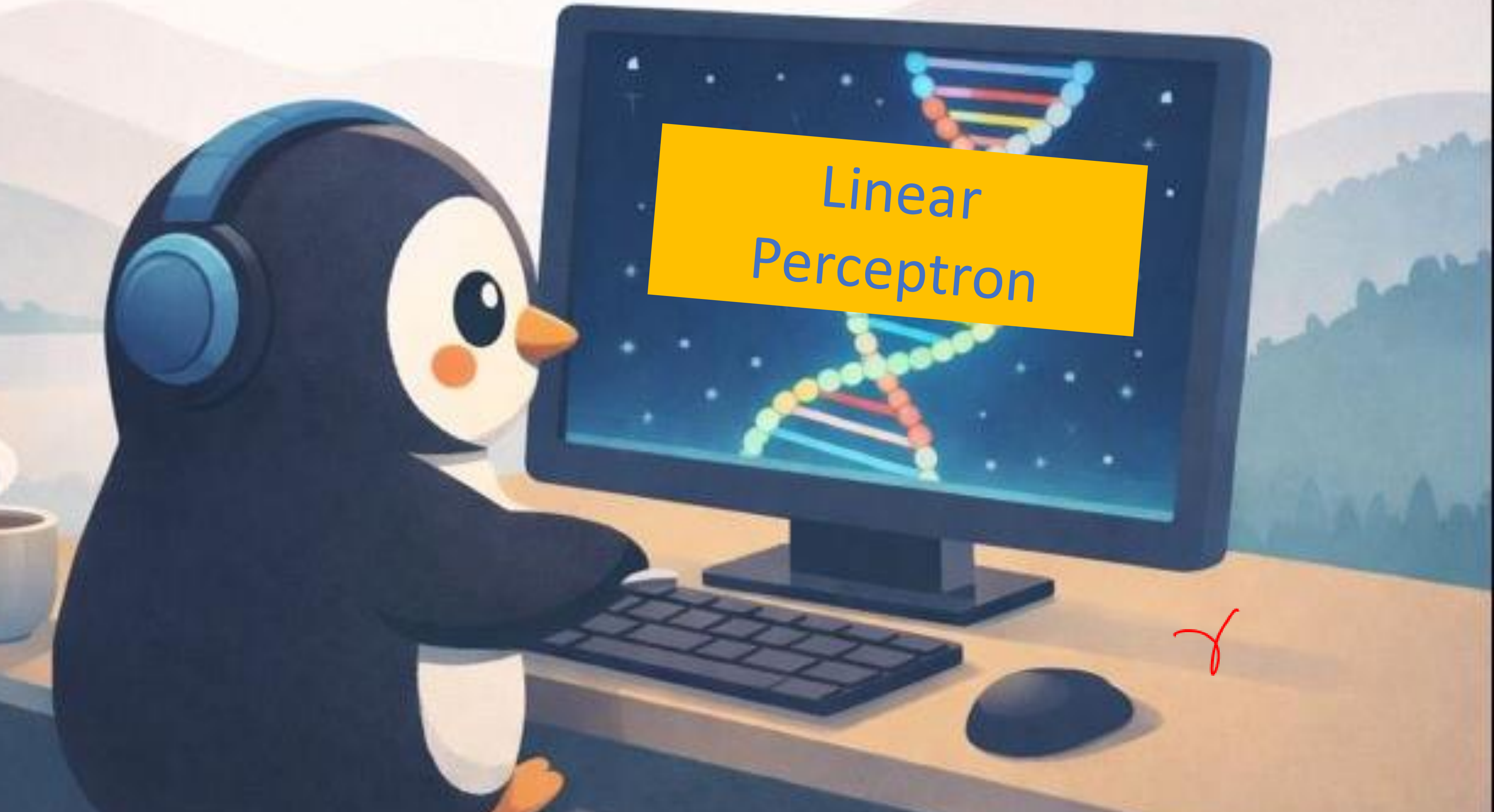
Machine Learning for Biology and Health

CSCI 1851
Spring 2026

Ritambhara Singh

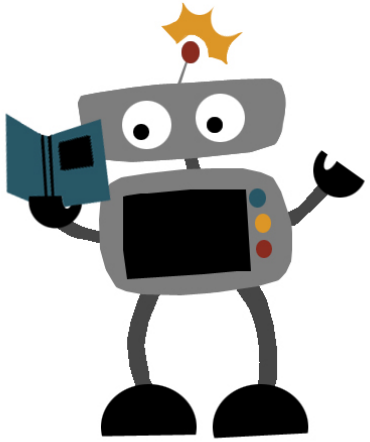
January 27, 2026
Tuesday

Linear Perceptron



ChatGPT prompt “minimalist landscape painting a cute penguin working on a computer with DNA on its screen”

Recap: What is Machine Learning?



Input: X

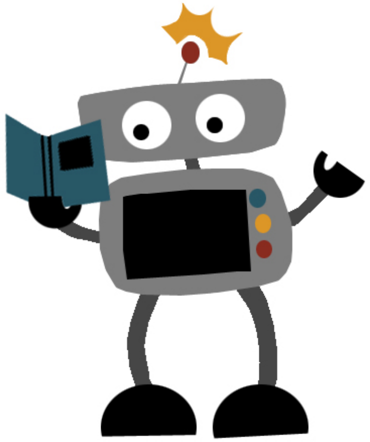
Output: Y

"Cooking?"

Function: f

$\boxed{f(X) \rightarrow Y}$

Recap: What is Machine Learning?



Supervised Learning

Input: X

Learned function: f

Output: Y

"Cooking?"

$f(X) \rightarrow Y$

Today’s goal - Learn about binary classification using perceptron

(1) Machine Learning problem – Heart Disease Classification

(2) Linear Perceptron

(3) Parameters – weights and biases

(4) Perceptron Learning Algorithm

(5) Class Activity: Perceptron in action

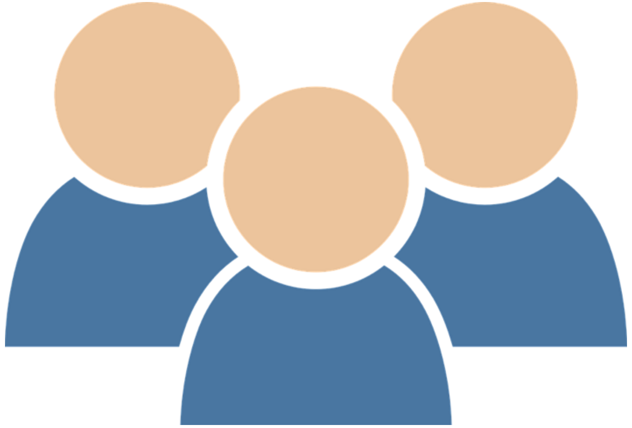
(6) Evaluating classification performance

Heart Disease Classification

Goal: Predict whether a patient has heart disease

Input: X

"Vitals/information about the patient"



Output: Y

"Heart Disease/No Heart Disease"

Function: f

$\boxed{f(X) \rightarrow Y}$

But first some notations…

$\mathbb{X}: A\ set\ of\ input\ data$

$\mathbb{Y}: Associated\ set\ of\ target\ values\ (outputs)\ for\ supervised\ learning$

$x^{(k)}$: kth example (input) from a dataset

$y^{(k)}: Target\ (output) associated\ with\ x^{(k)} for\ supervised\ learning$

$\mathbb{R}$: A set of real numbers

$\mathbb{Z}$: A set of integers

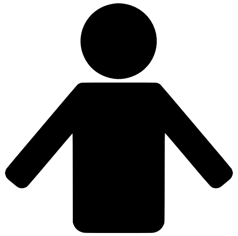
{0,1}: A set of binary numbers

Our Problem: Heart disease classification

What could be the input features?

Input: $\mathbb{X}$

Feature	Description
`age`	Age in years
`cholesterol`	Serum cholesterol (mg/dL)
`resting_bp`	Resting blood pressure (mm Hg)
`max_hr`	Maximum heart rate achieved



Function: f

$f(\mathbb{X}) \rightarrow \mathbb{Y}$

What could be the output labels?

Target: $\mathbb{Y}$

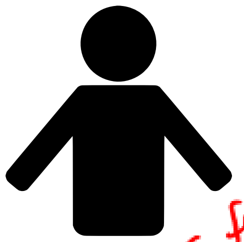
Do they have heart disease?

Yes/No

Our problem: Heart disease classification

binary

$y \in \{0,1\}$



$x \in \mathbb{X}$

features

Input: $\mathbb{X}$

y

What is our input space?

What is our output space?

What is our prediction task?

Target: $\mathbb{Y}$

samples

10

age,cholesterol,resting_bp,max_hr
45,180,120,170
48,185,118,172
50,190,125,165
52,200,130,160
55,210,135,158
57,220,140,150
58,235,142,152
60,240,145,148
62,250,150,140
65,260,155,138

Function: f

$f(\mathbb{X}) \rightarrow \mathbb{Y}$

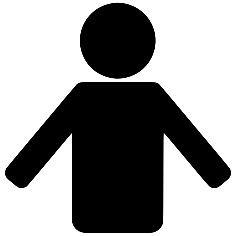
heart_disease
0
0
0
0
0
1
1
1
1
1

class 1

class 2

How many samples?

Simpler example: Let’s work in 2D



Input: $\mathbb{X}$

“Age, Cholestrol”

Binary classification

Target: $\mathbb{Y}$

“Has heart disease?”

$x^{(1)}$ 45, 180

$x^{(2)}$ 48, 185

$x^{(3)}$ 62, 250

$x^{(4)}$ 65, 260

$\mathbb{X} \in \mathbb{Z}$

Function: f

f(X)→Y

Do you see a trend here?

$y^{(1)}$ 0

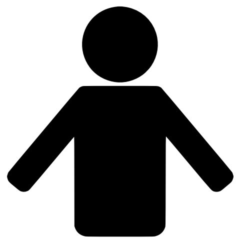
$y^{(2)}$ 0

$y^{(3)}$ 1

$y^{(4)}$ 1

$\mathbb{Y} \in \{0,1\}$

Simpler example: Learning function f



Input: $\mathbb{X}$

"Age, Cholestrol"

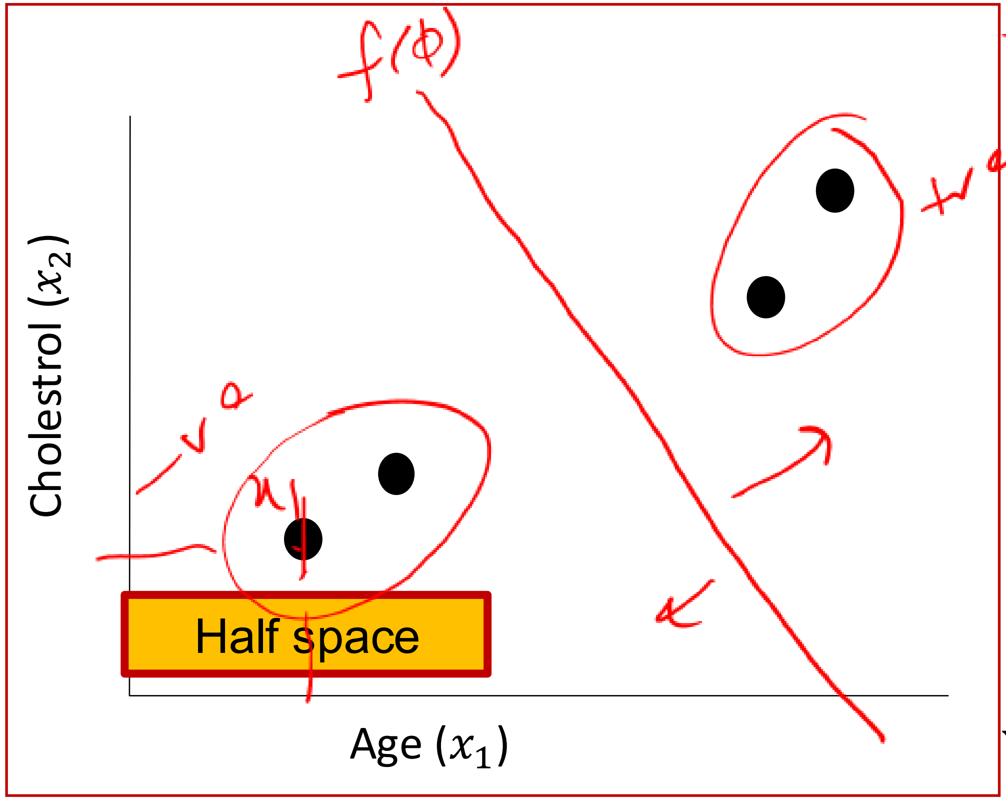
Binary classification

Target: $\mathbb{Y}$

"Has heart disease?"

Linear decision boundary

$x^{(1)}$	45, 180		$y^{(1)}$	0
$x^{(2)}$	48, 185		$y^{(2)}$	0
$x^{(3)}$	62, 250		$y^{(3)}$	1
$x^{(4)}$	65, 260		$y^{(4)}$	1



$\mathbb{X} \in \mathbb{Z}$

$\mathbb{Y} \in \{0,1\}$

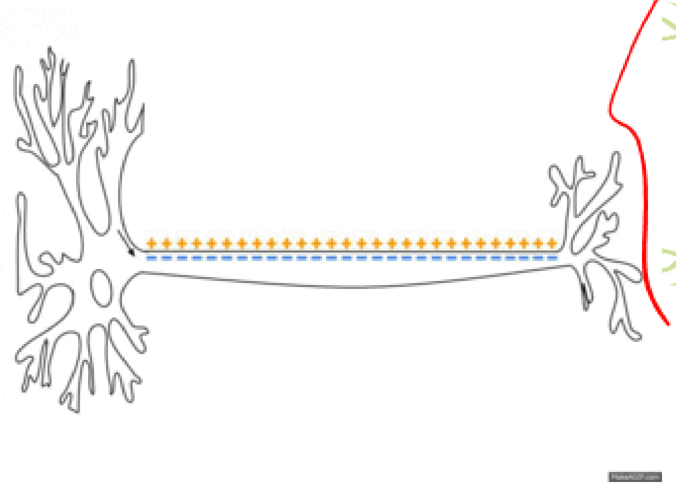
Any questions?

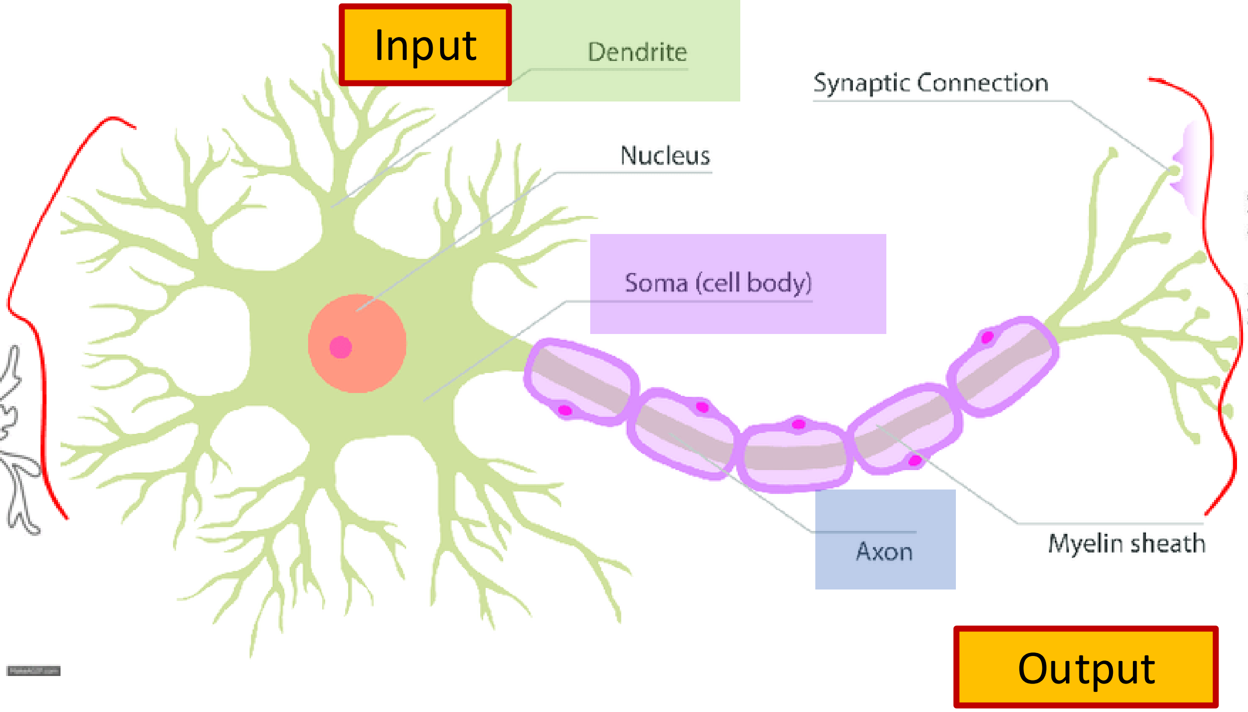
Linear Perceptron

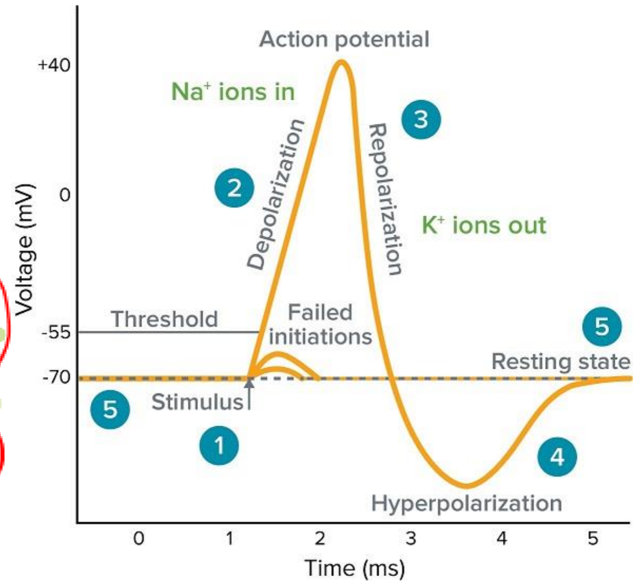
Biological motivation

- Loosely inspired by neurons, basic working unit of the brain
- Serve to transmit information between cells

Inputs are membrane depolarizations!



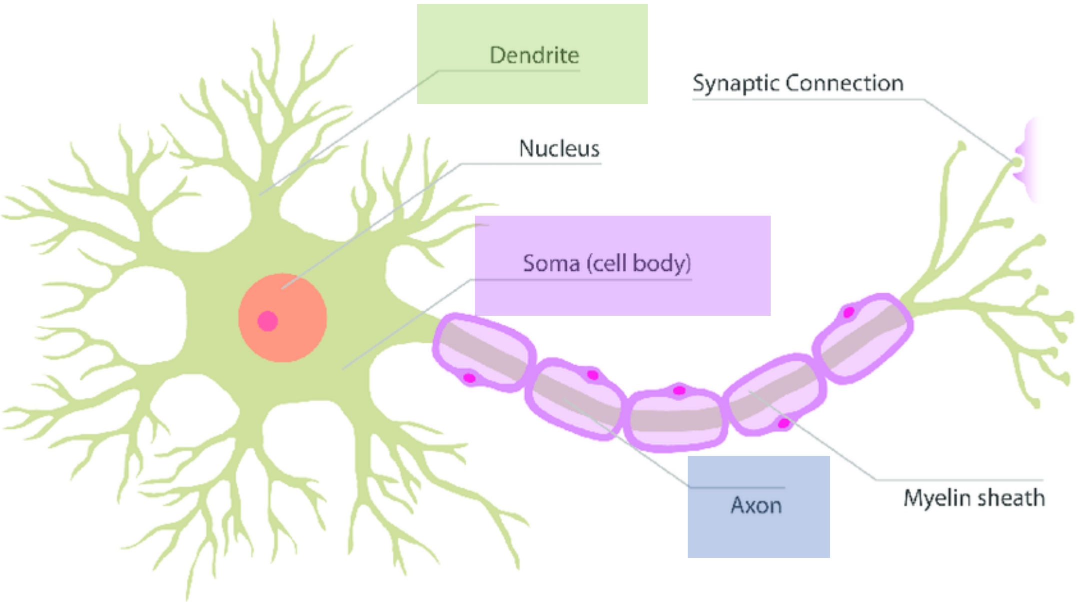




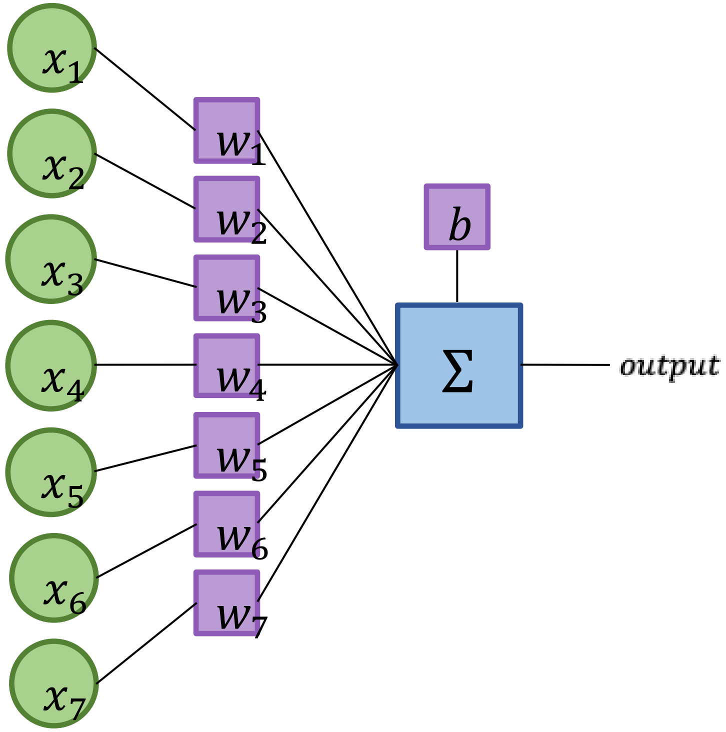
(Depolarization is a change within a cell, during which the cell undergoes a shift in electric charge distribution)
https://en.wikipedia.org/wiki/Depolarization

Image from [here]

The Perceptron



Biological Neuron

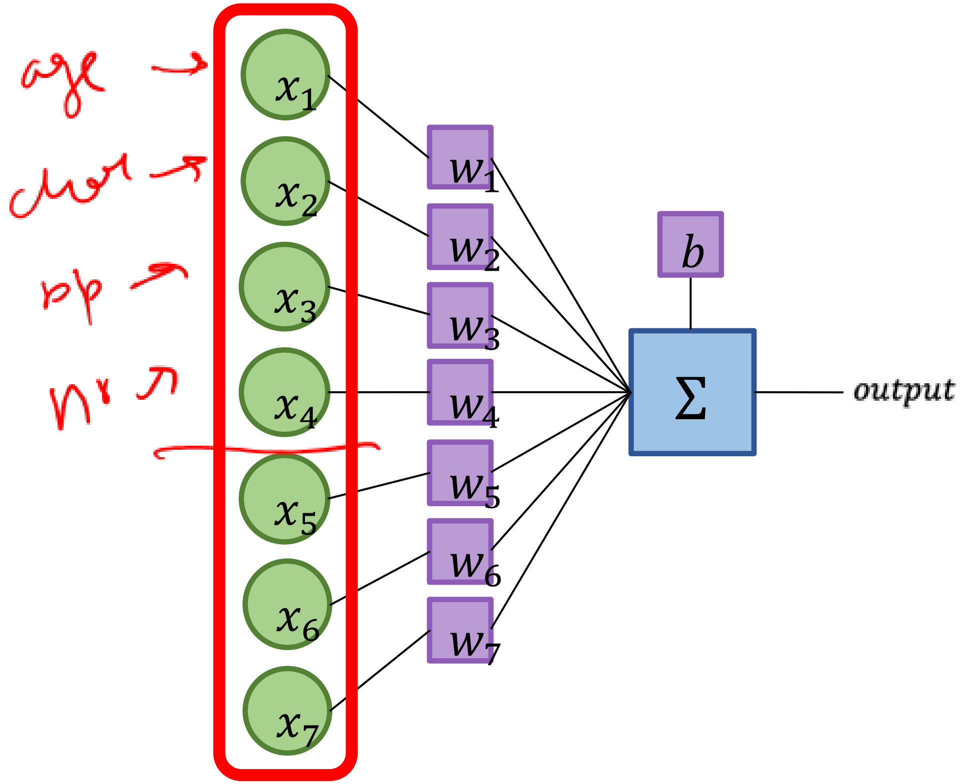


Artificial Neuron (Perceptron)

Input

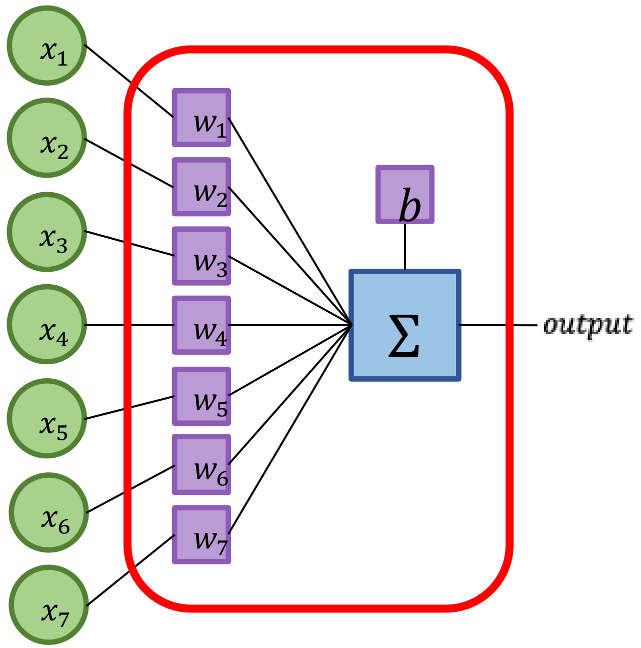
- Input: a vector of numbers x $= [x_1, x_2, \dots x_k]$

What is x_i for our problem?



Predicting with a Perceptron

1. Multiply each input x_i by its corresponding weight w_i, sum them up.
2. Add the bias b



Predicting with a Perceptron

1. Multiply each input x_i by its corresponding weight w_i, sum them up.
2. Add the bias b

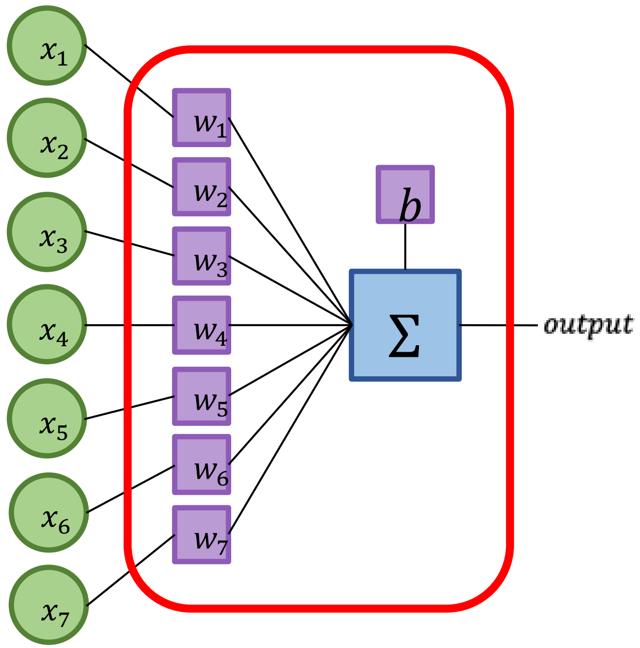
score = wx+b

y=mx+c

slope intercept

3. If the result value is greater than 0, return 1, otherwise return 0

$$f_{\Phi}(x) = \begin{cases} 1, & if\ b + \sum_{i=0}^{n} w_i\, x_i > 0 \\ 0, & otherwise \end{cases}$$



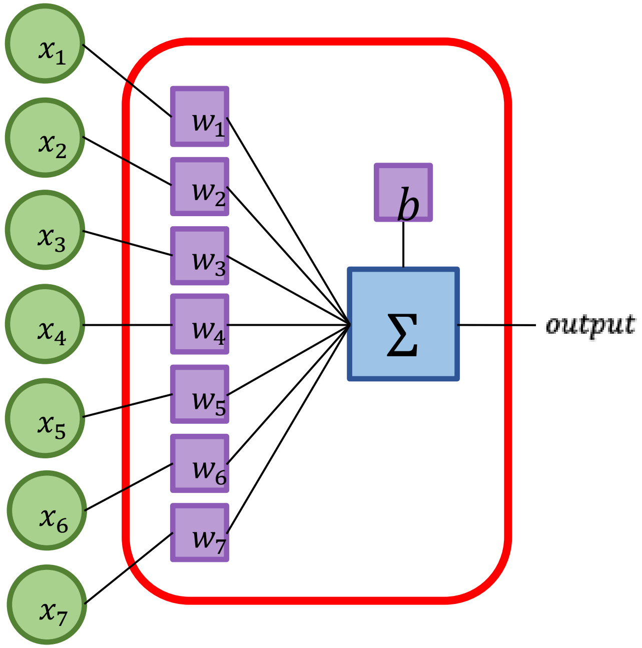
Does this equation remind you of anything?

Threshold value = 0

Performs binary classification!

Parameters

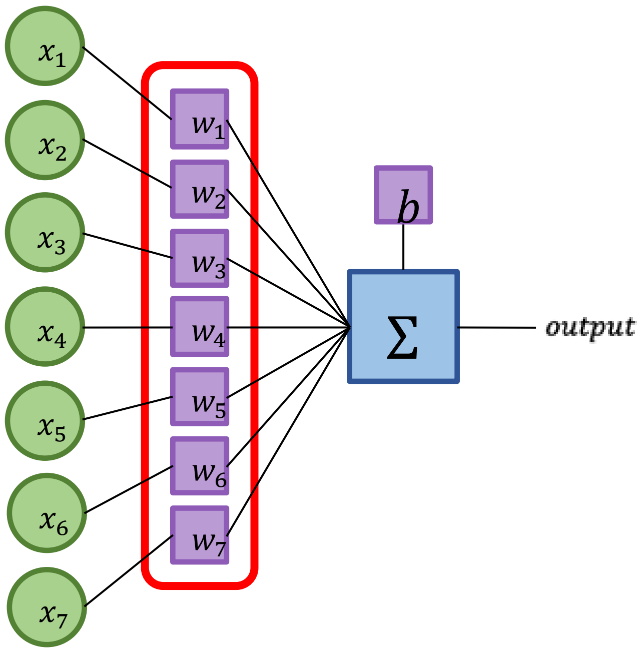
- w and b are parameters of the perceptron
  - Parameters: values we adjust during learning
  - Let $\Phi = \{w \cup b\}$ (the set of all parameters)



Parameters

- ***Weights*** — the importance of each input to determining the output
  - Weight near 0 imply this input has little influence on the output
  - Negative weight means?

$$f_{\Phi}(x) = \begin{cases} 1, & if\ b + \sum_{i=0}^{n} w_i\, x_i > 0 \\ 0, & otherwise \end{cases}$$



Option 1: Increasing input will increase output

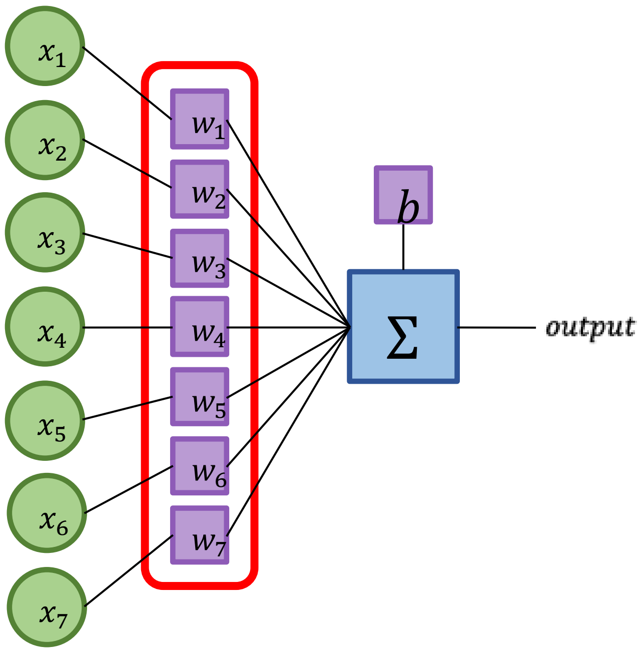
Option 2: Increasing input will decrease output

Option 3: Decreasing input will decrease output

Parameters

- ***Weights*** — the importance of each input to determining the output
  - Weight near 0 imply this input has little influence on the output
  - Negative weight means increasing the input will decrease the output

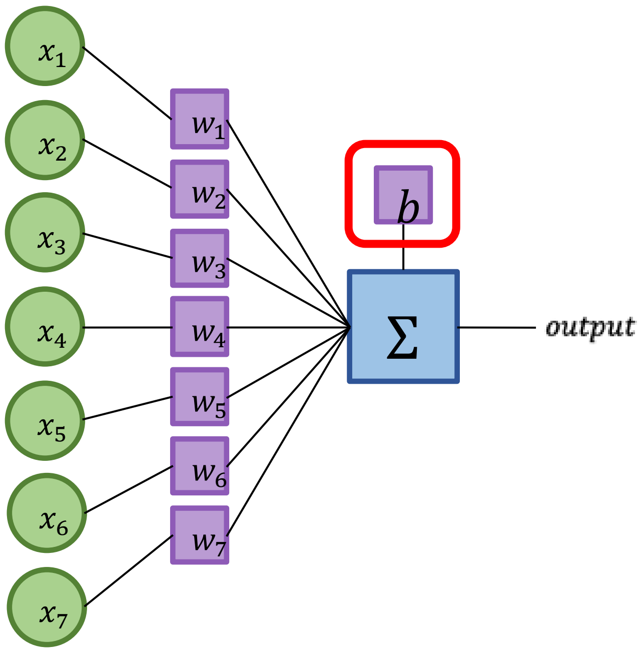
$$f_{\Phi}(x) = \begin{cases} 1, & if\ b + \sum_{i=0}^{n} w_i\, x_i > 0 \\ 0, & otherwise \end{cases}$$



Parameters

- ***Bias*** — What do we need this for?

$$f_{\Phi}(x) = \begin{cases} 1, & if\ b + \sum_{i=0}^{n} w_i\, x_i > 0 \\ 0, & otherwise \end{cases}$$

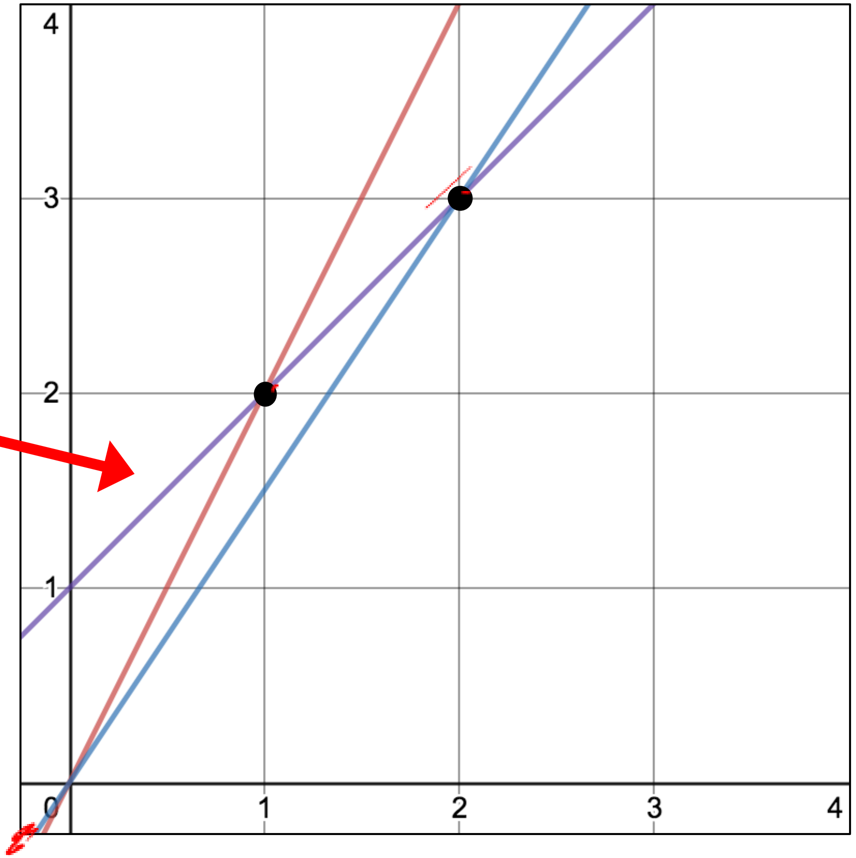


Bias: Geometric Explanation

w

- the bias is essentially the **b** term in **y = mx+b**

only the line with bias can fit the data



$$f(x) = 2x$$

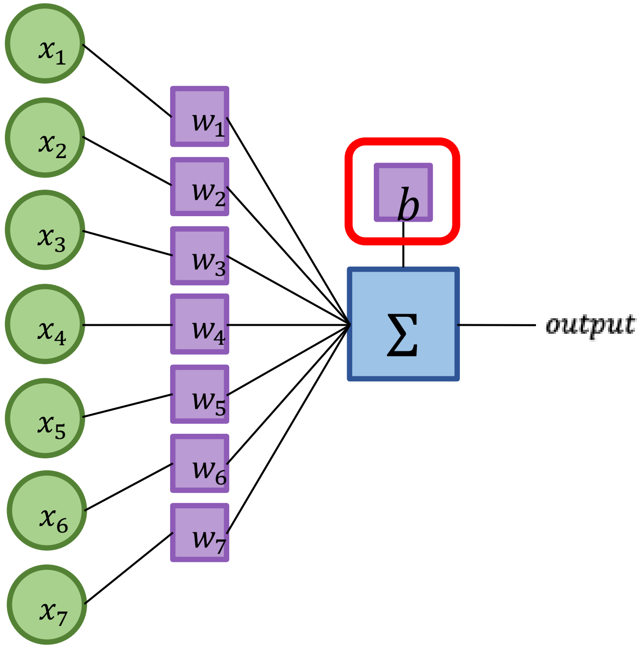
$$f(x) = \frac{3}{2}x$$

$$f(x) = x + 1$$

Bias: Conceptual Explanation

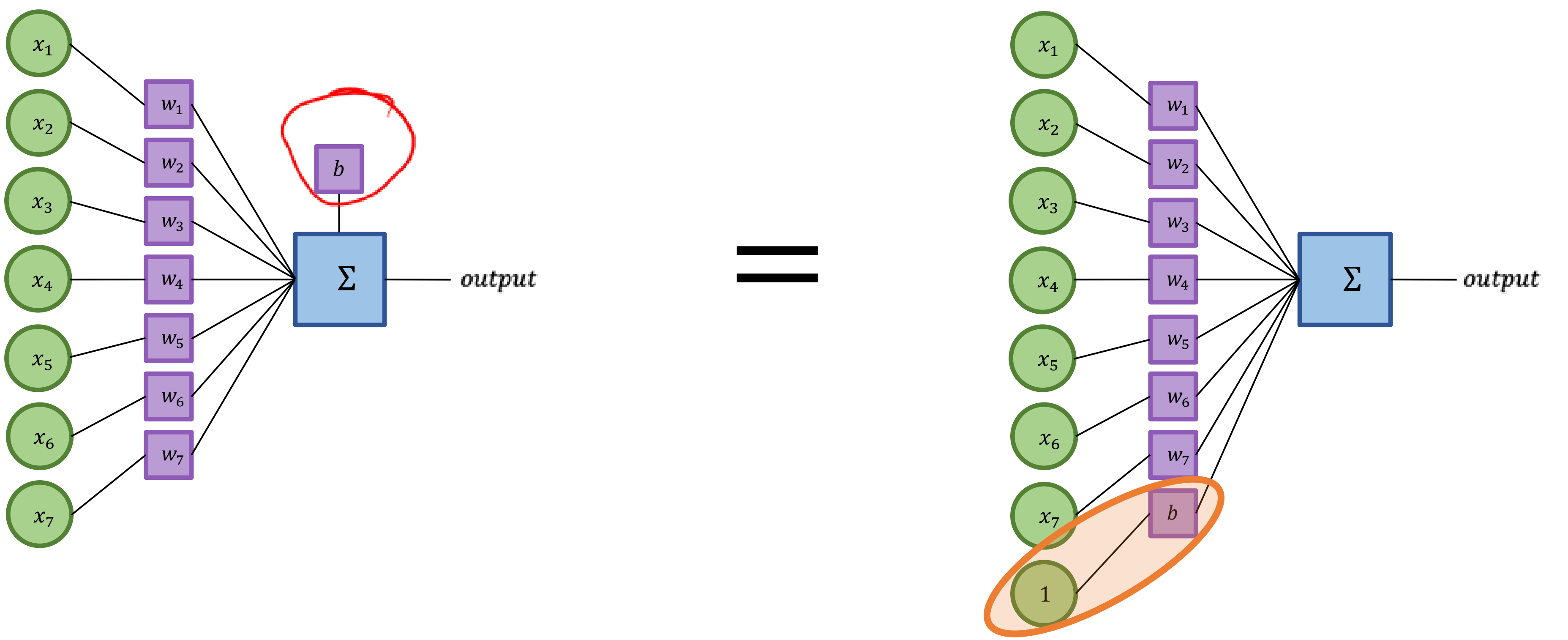
- ***Bias*** — the *a priori* likelihood of the positive class
  - Ensures that even if all inputs are 0, there will be some result value
  - Just because all inputs are 0, it does not mean there are no 1's in the world
  - Maybe there just happen to be more, say, 0's than 1's

$$f_{\Phi}(x) = \begin{cases} 1, & \text{if } b + \sum_{i=0}^{n} w_i x_i > 0 \\ 0, & \text{otherwise} \end{cases}$$



Bias as special type of weight

- Another way to think of bias is to represent it as an extra weight for an input/feature that is always 1



Bias as special type of weight

- Another way to think of bias is to represent it as an extra weight for an input/feature that is always 1

$$[x_0, x_1, x_2, \dots x_n] \cdot [w_0, w_1, w_2, \dots w_n] + b$$

$$= [x_0, x_1, x_2, \dots x_n, 1] \cdot [w_0, w_1, w_2, \dots w_n, b]$$

Recall

$\mathbf{a}=[a_1, a_2, \ \ldots, a_n]$ and $\mathbf{b}=[b_1, b_2, \ \ldots, b_n]$ with vector space n,

the dot product is

$$\mathbf{a}\cdot\mathbf{b}=\sum_{i=1}^{n} a_i b_i = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

Courtesy: MathType

Simplifying some notation...

- Recall: the dot product of two vectors of length n is $\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i$
- We can rewrite the perceptron function accordingly:

$$f_{\Phi}(x) = \begin{cases} 1, & \mathit{if}\ b + \sum_{i=0}^{n} w_i\, x_i > 0 \\ 0, & \mathit{otherwise} \end{cases}$$

$$f_{\Phi}(x) = \begin{cases} 1, & \mathit{if}\ b + \mathbf{w} \cdot \mathbf{x} > 0 \\ 0, & \mathit{otherwise} \end{cases}$$

Includes bias

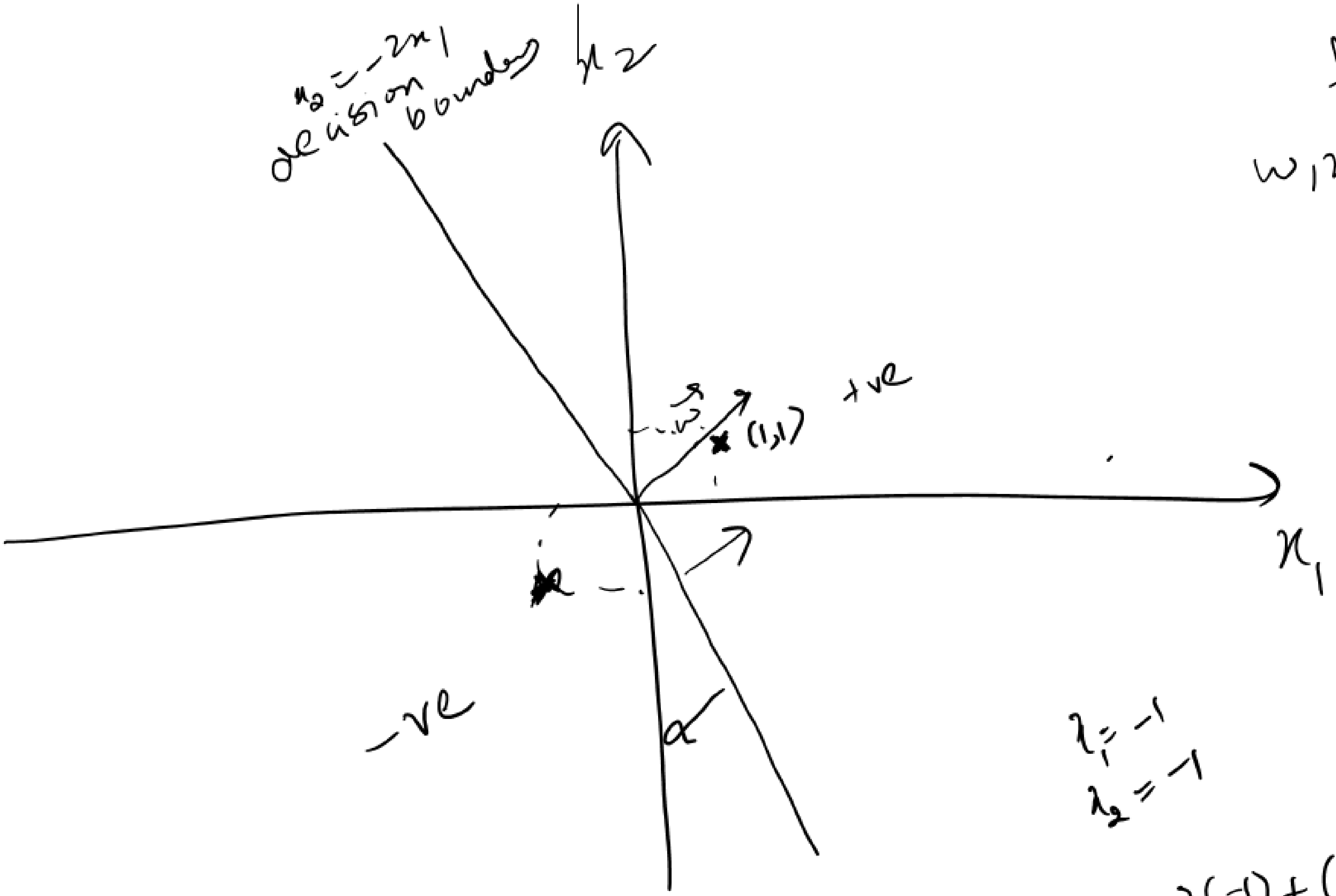
$w \cdot x > 0$

$f\phi(x) = \begin{cases} 1 \end{cases}$

Any questions?

- $b + \mathbf{w} \cdot \mathbf{x}$ is known as a ***linear unit***

Geometric interpretation of a perceptron



$f(x) = w \cdot x = 0$

$w_1x_1 + w_2x_2$

$w_1 = 2$

$w_2 = 1$

$w_1x_1 + w_2x_2 = 0$

$2x_1 + x_2 = 0$

$x_2 = -2x_1$

$x_1 = 1$

$x_2 = 1$

$2(1)+(1) = 3 > 0$

classify => 1

$x_1 = -1$

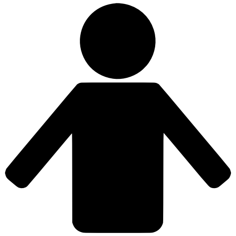
$x_2 = -1$

$2(-1)+(-1) = -3 < 0$

classify = -1

Our problem: Heart disease classification

$w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4$



Input: $\mathbb{X}$

Target: $\mathbb{Y}$

age,cholesterol,resting_bp,max_hr
45,180,120,170
48,185,118,172
50,190,125,165
52,200,130,160
55,210,135,158
57,220,140,150
58,235,142,152
60,240,145,148
62,250,150,140
65,260,155,138

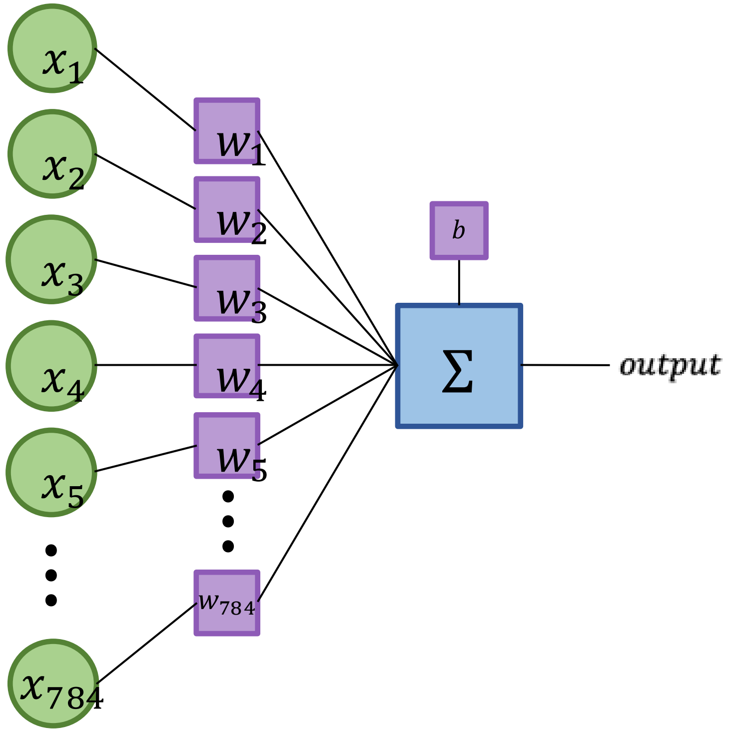
Function: f

$\boxed{f(\mathbb{X}) \rightarrow \mathbb{Y}}$

heart_disease
0
0
0
0
0
1
1
1
1
1

A Linear Perceptron for classifying heart disease

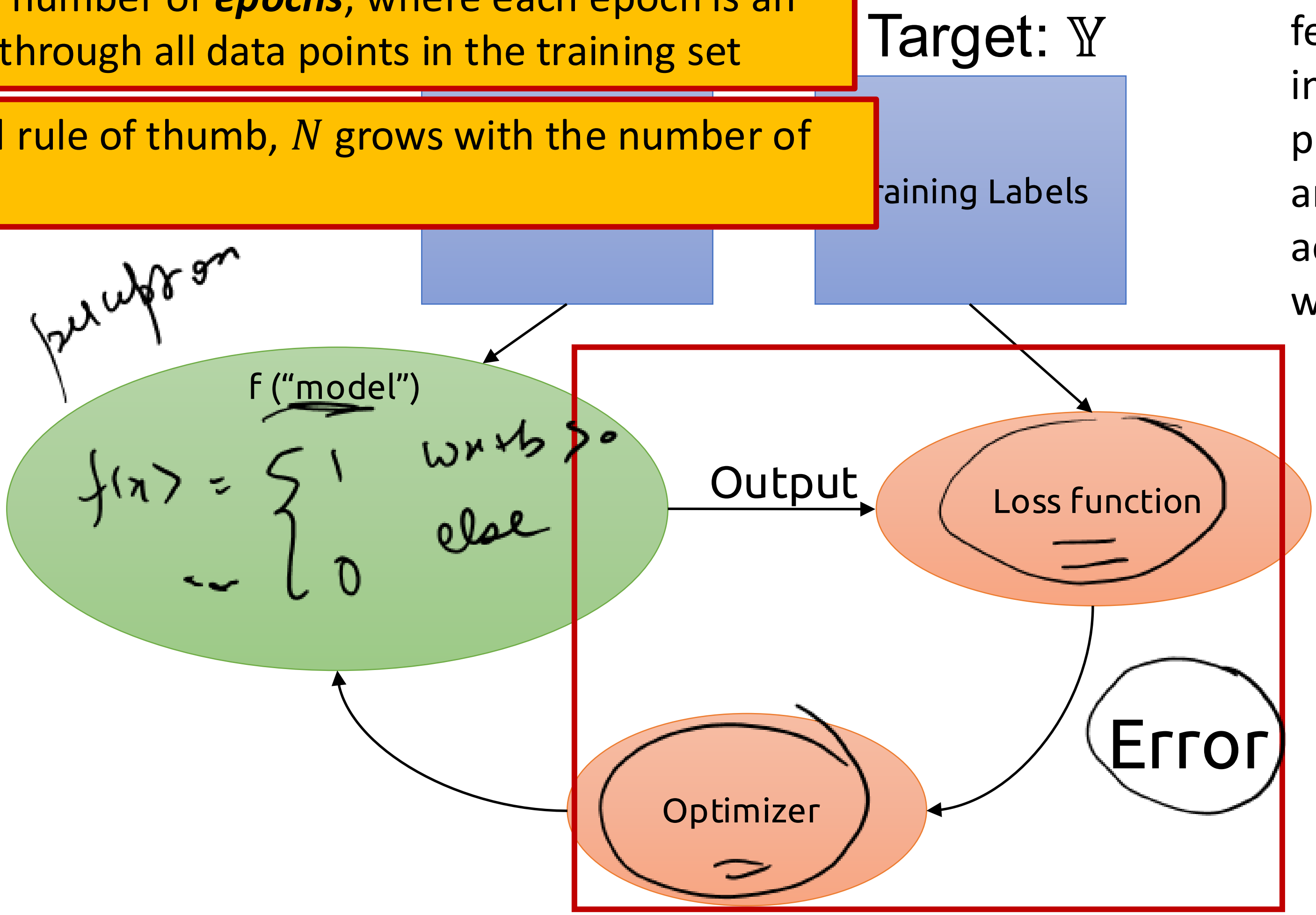
- $Inputs\,[x_1, x_2, \ldots x_n]$ are all positive
  - $n = 4$ (features)
- $output$ is either 0 or 1
  - 0 → no heart disease
  - 1 → has heart disease



Training a perceptron (ML training)

N is known as the number of ***epochs***, where each epoch is an iteration of going through all data points in the training set

As a general rule of thumb, N grows with the number of parameters



0. set the parameters $\Phi = \{w \cup b\}$ to 0

1. Iterate over training set several times, feeding in each training example into the model, producing an output, and adjusting the parameters according to whether that output was right or wrong

hyperparameters

2. Stop once we either (a) get every training example right or (b) after N iterations, a number set by the programmer.

The Perceptron Learning Algorithm

1. set w's to 0.
2. for N iterations, or until the weights do not change:
   a) for each training example $\mathbf{x}^n$ with label y^k
      i. if $y^k - f(\mathbf{x}^k) = 0$ continue
      ii. else for all weights w_i, $\Delta w_i = \left(y^k - f(\mathbf{x}^k)\right) x_i^k$

$w^T x + b = 1/0$

- b = bias
- w = weights
- N = maximum number of training iterations
- $\mathbf{x}^k$ = $\mathrm{k^{th}}$ training example

- y^k = label for the $\mathrm{k^{th}}$ example
- w_i = weight for the $\mathrm{i^{th}}$ input where $i \le n$
- n = number of features
- x_i^k = $\mathrm{i^{th}}$ input of the example where $i \le n$

The Perceptron Learning Algorithm

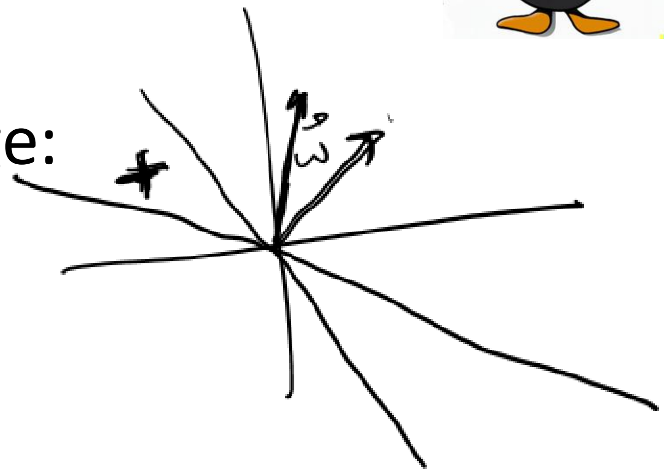
1. set w's to 0.
2. for N iterations, or until the weights do not change:
   a) for each training example $\mathbf{x}^k$ with label y^k
      i. **if $\boldsymbol{y^k - f(\mathbf{x}^k) = 0}$ continue**
      ii. else for all weights w_i, $\Delta w_i = \left(y^k - f(\mathbf{x}^k)\right) x_i^k$

- If the output of our model matches the label, we continue
- If the correct label is 1, and our output is 1, $1 - 1 = 0$
- If the correct label is 0, and our output is 0, $0 - 0 = 0$

Any questions?

The Perceptron Learning Algorithm

1. set w's to 0.
2. for N iterations, or until the weights do not change:
   a) for each training example $\mathbf{x}^k$ with label y^k
      i. if $y^k - f(\mathbf{x}^k) = 0$ continue
      ii. **else for all weights w_i, $\Delta w_i = \left(y^k - f(\mathbf{x}^k)\right) x_i^k$**



- If our label y^k is a 1, and our model's output is a 0, we update the i^{th} weight by:
  - $(1 - 0) \cdot x_i^k = x_i^k$
  - Output was 0 and should have been 1, so make the output more positive
- If our label y^k is a 0, and our model's output is a 1, we update the i^{th} weight by:
  - $(0 - 1) \cdot x_i^k = -x_i^k$
  - Output was 1 and should have been 0, so make the output more negative

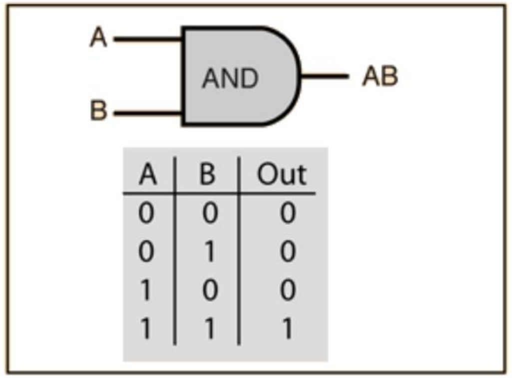
Class activity

[Linear Perceptron Demo]

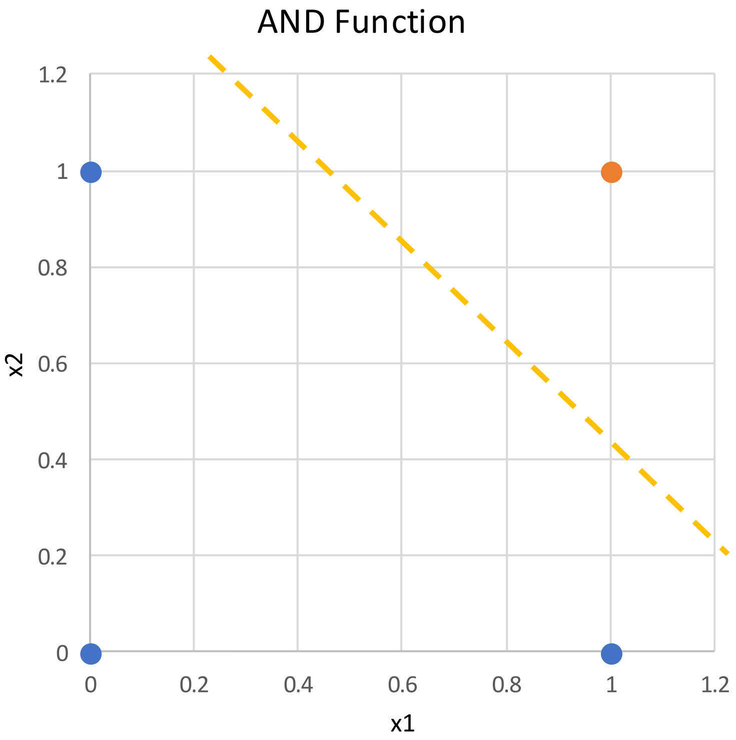
AND Function

The Perceptron Convergence Theorem:

Guarantees that if a dataset is linearly separable, the perceptron algorithm will find a separating hyperplane in a finite number of updates.



AND Gate

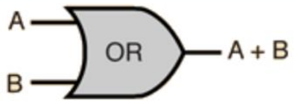


Output = 0

Output = 1

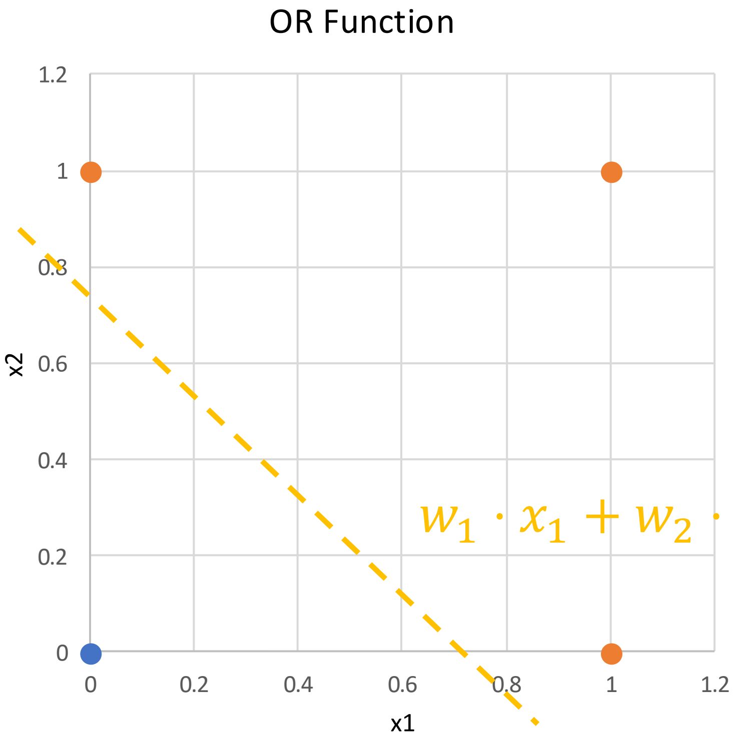
$w_1 \cdot x_1 + w_2 \cdot x_2 + b > 0$?

OR Function



A	B	Out
0	0	0
0	1	1
1	0	1
1	1	1

OR Gate



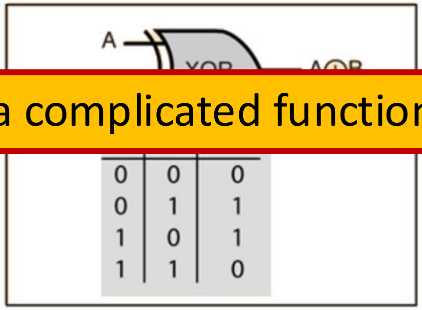
Output = 0

Output = 1

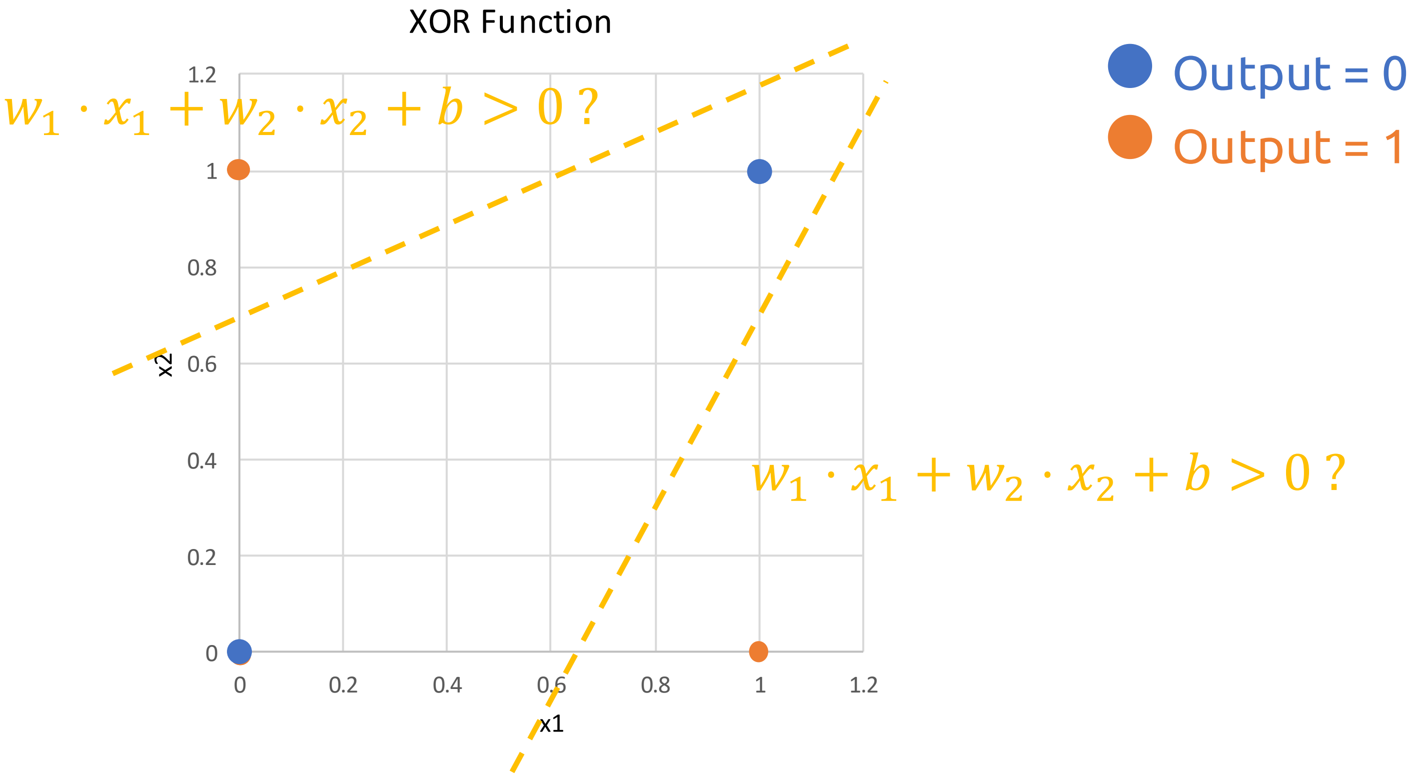
$$w_1 \cdot x_1 + w_2 \cdot x_2 + b > 0 \,?$$

XOR Function

Complicated data would need a complicated function!



XOR Gate



Need ***two*** linear decision boundaries to represent this function...

Evaluation

Calculating accuracy of classification

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

Do you know of any metric to evaluate classification?

Homework reading:
[Chapter 4 in the CIML book]

Recap

(1) Machine Learning problem – Heart Disease Classification

(2) Linear Perceptron

(3) Parameters – weights and biases

(4) Perceptron Learning Algorithm

(5) Class Activity: Perceptron in action

(6) Evaluating classification performance

Wrap up

What was the clearest point today?

What was the muddiest point today?