Learning Neural Networks - Part 1

Alice Gao

Lecture 8

Readings: RN 18.7, PM 7.5.

Outline

Learning Goals

By the end of the lecture, you should be able to

- Describe motivations for using a neural network model.
- Describe the simple mathematical model of a neuron.
- Describe desirable properties of an activation function.
  Give examples of activation functions and their properties.
- Distinguish feed-forward and recurrent neural networks.
- Learn a perceptron that represents a simple logical function.
- Determine the logical function represented by a perceptron.
- Explain why a perceptron cannot represent the XOR function.
- Construct a 3-layer neural network that represents the XOR function.

Learning Goals

Introduction to Artificial Neural Networks

Introduction to Perceptrons

Limitations of Perceptrons

Revisiting the Learning goals

History and Background

- Artificial Intelligence

- Machine Learning

- Deep Learning

  - ImageNet

  - The Cat Experiment

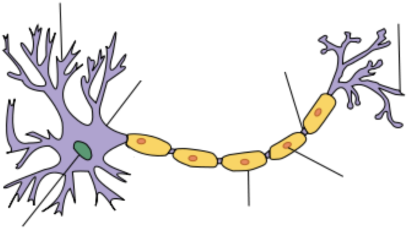
Learning complex relationships

- Image interpretation, speech recognition, and translation.
- The relationship between inputs and outputs can be extremely complex.
- How can we build a model to learn such complex relationships?

Humans can learn complex relationships well.
Can we build a model that mimics the human brain?

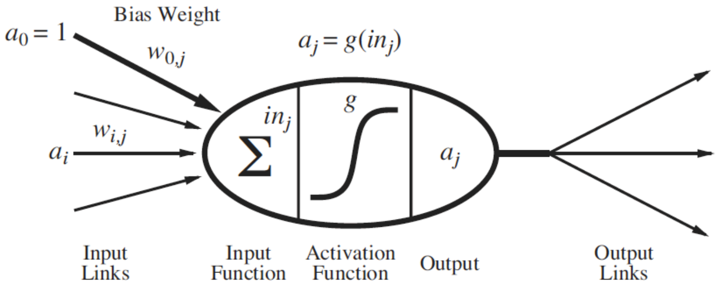
Human brains

- A brain is a set of densely connected neurons.
- Components of a neuron: dendrites, soma, axon, synapse
- Depending on the input signals, the neuron performs computations and decides to fire or not.

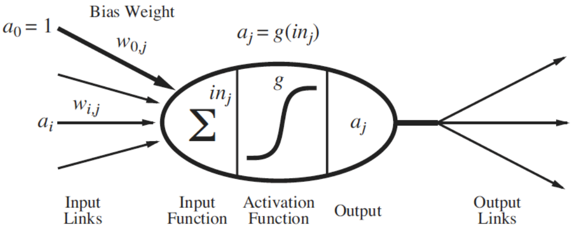


A simple mathematical model of a neuron

- McCulloch and Pitts 1943.
- A linear classfier — it “fires” when a linear combination of its inputs exceeds some threshold.



A simple mathematical model of a neuron



- Neuron j computes a weighted sum of its input signals. $in_j = \sum_{i=0}^{n} w_{ij} a_i$.
- Neuron j applies an activation function g to the weighted sum to derive the output. $a_j = g(in_j) = g\left(\sum_{i=0}^{n} w_{i,j} a_i\right)$.

Desirable Properties of The Activation Function

What are some desirable properties of the activation function?

- It should be non-linear.
- It should mimic the behaviour of real neurons.
- It should be differentiable almost everywhere.

Common activation functions

- Step function: $g(x) = 1$ if $x > 0$. $g(x) = 0$ if $x \leq 0$.

- Sigmoid function: $g(x) = \dfrac{1}{1 + e^{-kx}}$.

Common activation functions (continued)

- Rectified linear unit (ReLu): $g(x) = max(0, x)$.

- Leaky ReLU: $g(x) = (0.1 * x, x)$.

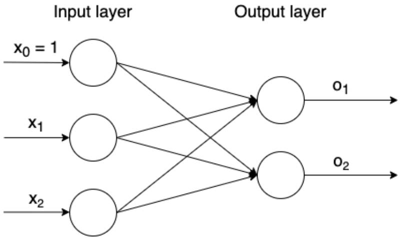
Connecting the neurons together into a network

- Feed-forward network

- Recurrent network

Perceptrons

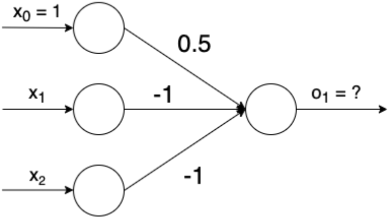
- Single-layer feed-forward neural network
- The inputs are connected directly to the outputs.
- Can represent logical functions, e.g. AND, OR, and NOT.



CQ: What does the perceptron compute?

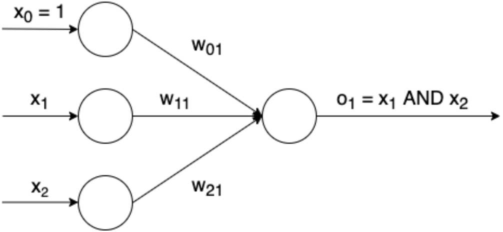
CQ: Consider the following perceptron, where the activation function is the step function. ($g(x) = 1$ if $x > 0$. $g(x) = 0$ if $x \leq 0$.). Which of the following logical function does the perceptron compute?

(A) $x_1 \land x_2$
(B) $\neg(x_1 \land x_2)$
(C) $x_1 \lor x_2$
(D) $\neg(x_1 \lor x_2)$



CQ: Learning a perceptron for the AND function

CQ: Consider the perceptron below where the activation function is the step function ($g(x) = 1$ if $x > 0$. $g(x) = 0$ if $x \leq 0$.). What should the weights w_{01}, w_{11} and w_{21} be such that the perceptron represents an AND function?



x_1	x_2	o_1
0	0	0
0	1	0
1	0	0
1	1	1

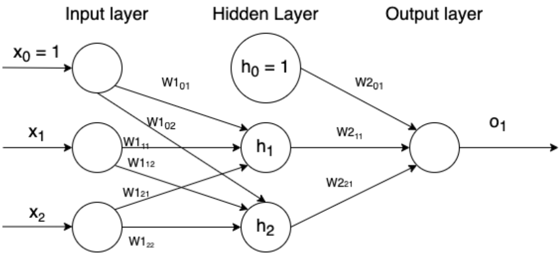
Limitations of perceptrons

- Perceptrons: An introduction to computational geometry. Minsky and Papert. MIT Press. Cambridge MA 1969.
- Results:
  - XOR cannot be represented using perceptrons. We need a deeper network.
  - No one knew how to train deeper networks.
- Led to the first AI winter.

CQ: Why can't a perceptron represent XOR?

XOR as a 2-Layer Neural Network

Can you come up with the weights such that
the following network represents the XOR function?



Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Describe motivations for using a neural network model.
- Describe the simple mathematical model of a neuron.
- Describe desirable properties of an activation function. Give examples of activation functions and their properties.
- Distinguish feed-forward and recurrent neural networks.
- Learn a perceptron that represents a simple logical function.
- Determine the logical function represented by a perceptron.
- Explain why a perceptron cannot represent the XOR function.
- Construct a 3-layer neural network that represents the XOR function.