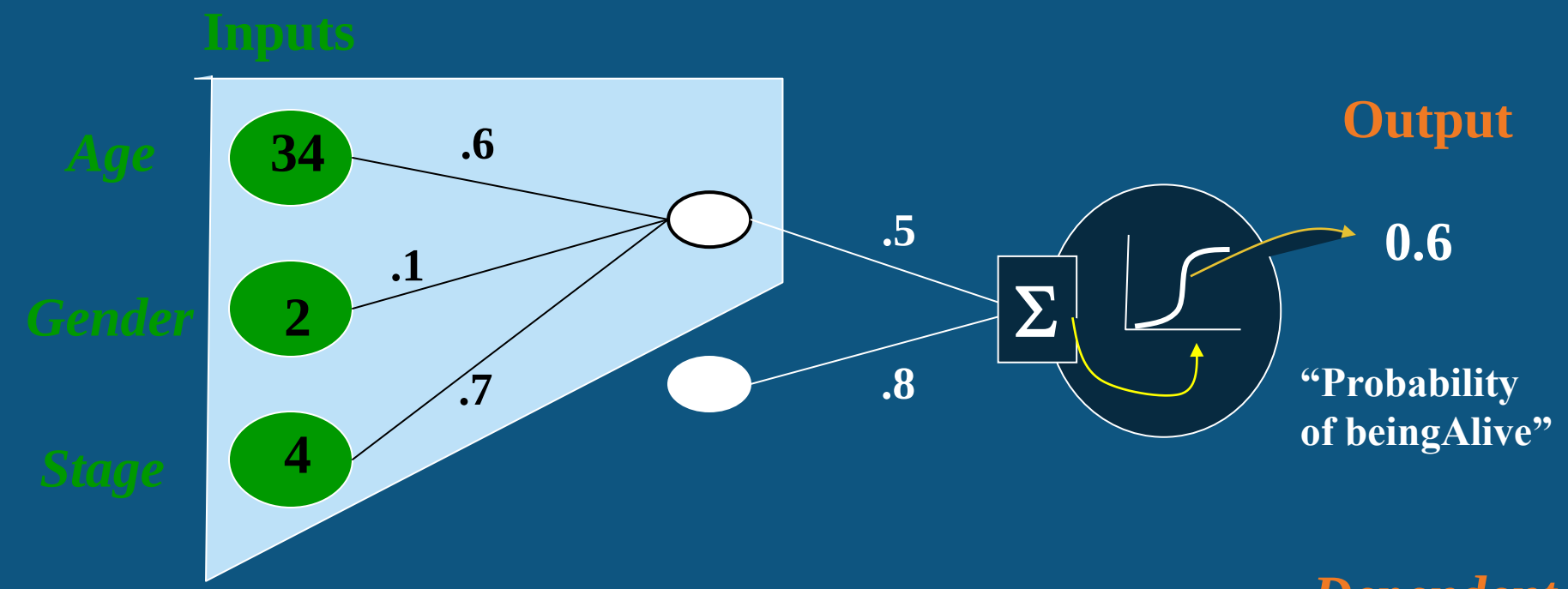
Hidden Layer

Weights

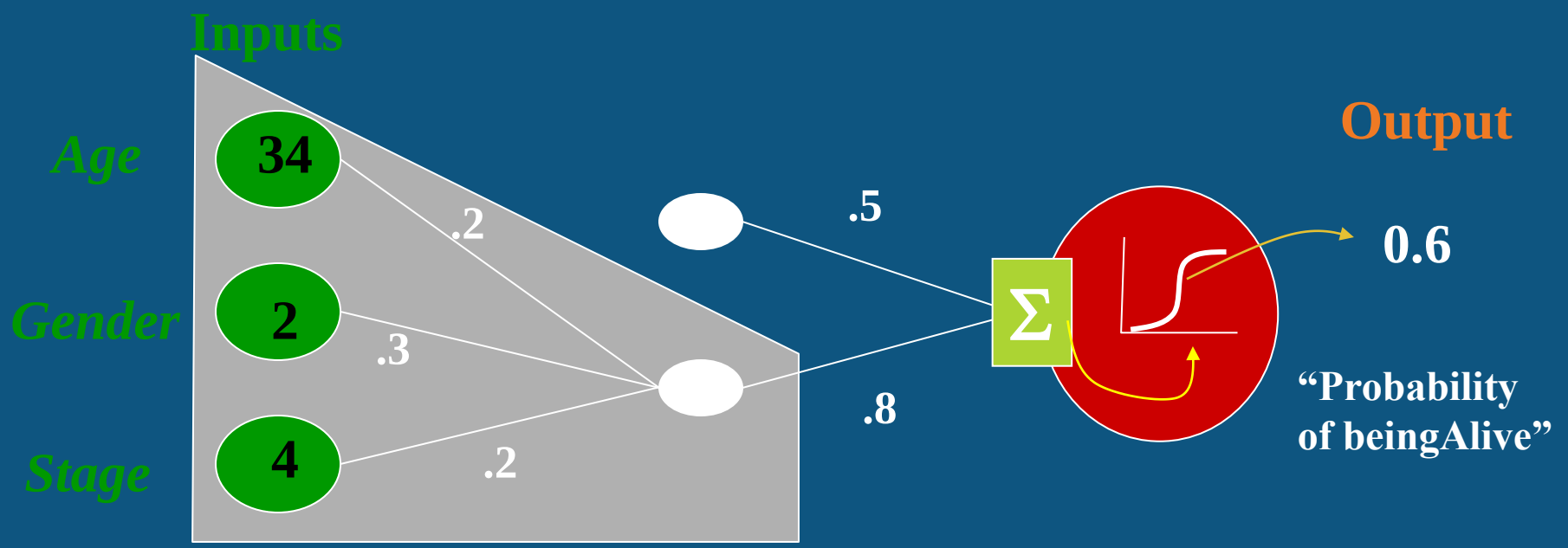
Dependent variable

Prediction

Getting an answer from a NN



Getting an answer from a NN



Independent variables

Weights

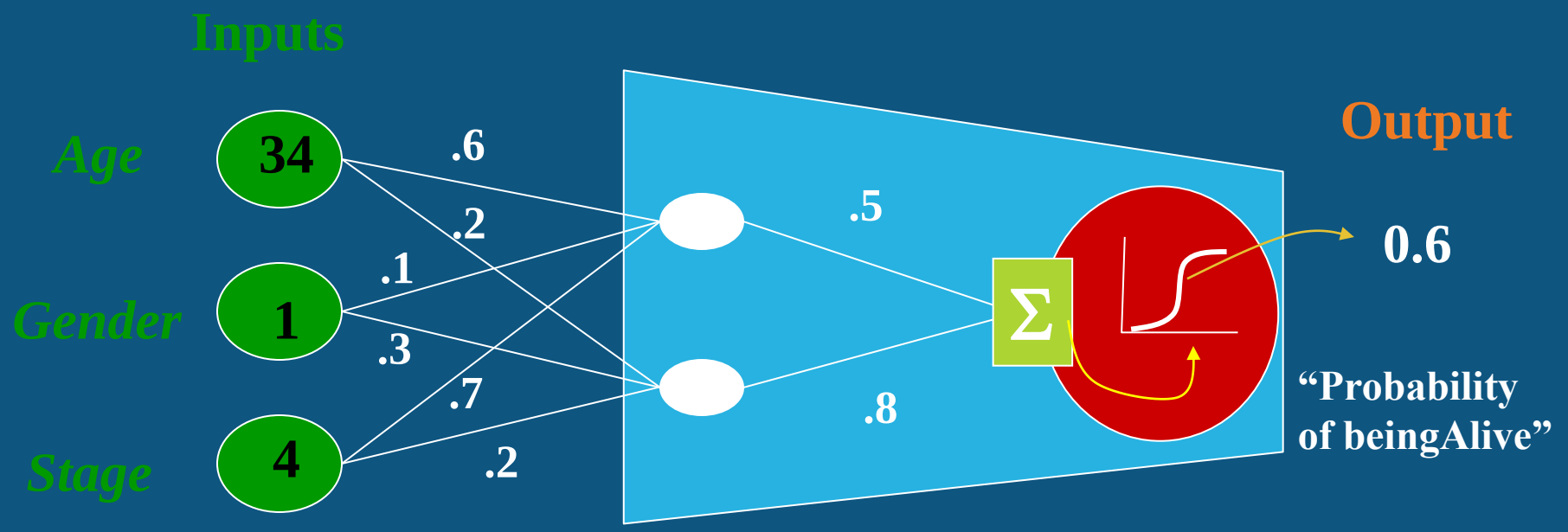
Hidden Layer

Weights

Dependent variable

Prediction

Getting an answer from a NN



Independent variables

Weights

Hidden Layer

Weights

Dependent variable

Prediction

Thank You!

References:

- Slide (Partial) Courtesy: University of Idaho
- Hung-yi Lee – Deep Learning Tutorial
- Ismini Lourentzou – Introduction to Deep Learning
- CS231n Convolutional Neural Networks for Visual Recognition (Stanford CS course) (link)
- James Hays, Brown – Machine Learning Overview
- Param Vir Singh, Shunyuan Zhang, Nikhil Malik – Deep Learning