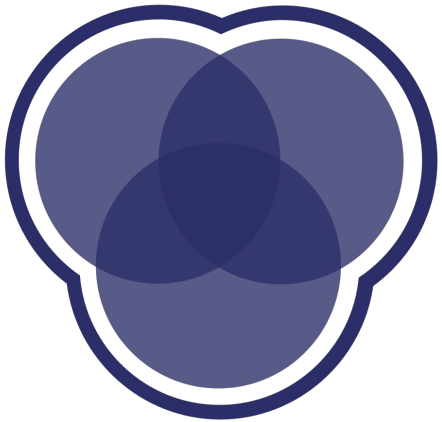
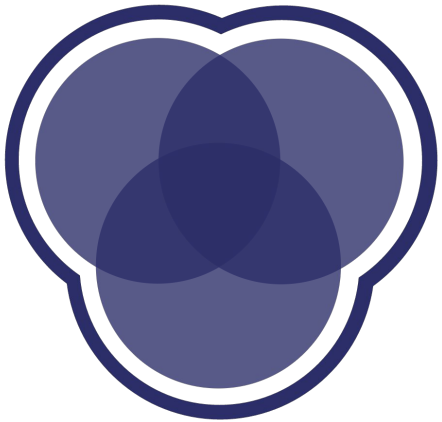
Intro to Machine Learning

Demo



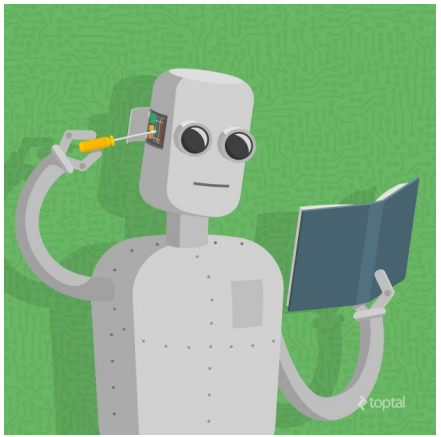
Machine Learning



Machine Learning

According to Wikipedia…

“Machine Learning is a subfield of computer science that gives computers the ability to learn without being explicitly programmed.”



ML Setup

Hypothesis: Some speculative relationship between the input space and output space

Input Space: Variable or set of variables(data)

Output Space: Target variable to estimate

Supervised vs Unsupervised

Supervised learning problems...

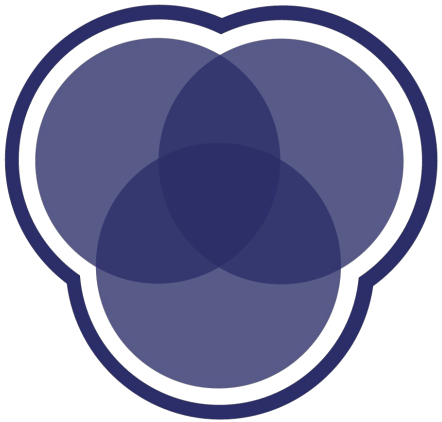
- Known target variable info
- Positive / Negative examples

Unsupervised learning problems...

- Unknown target variables
- Difficult to validate

- **Supervised learning:**
  given $(x_1, y_1), \ldots, (x_n, y_n)$, learn $f(x) = y$
- **Unsupervised learning:**
  given $x_1, \ldots, x_n$, learn patterns or structure
- **Online learning:** for $i = 1, \ldots, n$,
  given x_i, predict and observe y_i, learn $f(x) = y$
- **Active learning:** for $i = 1, \ldots, n$,
  choose x_i, predict and observe y_i, learn $f(x) = y$
- **Reinforcement learning:** for $i = 1, \ldots, n$,
  choose x_i, predict y_i, observe reward r_i, learn $f(x) = y$

Supervised Learning

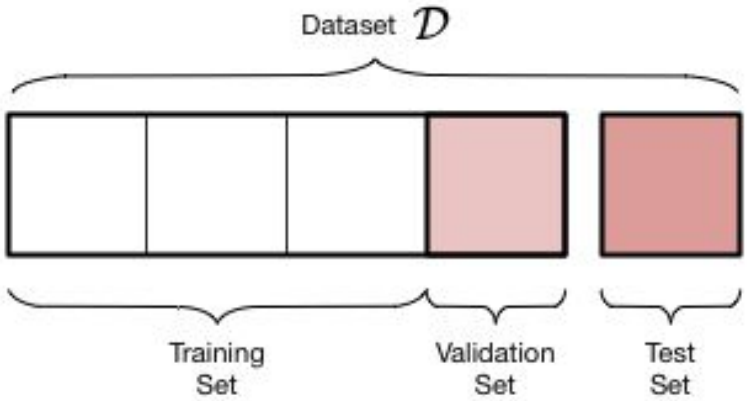


Setup

- Training / Validation split
- Feature variable(s)
- Target variable
- Train and Test

Validation Set

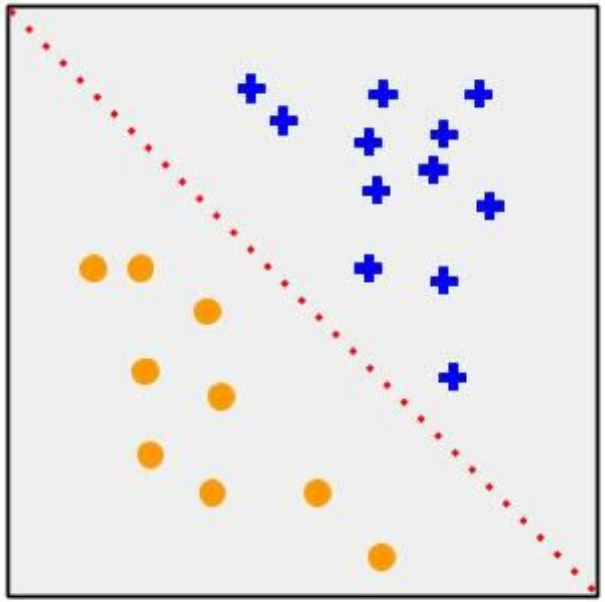
- Split data into two sets
- Train model on one and validate on another
- Advantages / Disadvantages?



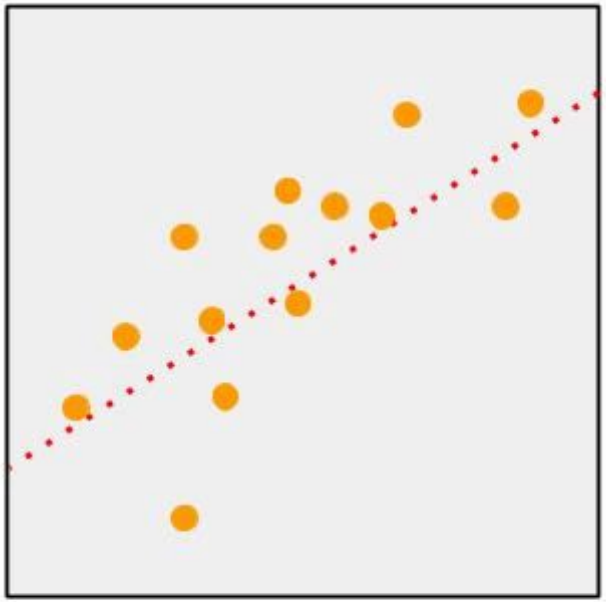
Output Space Properties

- **Continuous -** e.g. temperature, height, probability
- **Discrete -** e.g. car brands, race, Pokémon type, diagnosis

Regression vs Classification



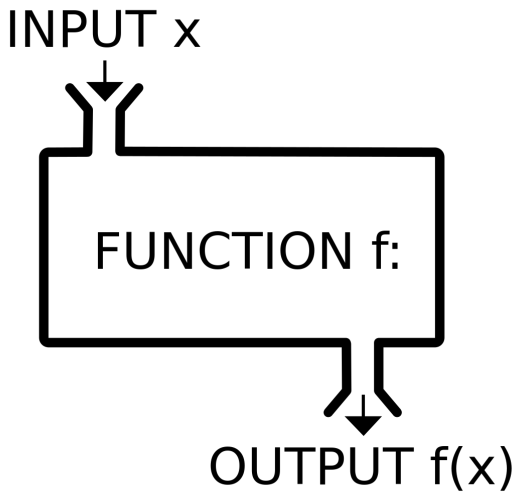
Classification



Regression

What is Learned

Function

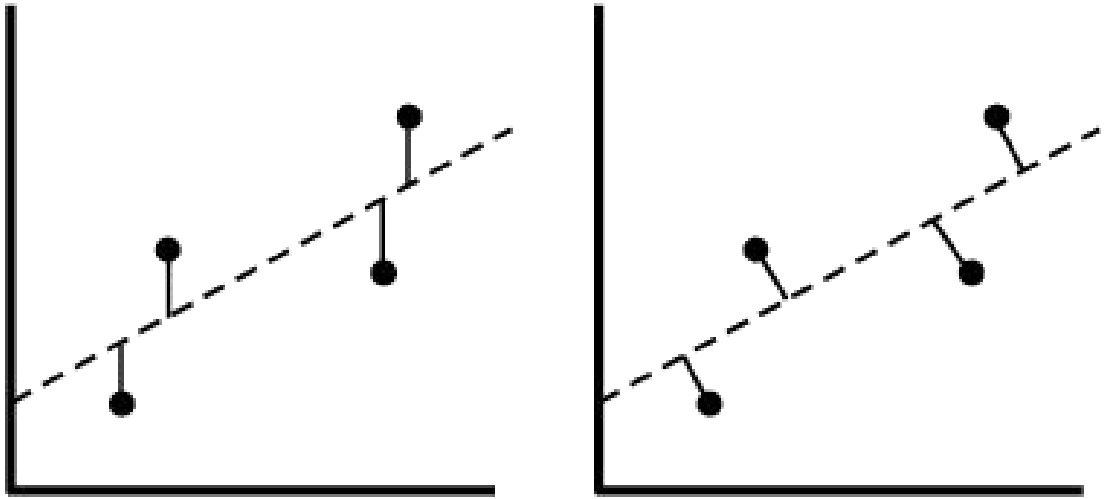


Weighted Sum

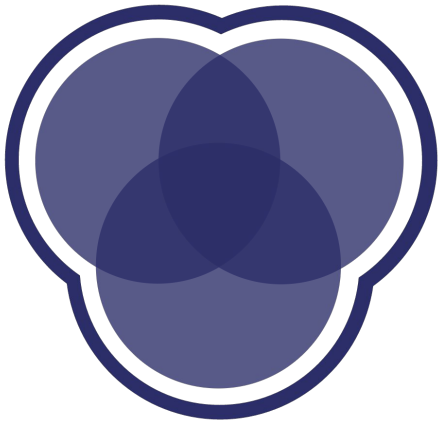
$$y = B_0 + B_1 x_1 + \ldots + B_p x_p$$

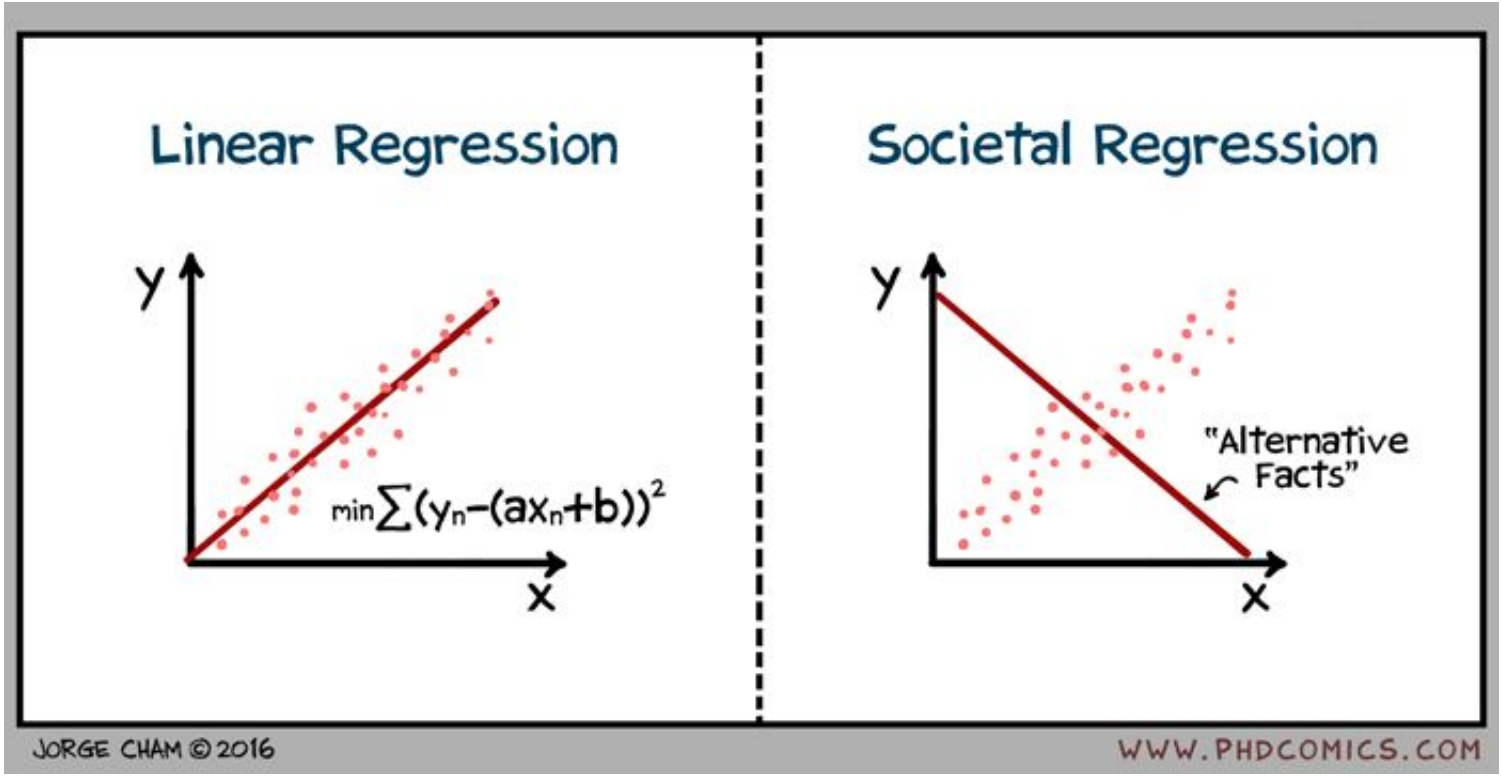
Objective function

- All ML problems are optimization problems
  - Format: Minimize/Maximize Obj in terms of x.
    - Subject to set of constraints
- Objective functions represent assumptions
- Value of objective is an estimation of error

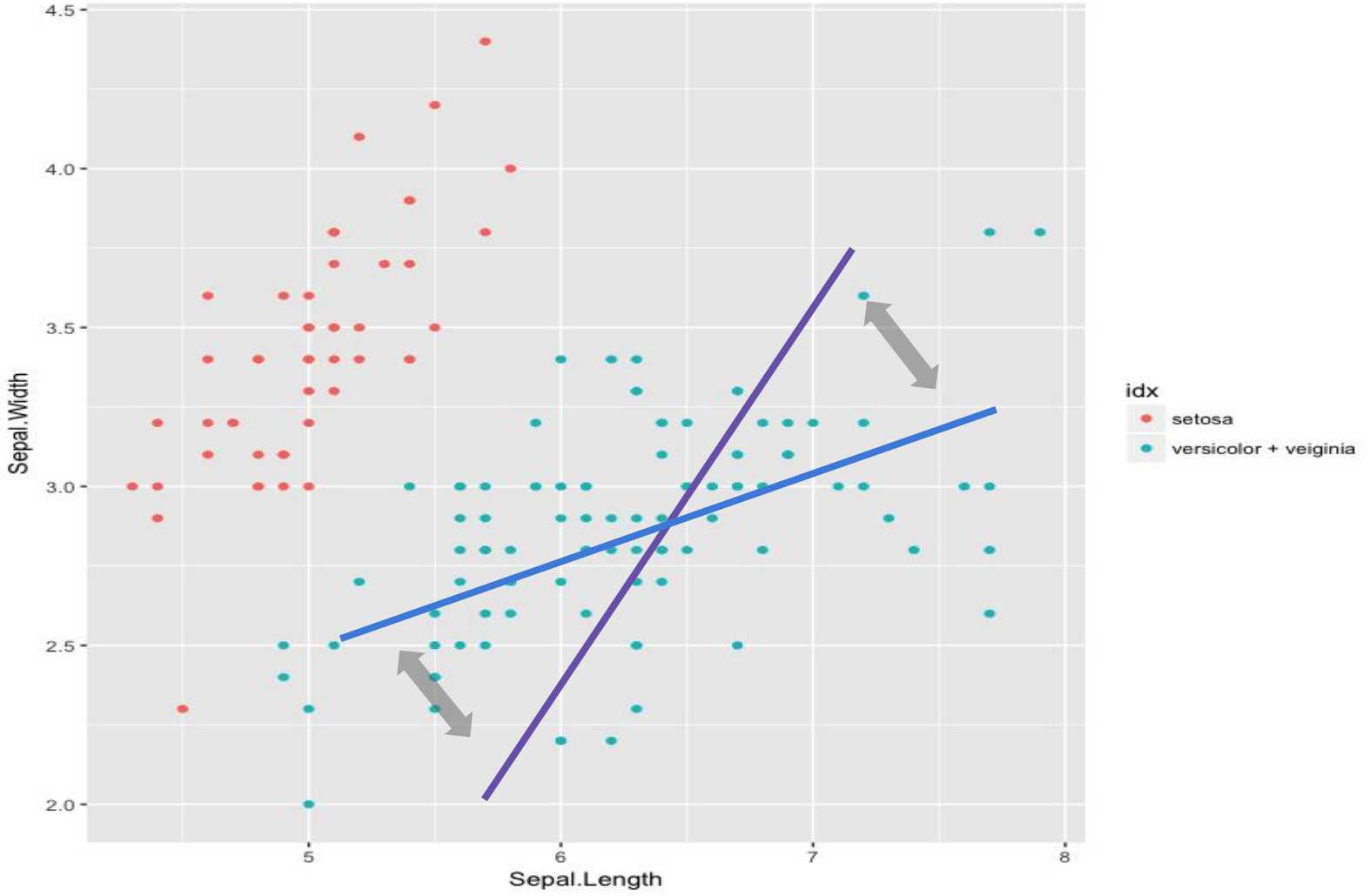

vertical offsets
perpendicular offsets

Linear Regression




Linear Regression
y
x
$\min\sum(y_n-(ax_n+b))^2$
Societal Regression
y
x
"Alternative Facts"
JORGE CHAM © 2016
WWW.PHDCOMICS.COM

Source


4.5
4.0
3.5
3.0
2.5
2.0
Sepal.Width
5
6
7
8
Sepal.Length
idx
setosa
versicolor + veiginia

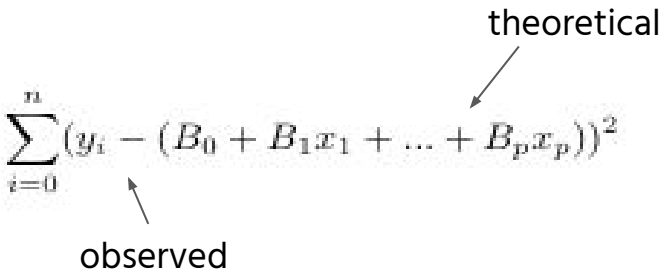
Linear Regression

$$y = B_0 + B_1 x_1 + \ldots + B_p x_p + \varepsilon$$

What are the assumptions?

- Linear relationship
  - B, the coefficient vector, does not depend on x
- There is an unremovable noise
- This noise is normally distributed about the line

Objective: Least Squares Error (L2)



What does this minimize?

Why this form?

Coming Up

Your problem set: Continue working on Project Part A

Next week: Introduction to classification

See you then!