Classification

Ahmet Hamza Emra

2020

Outline

What is supervised learning?

- Learning to map input x into output y

What is supervised learning?

- Learning to map input x into output y
- Function approximation

What is supervised learning?

- Learning to map input x into output y
- Function approximation

$$f(x) = y$$

What is supervised learning?

- Learning to map input x into output y
- Function approximation

$$f(x) = y$$

- y can be continuous or distinct class of numbers
- regressions vs classifications

Classification problems

Finding the function $f(x) = y$ where y belongs to distinct class

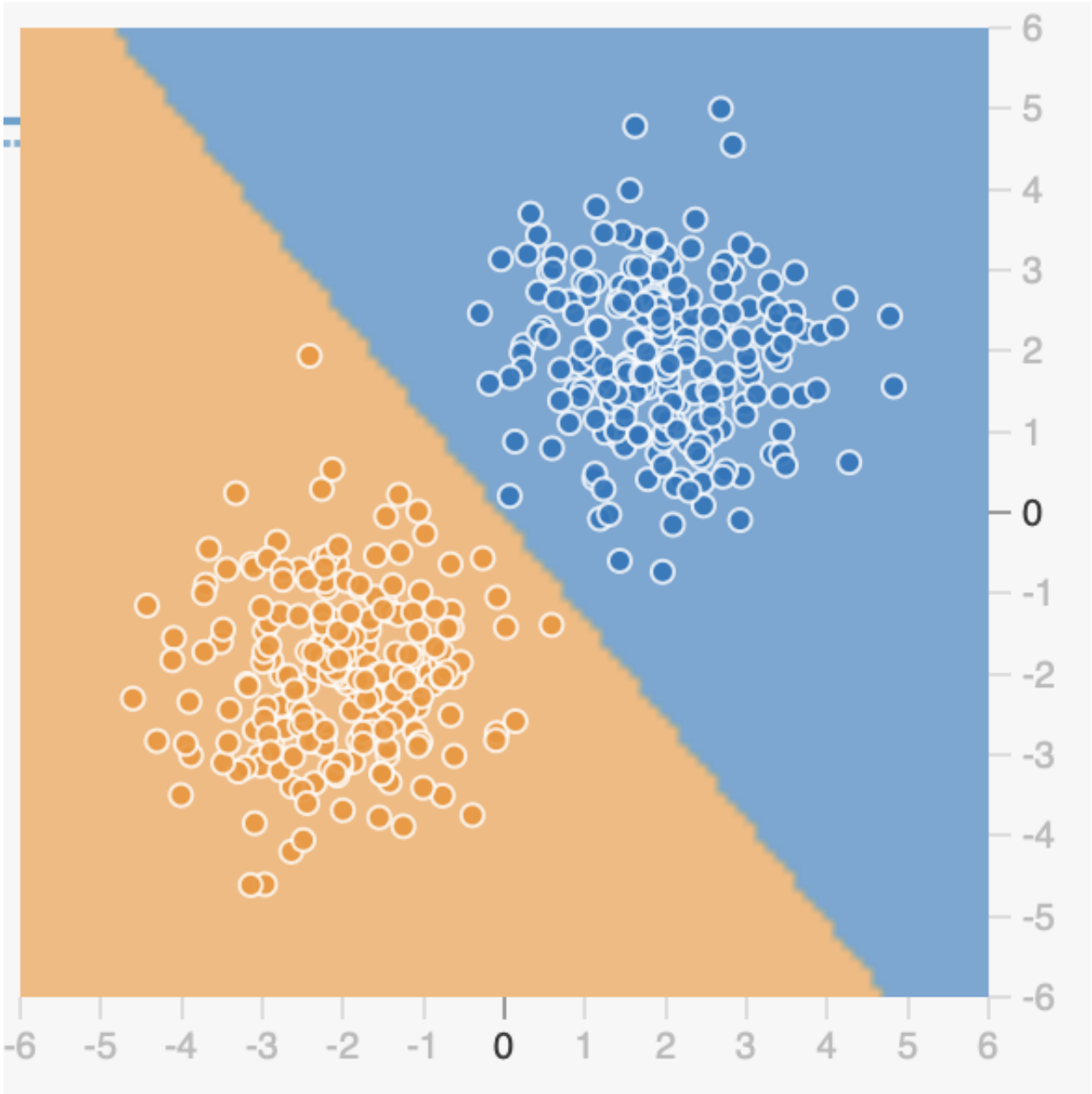
Classification problems

Finding the function $f(x) = y$ where y belongs to distinct class

Examples

- Predicting digit from hand written images
- Predicting size of the shirt(M - L - XL)
- Sentiment analysis
- Someone is cancer or not

Decision Boundary



Binary Classification Problem

How do we define $f(x)$?

Binary Classification Problem

How do we define $f(x)$?

$$y = \begin{cases} 1 & a.x + b > threshold \\ 0 & a.x + b < threshold \end{cases}$$

Binary Classification Problem

How do we define $f(x)$?

$$y = \begin{cases} 1 & a.x + b > threshold \\ 0 & a.x + b < threshold \end{cases}$$

Find best a and b that satisfies $f(x) = y$

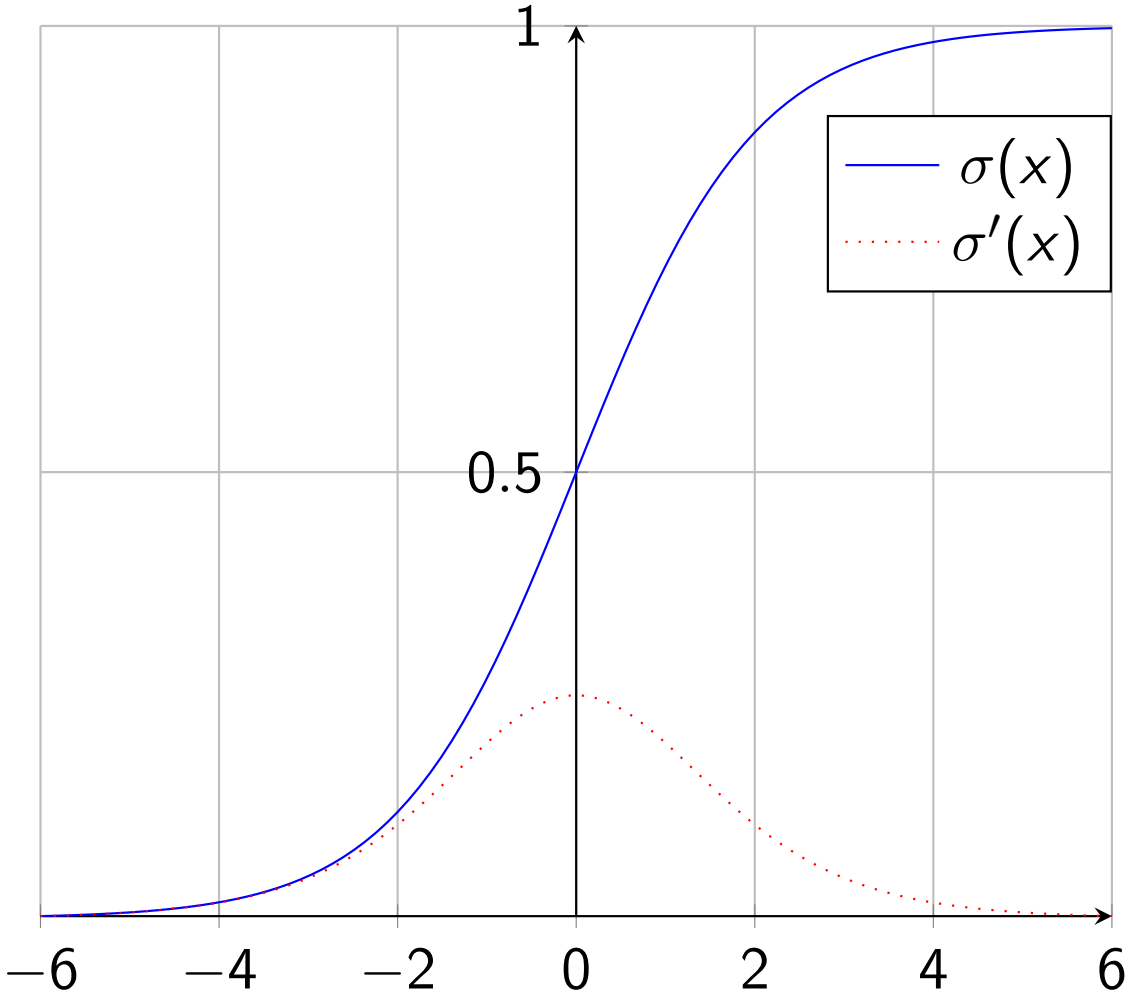
Logistic Regression

$$y = \begin{cases} 1 & \sigma(a.x + b) >= 0.5 \\ 0 & \sigma(a.x + b) < 0.5 \end{cases}$$

Sigmoid Function

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

Sigmoid Function



Cost Function

Error (cost) function

$$J(w, b) = -\frac{1}{m}\sum_{i=1}^{m}(y.log(y') + (1 - y).(1 - log(y'))$$

Learning

We learn with the same principal.

Learning

We learn with the same principal.
Get gradient, make changes.

Gradient Decent

Reminder

- Derivative of function is rate of change
- Basically it returns if the function is increasing or decreasing in given point.

Gradient Decent

Reminder

- Derivative of function is rate of change
- Basically it returns if the function is increasing or decreasing in given point.

By using derivative of the cost function one can determine if he/she should increase the weights and biases or decrease

Gradient Decent

Reminder

- Derivative of function is rate of change
- Basically it returns if the function is increasing or decreasing in given point.

By using derivative of the cost function one can determine if he/she should increase the weights and biases or decrease

$$W = W - \alpha.\Delta W$$

$$b = b - \alpha.\Delta b$$

Summary

1. pick random W and b
2. make predictions y‘
3. calculate $J(W, b)$
4. use gradient decent to make update
5. repeat until J settles

Implementation

Python Implementation

Overfitting vs Underfitting