
Machine Learning Essentials

Mehrdad Zakershahrak
Guest Speaker
Arizona State University

What is Machine Learning?

Machine Learning

- Field of study that gives computers the ability to **learn without being explicitly programmed.**
- Arthur Samuel (1959)

Applications of Machine Learning



- | **Spam filter in email systems**
- | **Computer vision**
- | **Language translation**
- | **Amazon and Netflix suggestions**

Machine Learning Algorithms

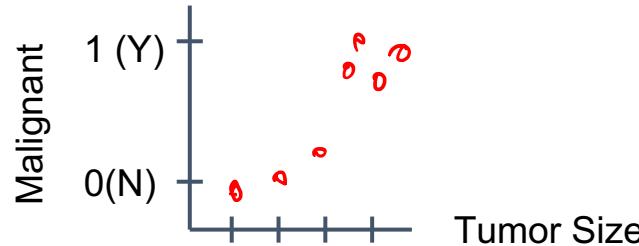


- | **Supervised learning**
- | **Unsupervised learning**
- | **Reinforcement learning**
- | **Other**

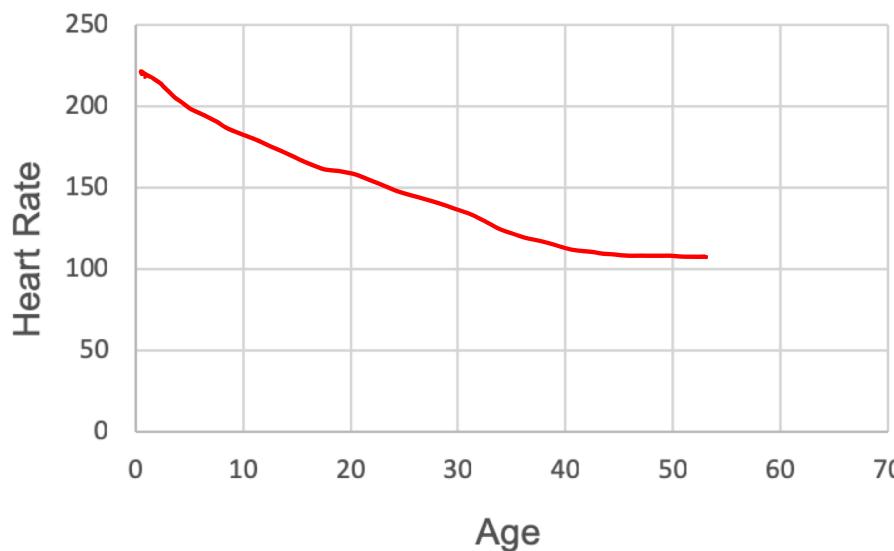
Supervised Learning

Right answer is given → data is labeled

Regression: Predict continuous valued output

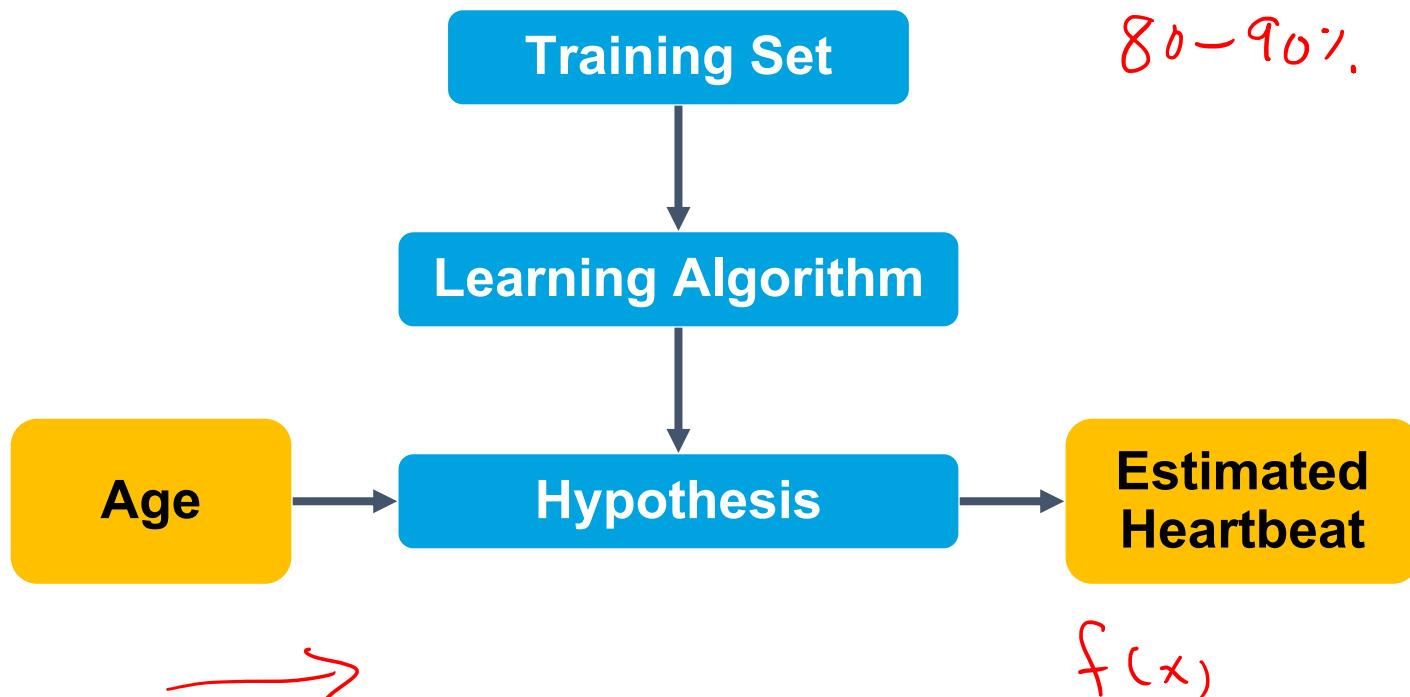


Classification: Predict discrete valued output



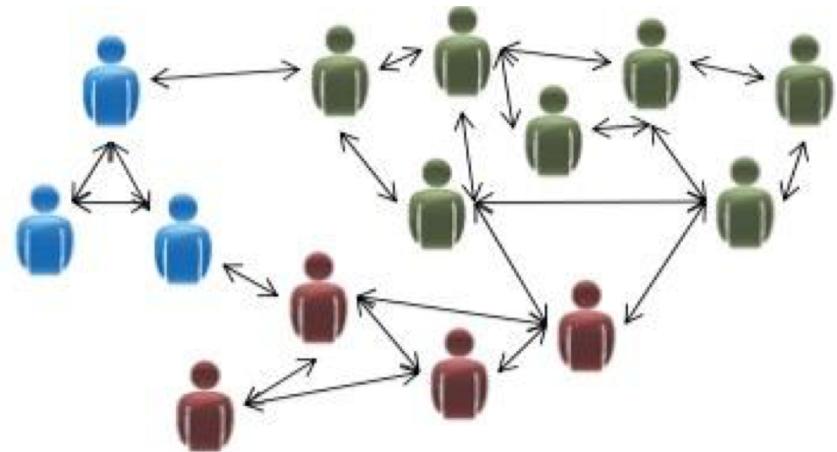
Model Representation

| How do we represent the hypothesis?



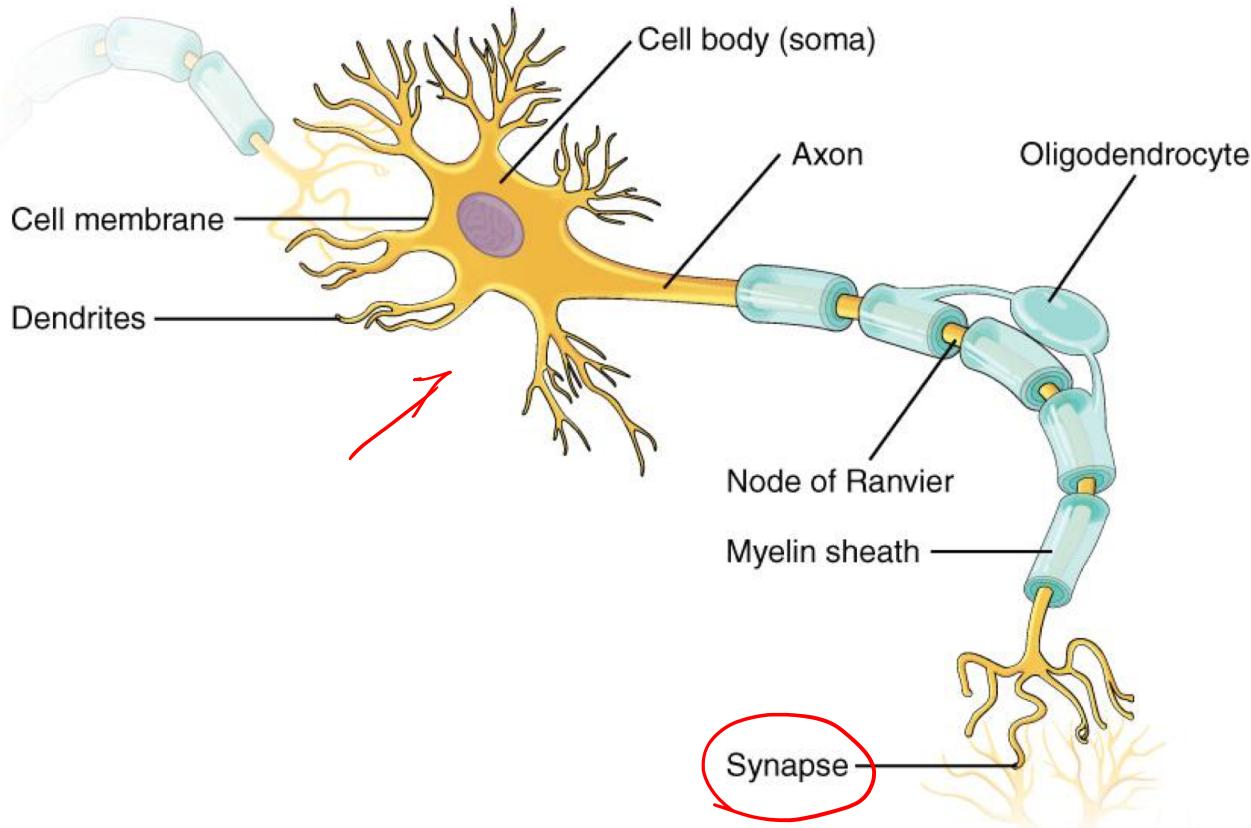
Unsupervised Learning

- | Social media analysis
- | Market segmentation

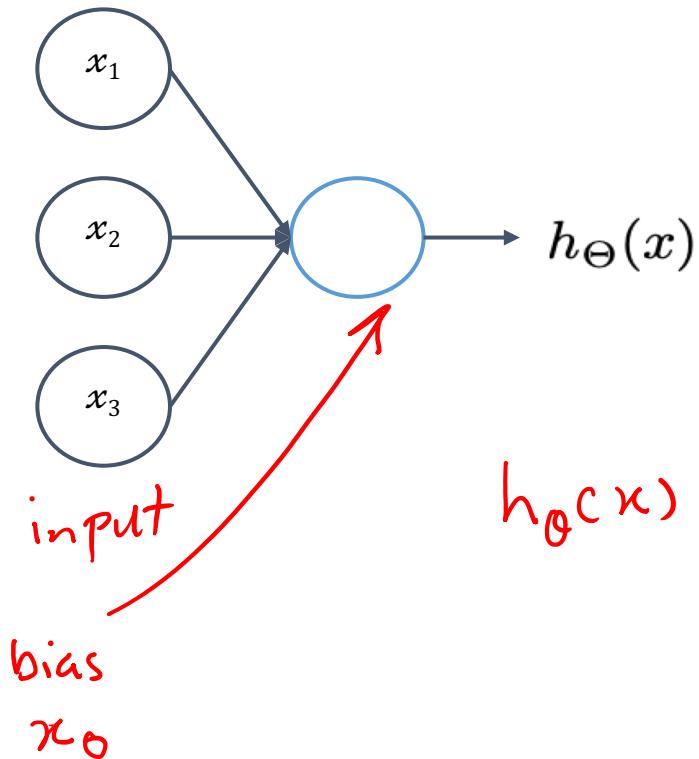


Neural Networks

Mimic the neurons structure in the brain.



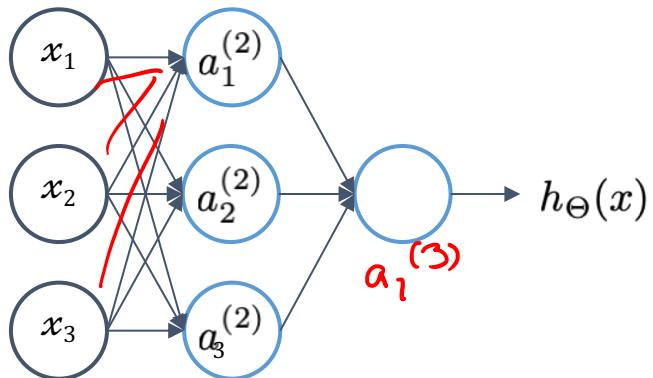
Neural Networks



$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

$$h_{\Theta}(x) = \frac{1}{1 + e^{-x\theta}}$$

Neural Networks



$a_i^{(j)}$ = “activation” of unit i in layer j

$\Theta^{(j)}$ = matrix of weights controlling
function mapping from layer j to
layer $j + 1$

input

$$a_1^{(2)} = g(\underline{\Theta_{10}^{(1)}} \underline{x_0} + \Theta_{11}^{(1)} x_1 + \Theta_{12}^{(1)} x_2 + \Theta_{13}^{(1)} x_3)$$

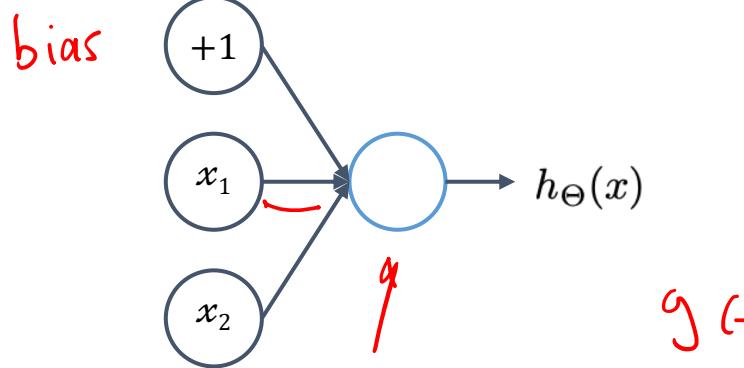
$$a_2^{(2)} = g(\Theta_{20}^{(1)} x_0 + \underline{\Theta_{21}^{(1)}} x_1 + \Theta_{22}^{(1)} x_2 + \Theta_{23}^{(1)} x_3)$$

$$\underline{h_\Theta(x)} = a_1^{(3)} = g(\underline{\Theta_{10}^{(2)}} \underline{a_0^{(2)}} + \underline{\Theta_{11}^{(2)}} \underline{a_1^{(2)}} + \underline{\Theta_{12}^{(2)}} \underline{a_2^{(2)}} + \underline{\Theta_{13}^{(2)}} \underline{a_3^{(2)}})$$

Example

Simple example: AND

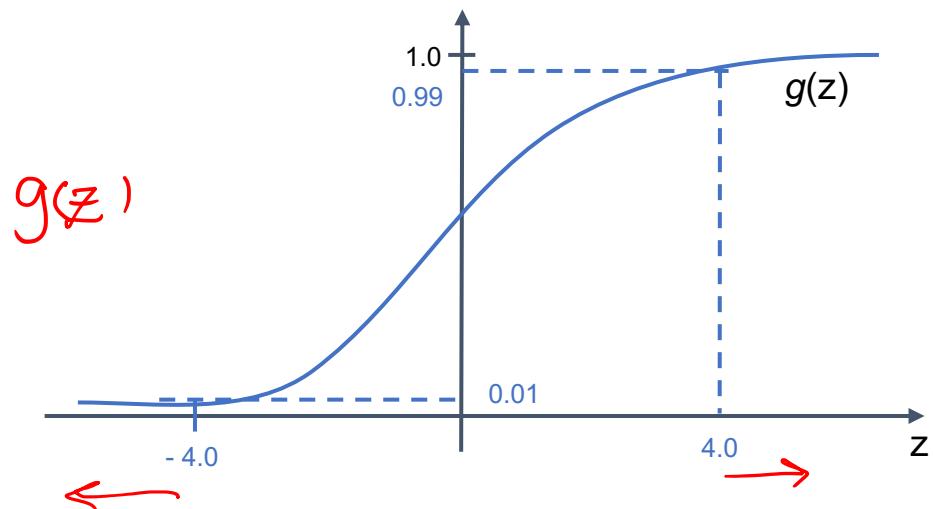
- $x_1, x_2 \in \{0, 1\}$
- $y = x_1 \text{AND} x_2$



$$h_{\Theta}(x) = g(-30 + 20x_1 + 20x_2)$$

$\Theta_{10}^{(1)}$ $\Theta_{11}^{(1)}$ $\Theta_{12}^{(1)}$

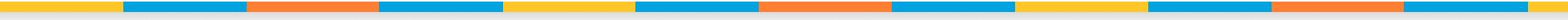
=



x_1	x_2	$h_{\Theta}(x)$
0	0	$g(-30) \approx 0$
0	1	$g(-10) \approx 0$
1	0	$g(-10) \approx 0$
1	1	$g(10) \approx 1$

$h_{\Theta}(x) \approx X_1 \text{ AND } X_2$

Summary



- | **Machine learning definition**
- | **Applications of machine learning**
- | **Machine learning algorithms**
 - Supervised
 - Unsupervised
- | **Neural networks**