Artificial Intelligence Foundation – JC3001

Lecture 41: Neural Networks -I

Prof. Aladdin Ayesh (aladdin.ayesh@abdn.ac.uk)

Dr. Binod Bhattarai (binod.bhattarai@abdn.ac.uk)

Dr. Gideon Ogunniye, (g.ogunniye@abdn.ac.uk)

November 2025

Material adapted from:

Russell and Norvig (AIMA Book): Chapter 21

Andrew Ng (Stanford University / Coursera)

Course Progression

- Part 1: Introduction
  1. Introduction to AI ✓
  2. Agents ✓
- Part 2: Problem-solving
  1. Search 1: Uninformed Search ✓
  2. Search 2: Heuristic Search ✓
  3. Search 3: Local Search ✓
  4. Search 4: Adversarial Search ✓
- Part 3: Reasoning and Uncertainty
  1. Reasoning 1: Constraint Satisfaction ✓
  2. Reasoning 2: Logic and Inference ✓
  3. Probabilistic Reasoning 1: BNs ✓
  4. Probabilistic Reasoning 2: HMMs ✓
- Part 4: Planning
  1. Planning 1: Intro and Formalism ✓
  2. Planning 2: Algos and Heuristics ✓
  3. Planning 3: Hierarchical Planning ✓
  4. Planning 4: Stochastic Planning ✓
- Part 5: Learning
  1. Learning 1: Intro to ML ✓
  2. Learning 2: Regression ✓
  3. **Learning 3: Neural Networks**
  4. Learning 4: Reinforcement Learning
- Part 6: Conclusion
  1. Ethical Issues in AI
  2. Conclusions and Discussion

Objectives

- From regression to perceptrons
- Neural Networks
- Deep Learning basics

- Recap
- Artificial Neural Networks
- Perceptron

- Machine Learning is a subfield of artificial intelligence that allows the computer to learn about a specific subject through data examples;
- We can apply it using supervised learning, unsupervised learning, and reinforcement learning;

Classification: $y = 0$ or 1

$h_{\boldsymbol{w}}(x)$ can be > 1 or < 0

Logistic regression: $0 \leq h_{\boldsymbol{w}}(x) \leq 1$

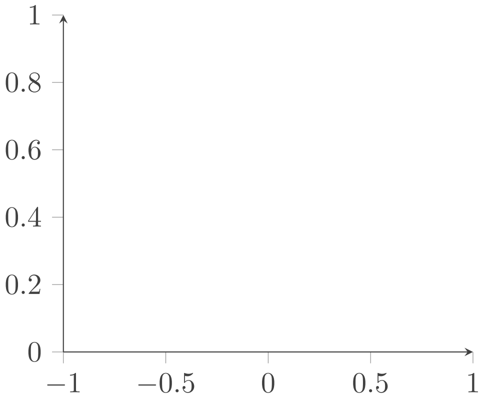
Want $0 \le h_{\boldsymbol{w}}(x) \le 1$

$$h_{\boldsymbol{w}} = g(\boldsymbol{w}^T x)$$

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

$$h_{\boldsymbol{w}} = \frac{1}{1 + e^{-\boldsymbol{w}^T x}}$$

Sigmoid Function
Logistic Function



Want $0 \leq h_{\boldsymbol{w}}(x) \leq 1$

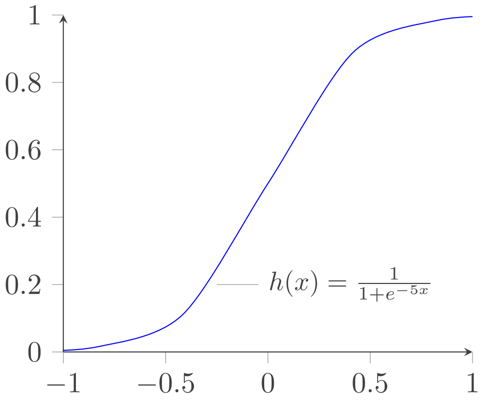
$$h_{\boldsymbol{w}} = g(\boldsymbol{w}^T x)$$

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

$$h_{\boldsymbol{w}} = \frac{1}{1 + e^{-\boldsymbol{w}^T x}}$$

Sigmoid Function

Logistic Function



Logistic Regression – Cost Function

Linear regression: $Loss(\boldsymbol{w}) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{2}\left(h_{\boldsymbol{w}}(x^{(i)}) - y^{i}\right)^{2}$

Logistic regression: $Loss(\boldsymbol{w}) = \frac{1}{m}\sum_{i=1}^{m} Cost(h_{\boldsymbol{w}}(x), y)$

$$Cost(h_{\boldsymbol{w}}(x), y) = \begin{cases} -\log(h_{\boldsymbol{w}}(x)) & \text{if } y = 1 \\ -\log(1 - h_{\boldsymbol{w}}(x)) & \text{if } y = 0 \end{cases}$$

Logistic Regression – Cost Function

$$Loss(\boldsymbol{w}) = \frac{1}{m}\sum_{i=1}^{m} Cost(h_{\boldsymbol{w}}(x), y)$$

$$Cost(h_{\boldsymbol{w}}(x), y) = \begin{cases} -\log(h_{\boldsymbol{w}}(x)) & \text{if } y = 1 \\ -\log(1 - h_{\boldsymbol{w}}(x)) & \text{if } y = 0 \end{cases}$$

Note: $y = 0$ or 1 always

$$Cost(h_{\boldsymbol{w}}(x), y) = -y \log h_{\boldsymbol{w}}(x) - (1 - y)\log(1 - h_{\boldsymbol{w}}(x))$$

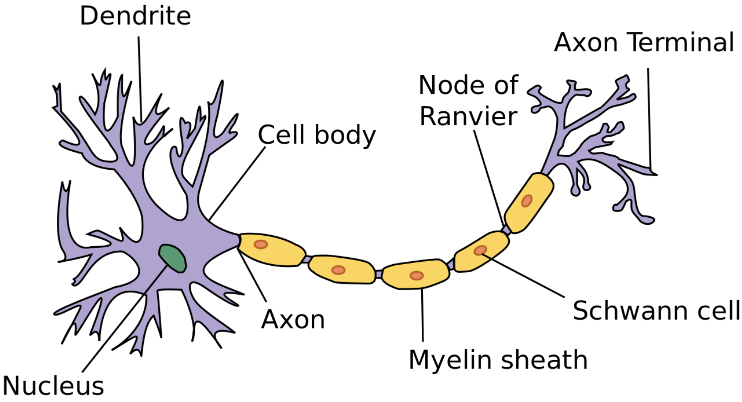
$$Loss(\boldsymbol{w}) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_{\boldsymbol{w}}(x)^{(i)} + (1 - y^{(i)})\log(1 - h_{\boldsymbol{w}}(x)^{(i)})\right]$$

- Recap
- **Artificial Neural Networks**
- Perceptron

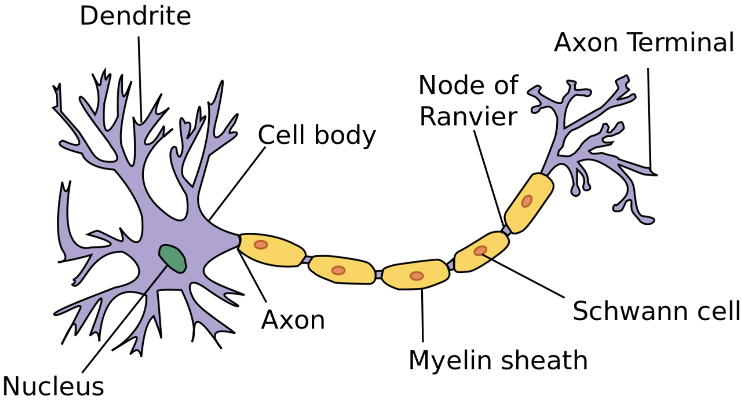
- Supervised learning algorithm that was originally motivated by the goal of having machines that can mimic the brain

- Supervised learning algorithm that was originally motivated by the goal of having machines that can mimic the brain
  But take this analogy lightly
- Used for classification and regression

- Origins: Algorithms that try to mimic the brain
- Widely used in 80s and early 90s; popularity diminished in 90s
- Recent resurgence: State-of-the-art technique for many applications


Dendrite
Axon Terminal
Node of Ranvier
Cell body
Axon
Schwann cell
Myelin sheath
Nucleus

Neurons in the brain

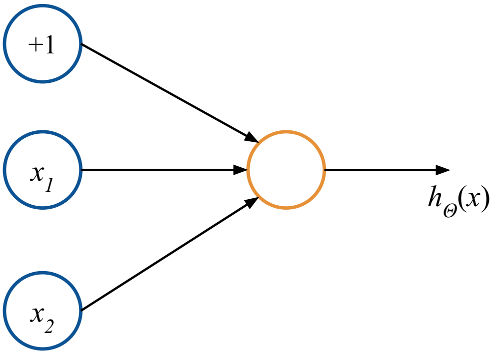


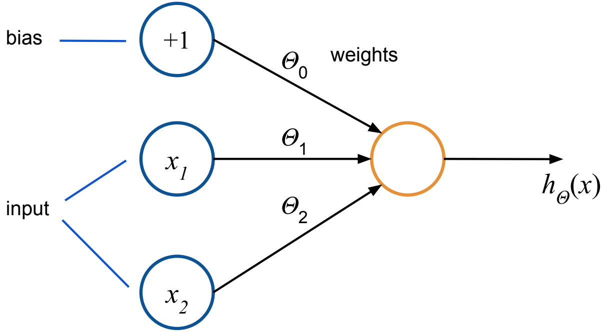
Comparing this with an artificial “neuron” is nonsense

Outline

- Recap
- Artificial Neural Networks
- Perceptron

- Created by Frank Rosenblatt in 1957:

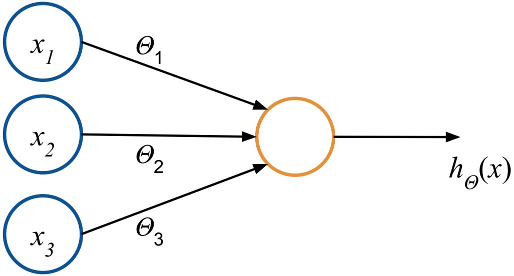




- $\sum_{i=0}^{k} x_i * \theta_i$
- A threshold: σ
- $h_\theta(x) \geq \sigma : y = 1; h_\theta < \sigma : y = 0$

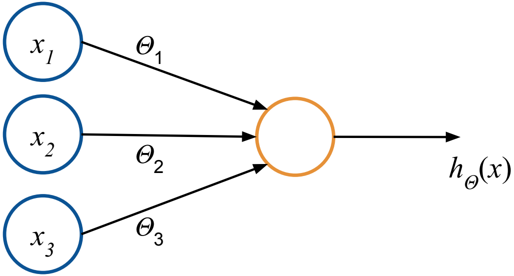
Perceptron

Quiz



- $\sigma = 0$
- $h_{\boldsymbol{w}}(x) = ?$

x	w
1	$\frac{1}{2}$
0	$\frac{3}{5}$
-1.5	1



x	w
1	$\frac{1}{2}$
0	$\frac{3}{5}$
-1.5	1

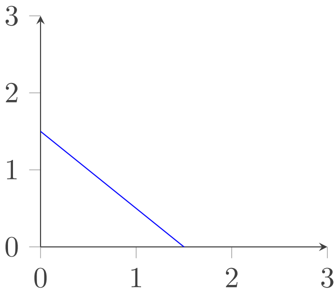
- $\sigma = 0$
- $h_{\boldsymbol{w}}(x) = 0$

How powerful is a perceptron unit?

- $x_1, x_2 \in \{0, 1\}$
- $\boldsymbol{w}_1 = 1/2$
- $\boldsymbol{w}_2 = 1/2$
- $\sigma = 3/4$
- Return 1
- Return 0

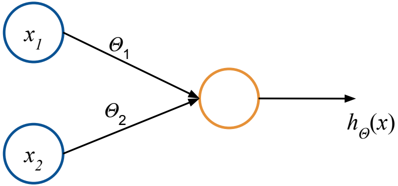
How powerful is a perceptron unit?

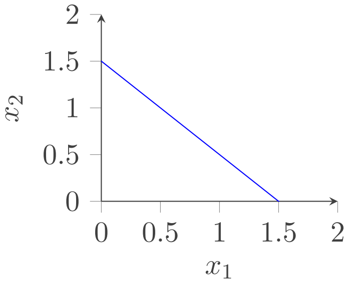
- $x_1, x_2 \in \{0, 1\}$
- $\boldsymbol{w}_1 = 1/2$
- $\boldsymbol{w}_2 = 1/2$
- $\sigma = 3/4$
- Return 1
- Return 0



How powerful is a perceptron unit?

Quiz



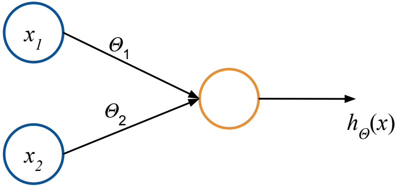


Given:

- $\boldsymbol{w}_1 = \frac{1}{2}$;
- $\boldsymbol{w}_2 = \frac{1}{2}$;
- $\sigma = \frac{3}{4}$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_{\boldsymbol{w}}(x)$ = ?

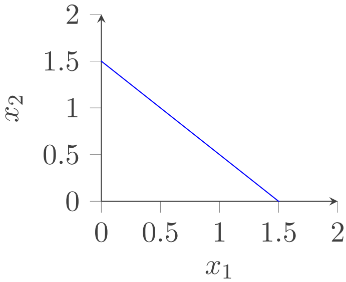
How powerful is a perceptron unit?

Quiz



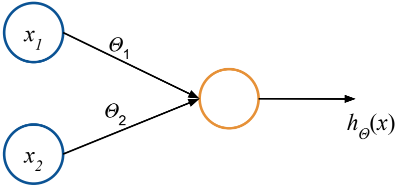
Given:

- $\boldsymbol{w}_1 = \frac{1}{2}$;
- $\boldsymbol{w}_2 = \frac{1}{2}$;
- $\sigma = \frac{3}{4}$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_{\boldsymbol{w}}(x)$ = AND



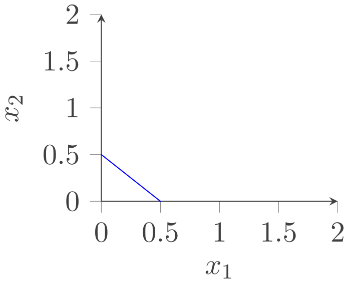
How powerful is a perceptron unit?

Quiz



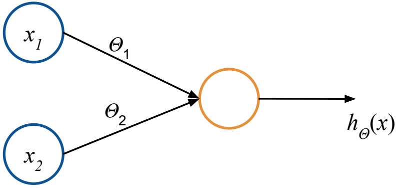
Given:

- $\boldsymbol{w}_1 = ?$;
- $\boldsymbol{w}_2 = ?$;
- $\sigma = ?$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_{\boldsymbol{w}}(x)$ = OR



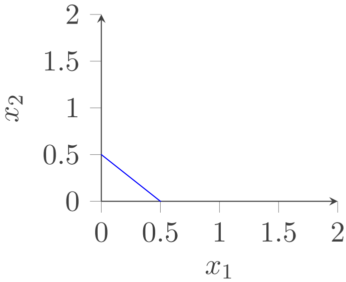
How powerful is a perceptron unit?

Quiz



Given:

- $w_1 = 1$;
- $w_2 = 1$;
- $\sigma = 0$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_w(x)$ = OR



How powerful is a perceptron unit (NOT)?

Quiz

Given:

- $\boldsymbol{w}_1 = ?$;
- $\sigma = ?$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_{\boldsymbol{w}}(x)$ = NOT

How powerful is a perceptron unit (NOT)?

Quiz

Given:

- $\boldsymbol{w}_1 =$ -1;
- $\sigma =$ 0; and
- $x_1, x_2 \in \{0, 1\}$
- $h_{\boldsymbol{w}}(x)$ = NOT

How powerful is a perceptron unit (NOT)?

Quiz

Given:

- $\boldsymbol{w}_1 =$;
- $\sigma =$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_{\boldsymbol{w}}(x)$ = NOT

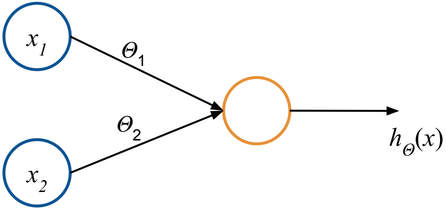
AND, OR, NOT: expressible as perceptron units.
Can represent anything!

How powerful is a perceptron unit?

Quiz

Given:

- $w_1 = ?$;
- $w_2 = ?$;
- $\sigma = ?$; and
- $x_1, x_2 \in \{0, 1\}$
- $h_w(x)$ = XOR



To continue in the next session.