COSC 522 – Machine Learning

Biological Neuron and Perceptron

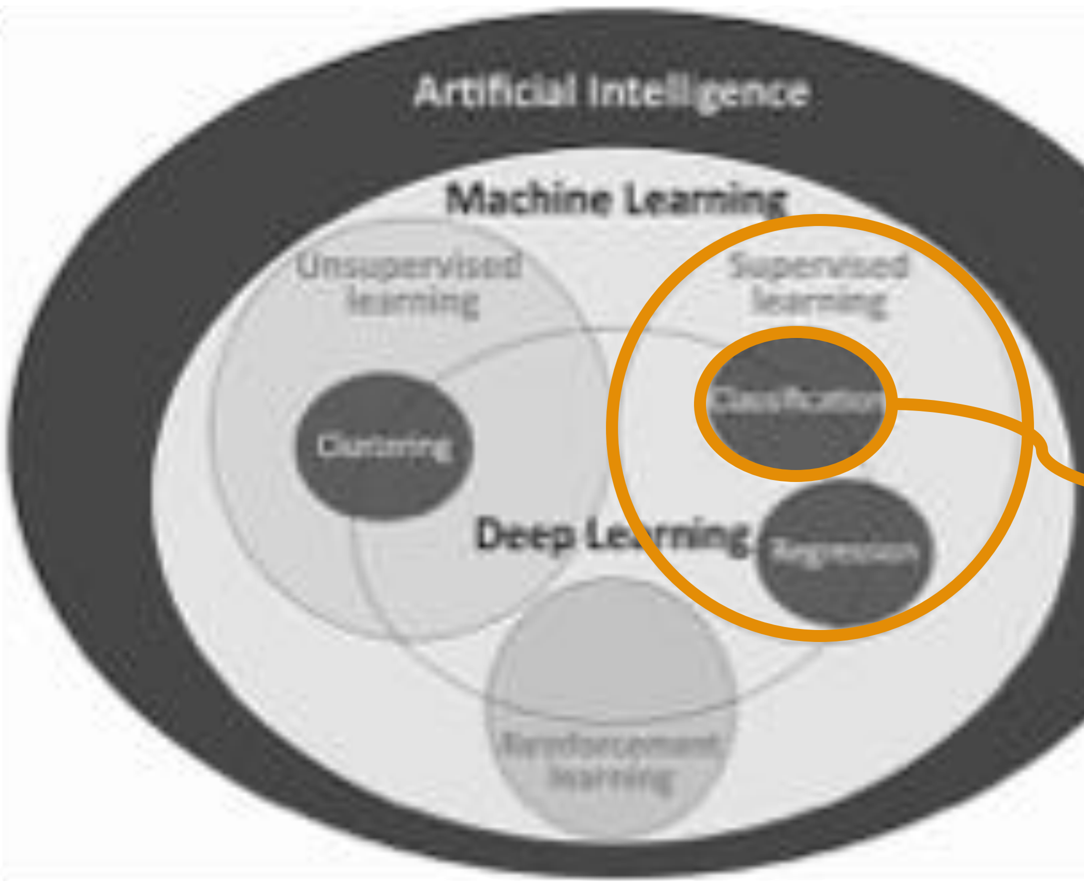
Hairong Qi, Gonzalez Family Professor
Electrical Engineering and Computer Science
University of Tennessee, Knoxville
https://www.eecs.utk.edu/people/hairong-qi/
Email: hqi@utk.edu

Questions

- The anatomy of biological neuron
- What is action potential and how does it work?
- The analogy between biological neuron and perceptron

- What is the cost function of perceptron?
- How is perceptron trained?
- What is the limitation of perceptron?
- What is an epoch?

Where Are We?



Part 1: Statistical Methods	
Baysian Learning	
	Introduction
	Baysian Decision Theory and Parametric Learning
	Baysian Decision Theory and Non-Parametric Learning
	Case Study: Representation for Natural Language (taught by Andrei Cozma)
	Parametric vs. Non-Parametric Learning: Some In-Depth Discussion
	Homework and Project Discussion (taught by Fanqi Wang)
Neural Networks	
	Biological Neuron and Perceptron
	Back Propagation and Gradient Descent
	Kernel Methods
	Support Vector Machine
	SVM

M. Mafu, "Advances in artificial intelligence and machine learning for quantum communication applications," IET Quantum Communication, 2024, DOI: 10.1049/qtc2.12094

AICIP RESEARCH

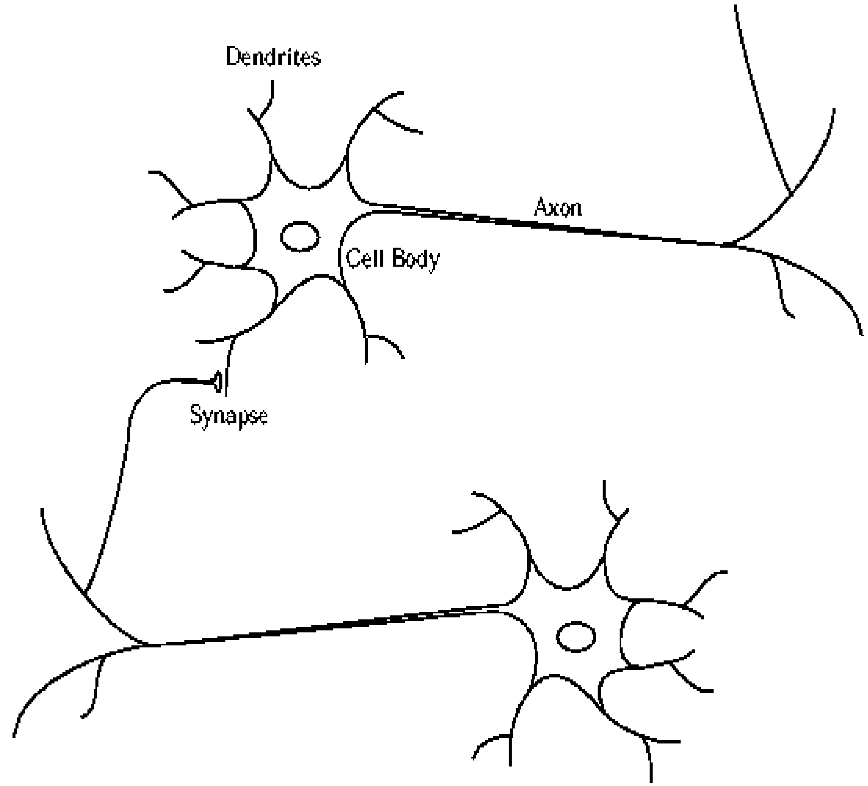
Part I: Biological Neuron

A wrong direction – The first AI winter

- One argument: Instead of understanding the human brain, we understand the computer. Therefore, NN dies out in 70s.
- 1980s, Japan started “the fifth generation computer research project”, namely, “knowledge information processing computer system”. The project aims to improve logical reasoning to reach the speed of numerical calculation. This project proved an abortion, but it brought another climax to AI research and NN research.

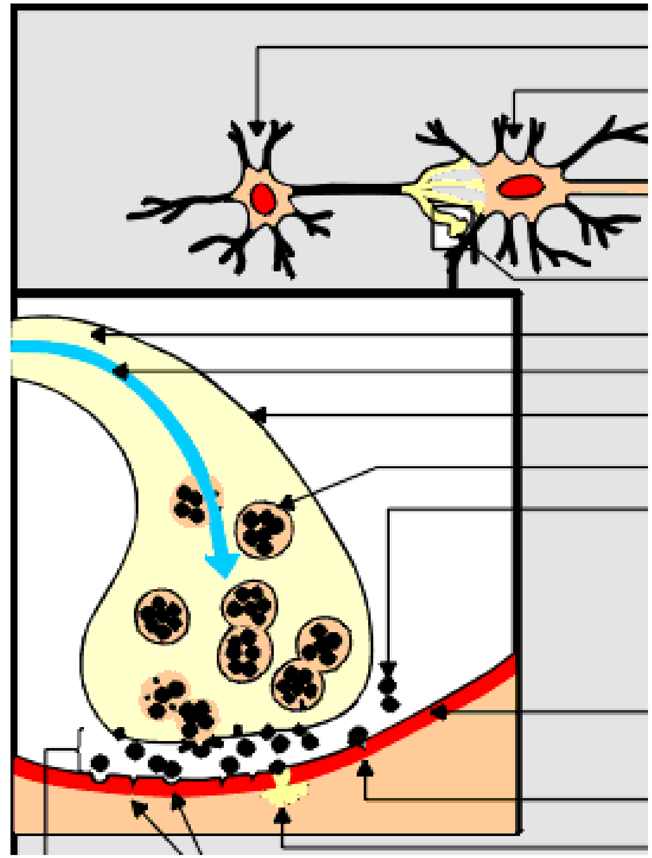
Biological neuron

- Dendrites: tiny fibers which carry signals to the neuron cell body
- Cell body: serves to integrate the inputs from the dendrites
- Axon: one cell has a single output which is axon. Axons may be very long (over a foot)
- Synaptic junction: an axon impinges on a dendrite which causes input/output signal transitions



Synapse

- Communication of information between neurons is accomplished by movement of chemicals across the synapse.
- The chemicals are called neurotransmitters (generated from cell body)
- The neurotransmitters are released from one neuron (the presynaptic nerve terminal), then cross the synapse and are accepted by the next neuron at a specialized site (the postsynaptic receptor).

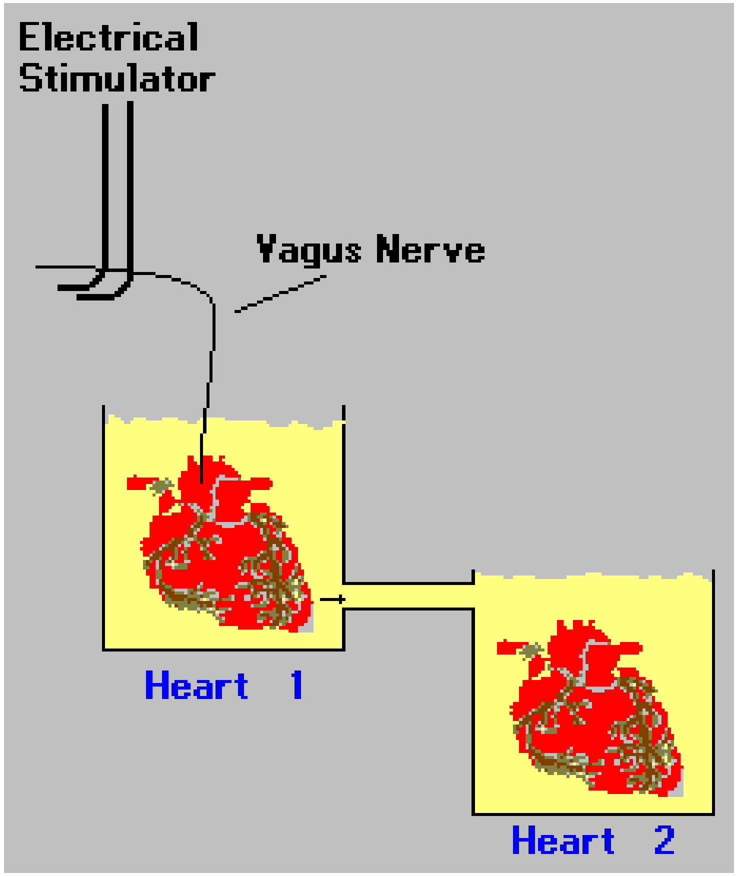


http://faculty.washington.edu/chudler/chnt1.html

The discovery of neurotransmitters

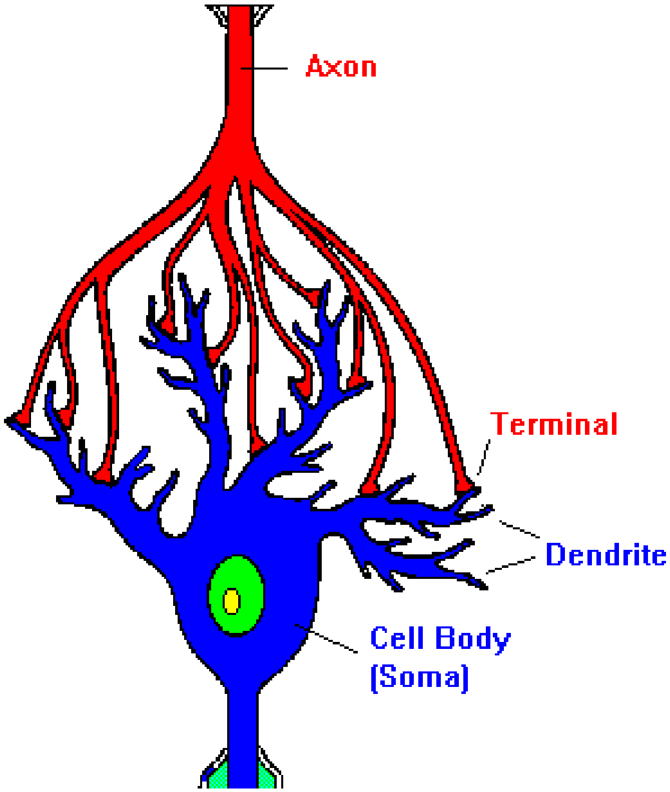
AICIP RESEARCH

- Otto Loewi's Experiment (1920)
- Heart 1 is connected to vagus nerve, and is put in a chamber filled with saline
- Electrical stimulation of vagus nerve causes heart 1 to slow down. Then after a delay, heart 2 slows down too.
- Acetylcholine

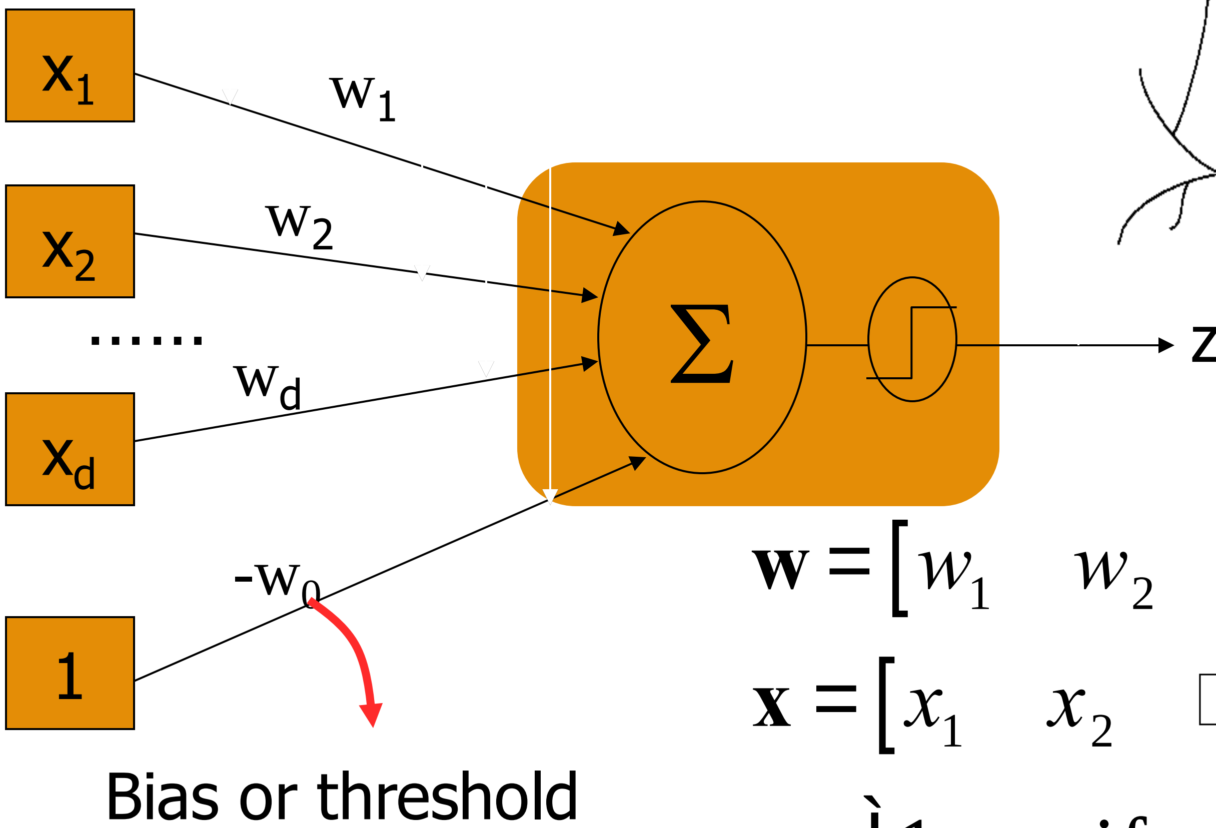


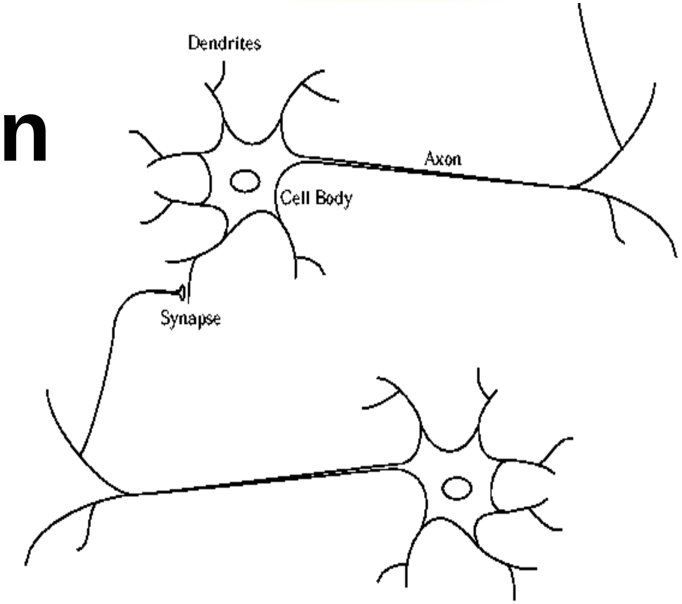
Action potential

- When a neurotransmitter binds to a receptor on the postsynaptic side of the synapse, it results in a change of the postsynaptic cell's excitability: it makes the postsynaptic cell either more or less likely to fire an action potential. If the number of excitatory postsynaptic events are large enough, they will add to cause an action potential in the postsynaptic cell and a continuation of the "message."
- Many psychoactive drugs and neurotoxins can change the properties of neurotransmitter release, neurotransmitter reuptake and the availability of receptor binding sites.



The analogy: Biological neuron and Perceptron



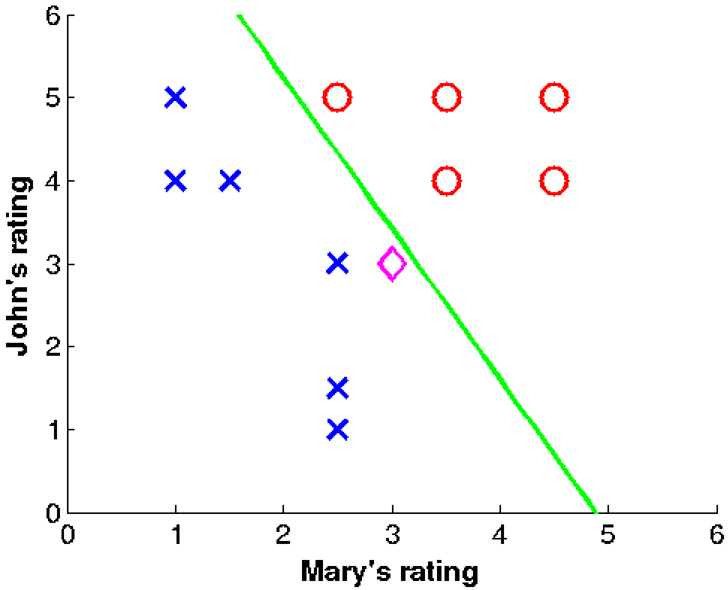


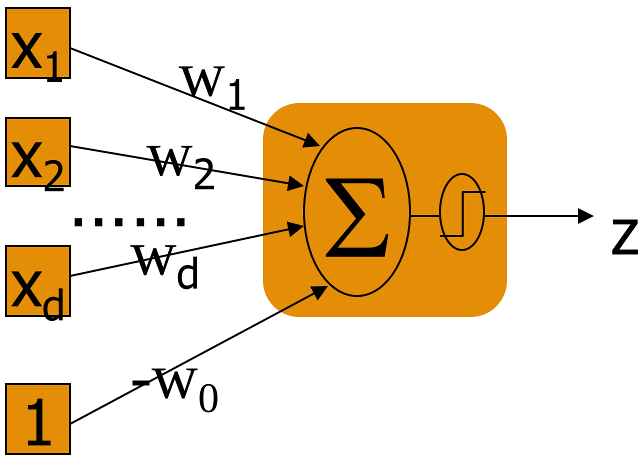
$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_d \end{bmatrix}$$

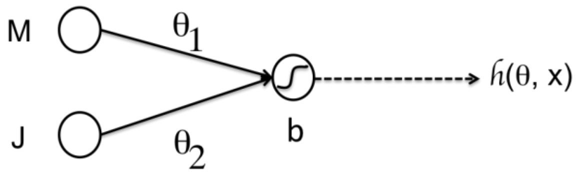
$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix}$$

$$z = \begin{cases} 1 & \text{if } w^T x - w_0 > 0 \\ 0 & \text{otherwise} \end{cases}$$

Revisit the toy example





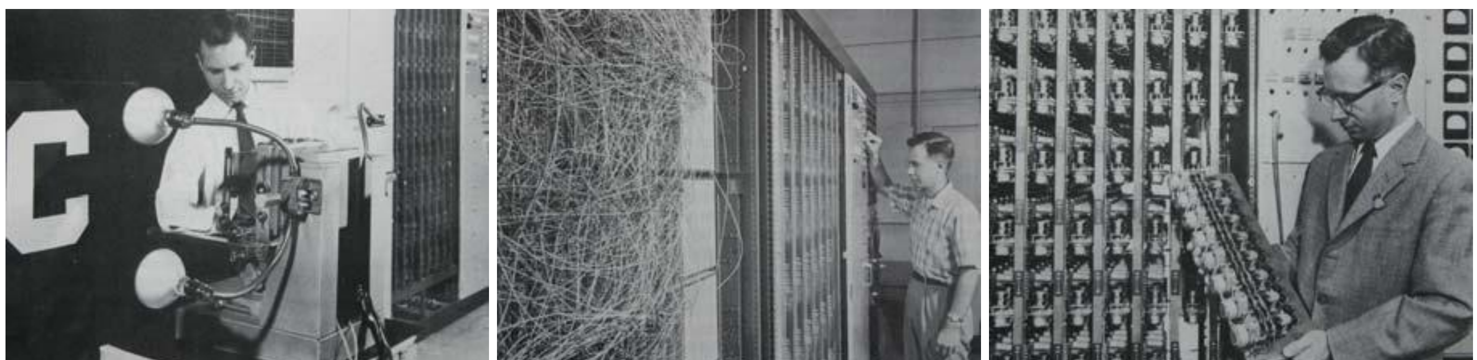


AICIP RESEARCH

Part II: Perceptron

Rosenblatt and Perceptron

- A program that learns “concepts” based on examples and correct answers
- It can only respond with “true” or “false”
- Single layer neural network
- By training, the weight and bias of the network will be changed to be able to classify the training set with 100% accuracy



Rosenblatt and Mark 1 Perceptron (from Bishop’s book)

Perceptron cost function – the perceptron criterion

Initial thought:

$J(\mathbf{w}, w_0) = \mathbf{w}^T\mathbf{x} - w_0 = \mathbf{a}^T\mathbf{y}$

Augmentation:

$\mathbf{a} = [w_1 \; w_2 \; \dots \; w_d \; -w_0]$

$\mathbf{y} = [x_1 \; x_2 \; \dots \; x_d \; 1]$

$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_d \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix}$$

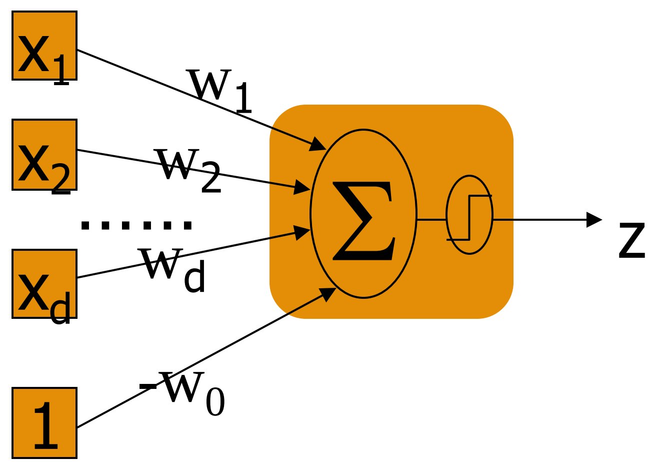
$$z = \begin{cases} 1 & \text{if } w^T x - w_0 > 0 \\ -1 & \text{otherwise} \end{cases}$$

Perceptron criterion: t = {1, -1}

$$\mathbf{J}_p(\mathbf{a}) = \sum_{\mathbf{y} \in Y} (-\mathbf{a}^T\mathbf{y})\,t$$

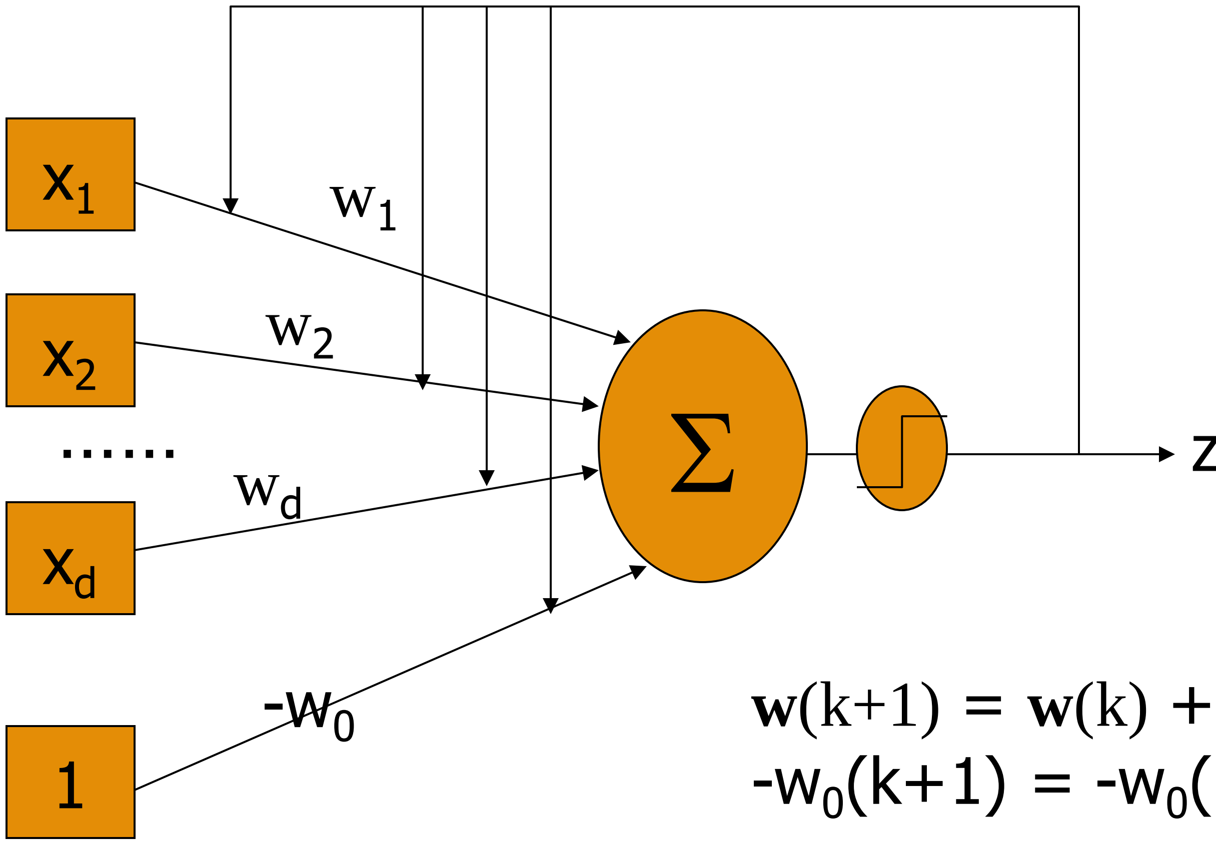
$$\nabla \mathbf{J}_p = \sum_{\mathbf{y} \in Y} (-\mathbf{y})\,t$$

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \eta(k) \sum_{\mathbf{y} \in Y_k} \mathbf{y}\,t$$



Gradient descent learning

Perceptron Learning Rule



$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Sigma\mathbf{x}(T-z)$$

$$-w_0(k+1) = -w_0(k) + \Sigma(T-z)$$

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \eta(k) \sum_{\mathbf{y} \in Y_k} \mathbf{y}\, t$$

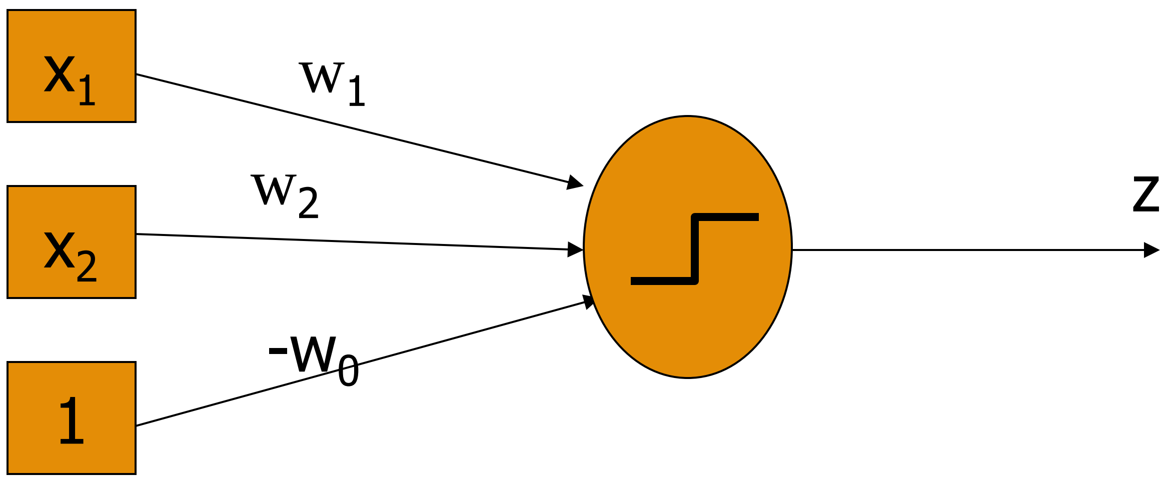
T is the expected output

z is the real output

Training

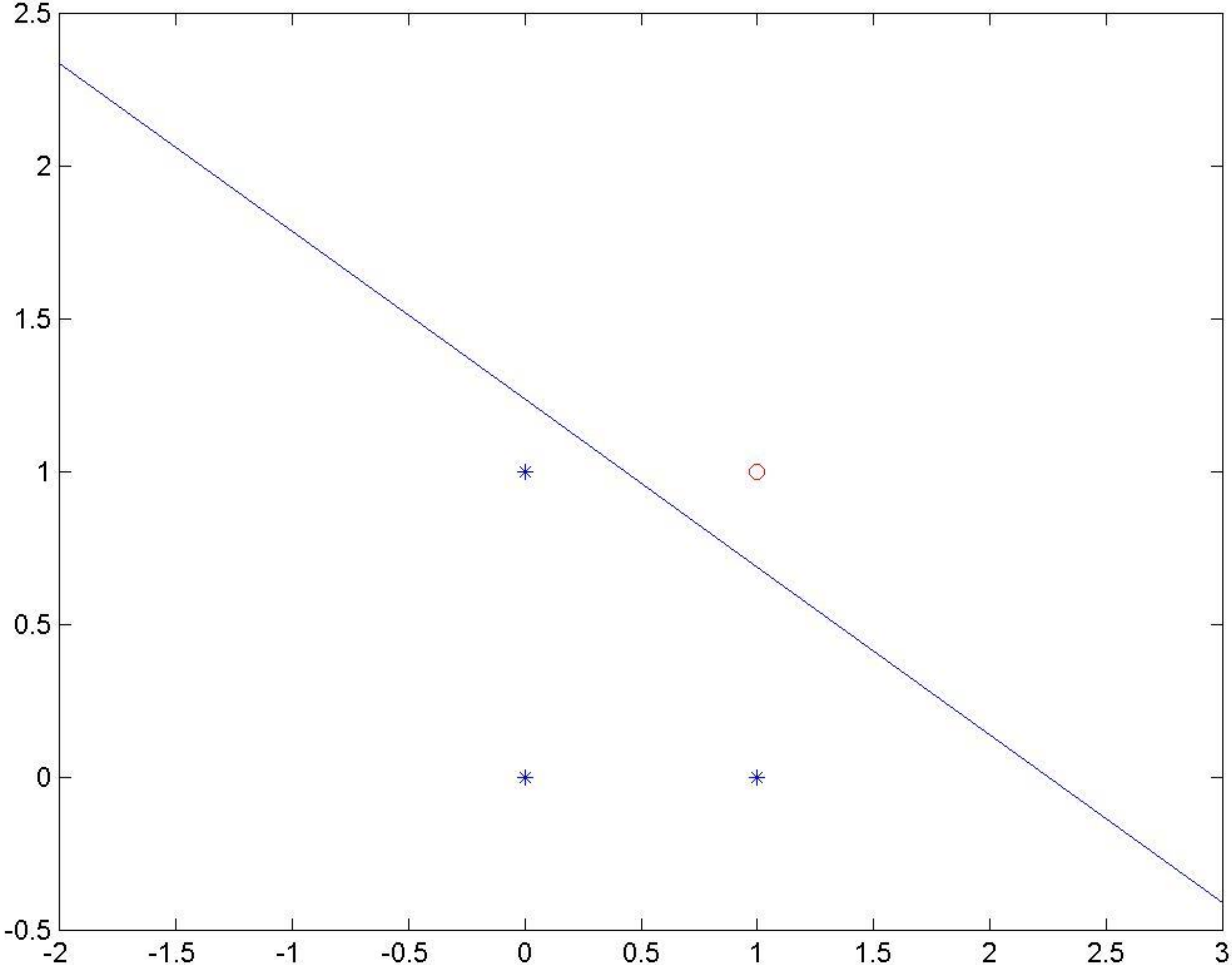
- Step1: Samples are presented to the network
- Step2: If the output is correct, no change is made; Otherwise, the weight and biases will be updated based on perceptron learning rule. That is,
  - For Class 1, add **x** onto the current estimate **w**
  - For Class -1, subtract **x** from **w**
- Step3: An entire pass through all the training set is called an “epoch”. If no change has been made for the epoch, stop. Otherwise, go back Step1

Exercise (AND Logic)



x1	x2	T
0	0	0
1	0	0
0	1	0
1	1	1

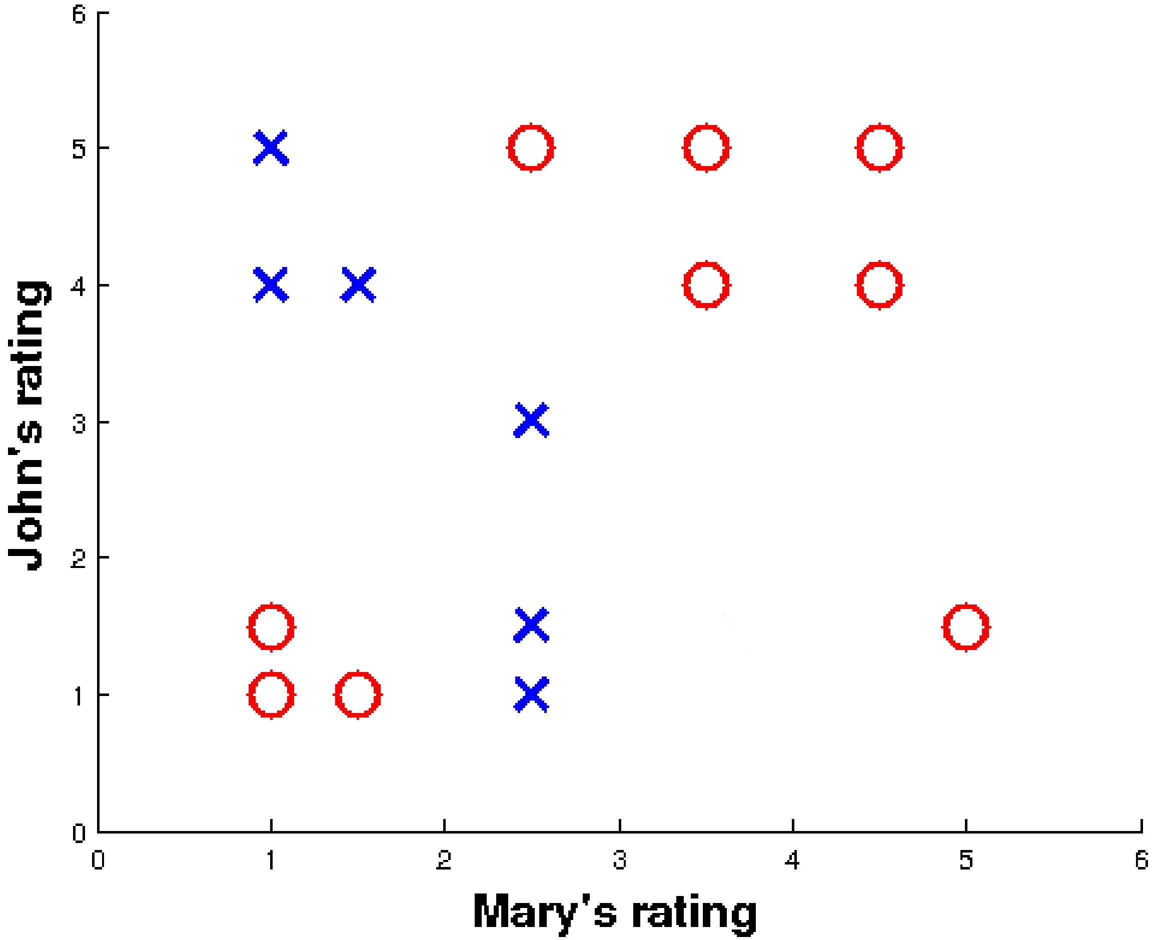
w1 w2 $-w_0$


2.5
2
1.5
1
0.5
0
-0.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3

Limitations (Limited View)

- The output only has two values (1 or 0)
- Can only classify samples which are linearly separable (straight line or straight plane)
- Can’t train network functions like XOR

Back to the toy example



?