Machine Learning from Data

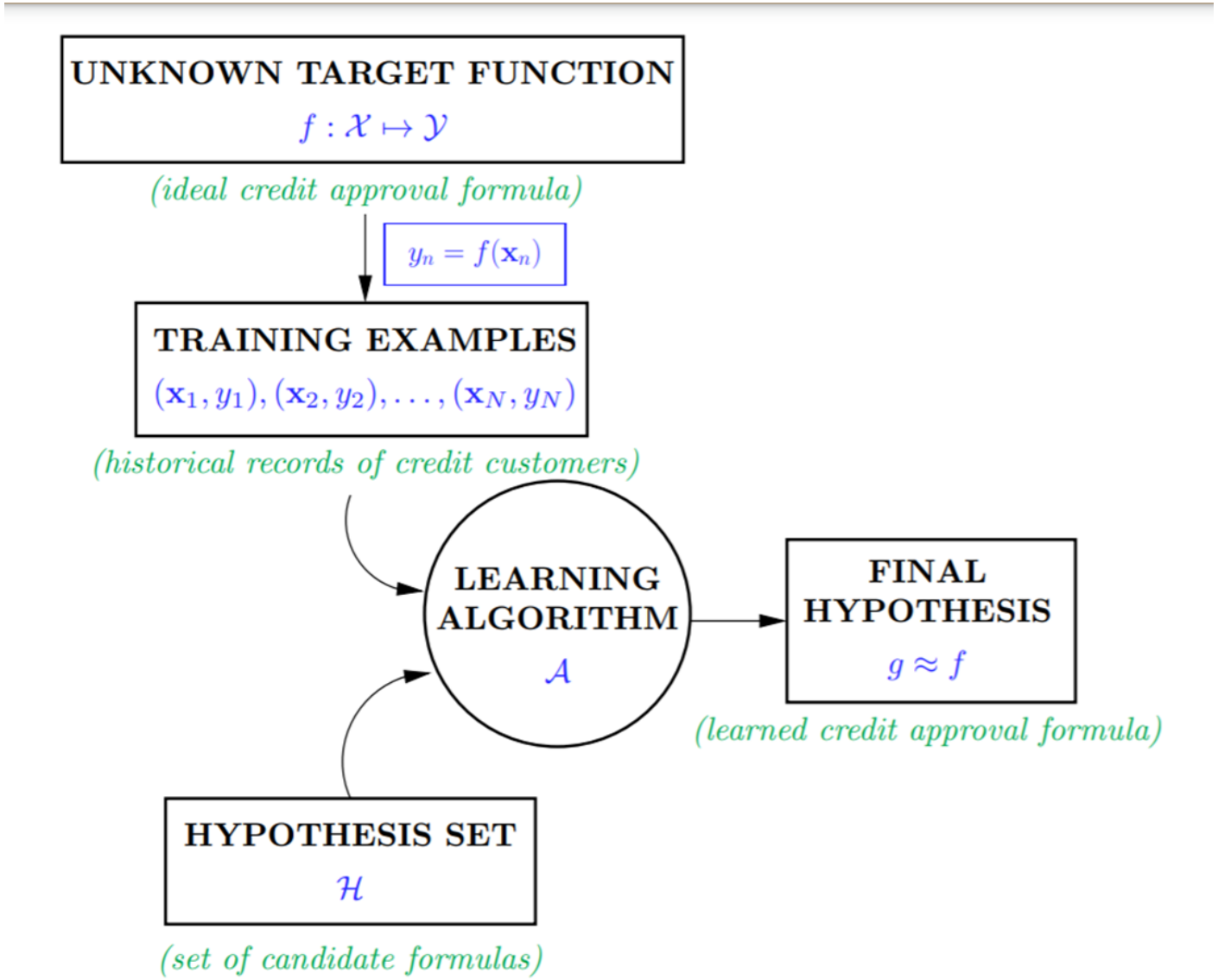
Lecture 2: Spring 2021

Today's Lecture

- The Perceptron
- Learning Set-Up
- The PLA
- Other Views of Learning

In the Previous Lecture

- Formalized Components of Learning
- Learning Process



Formalize Components of Learning:

$$\textit{input } \mathbf{x} \in \mathbb{R}^d = \mathcal{X}.$$

$$\textit{output } y \in \{-1, +1\} = \mathcal{Y}.$$

$$\textit{target function } f : \mathcal{X} \mapsto \mathcal{Y}.$$

(The target f is *unknown*.)

$$\textit{data set } \mathcal{D} = (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N).$$

$(y_n = f(\mathbf{x}_n).)$

- **Input**: Salary, debt, years
  Output: Approve or not
  Target function: Relationship between X and Y
- Data on customers
- X, Y and D will be given by the learning problem.

A Simple Learning Model: The Perceptron

- Let us consider the Credit Example: We as Data Scientists/ML Practitioners want to approve/decline an incoming application.
- Formalize the problem:

Simple Learning Model

•

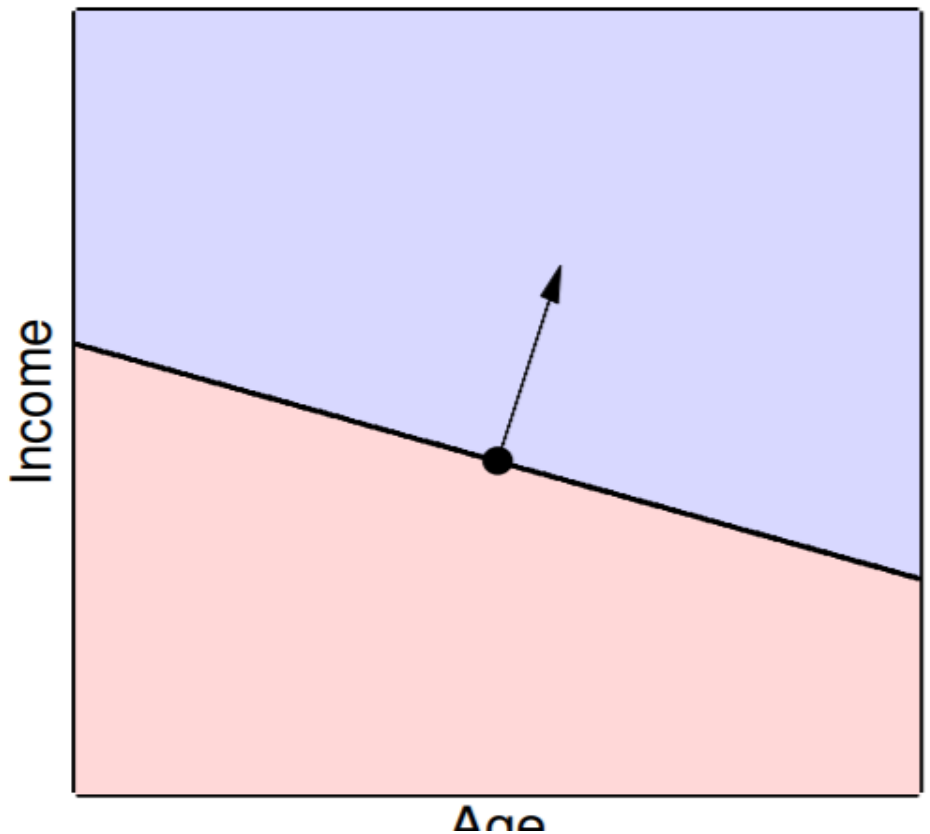
The Perceptron Hypothesis Set

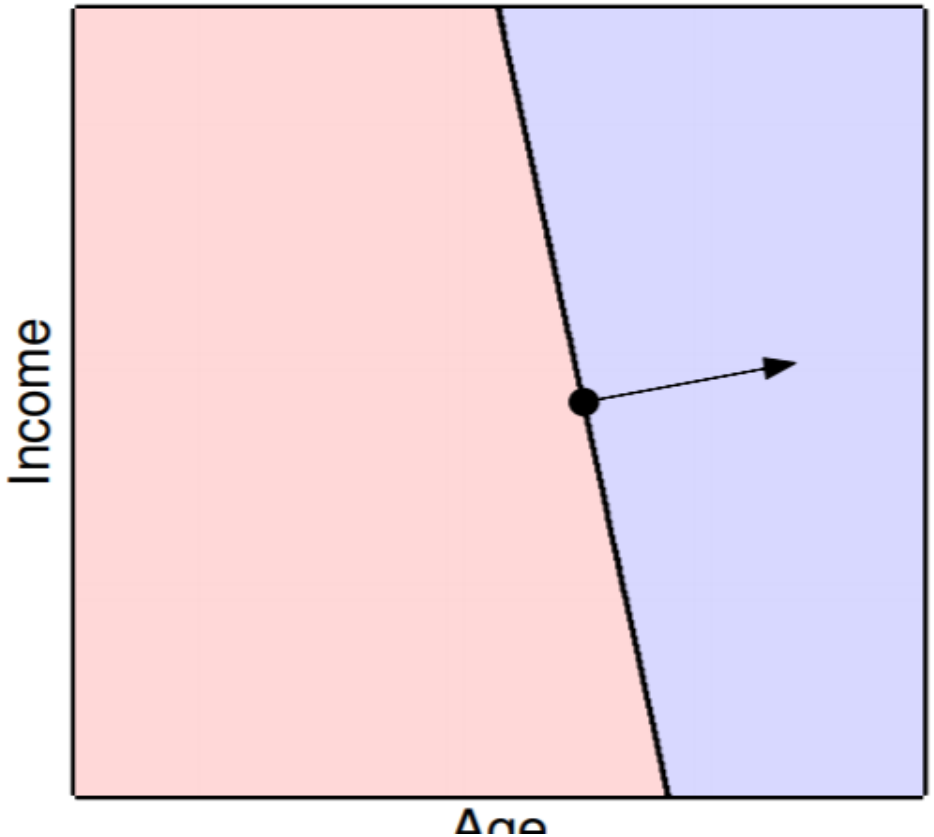
•

Geometry of the Perceptron

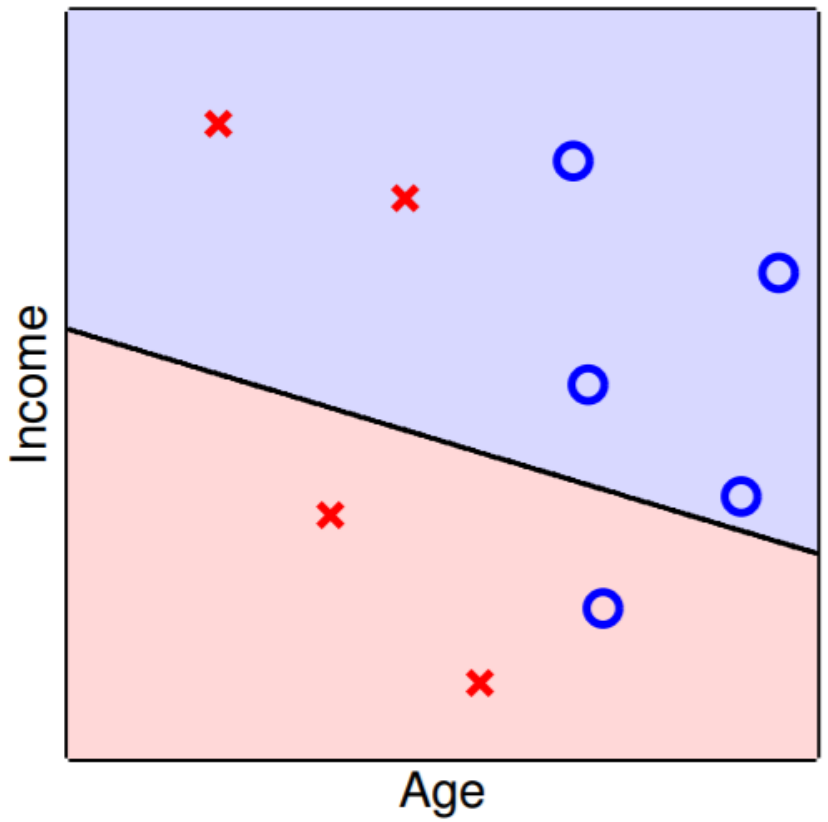
•

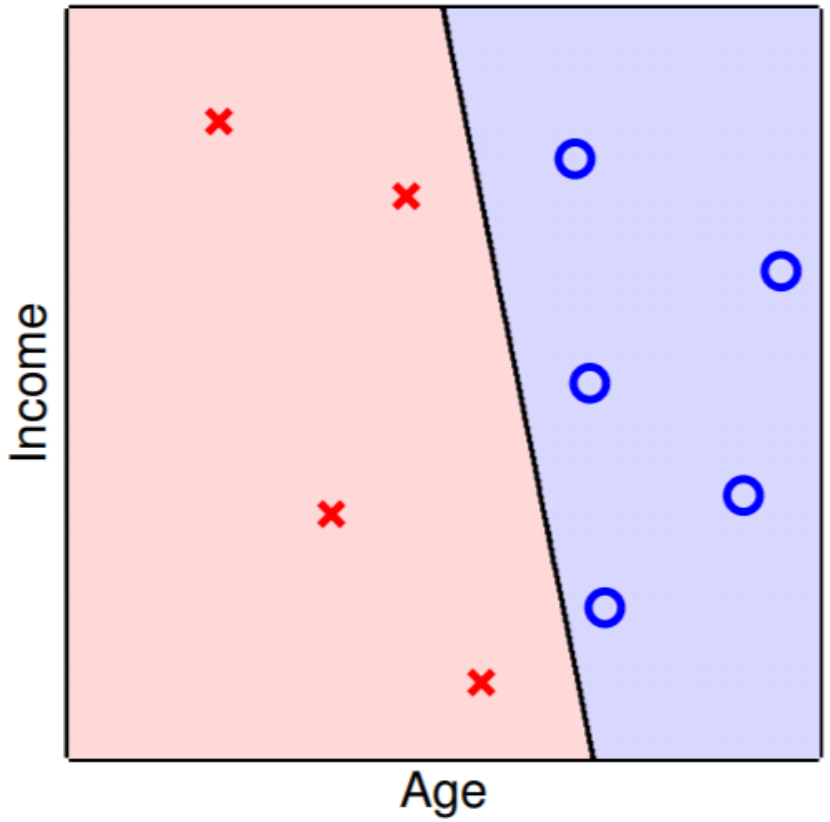
Geometry of Perceptron





Use Data to Pick the Perfect Line





How to Learn a Final Hypothesis *g* from *H*

- What do we want?
- What does that mean in terms of the data?
- How do we find this g in H(which is infinite)?
- How can we get started?

Perceptron Learning Algorithm

•

PLA Summarized

A simple iterative method.

1: $\mathbf{w}(1) = \mathbf{0}$
2: **for** iteration $t = 1, 2, 3, \ldots$
3: the weight vector is $\mathbf{w}(t)$.
4: From $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)$ pick any misclassified example.
5: Call the misclassified example $(\mathbf{x}_*, y_*)$,

$$\text{sign}\left(\mathbf{w}(t) \bullet \mathbf{x}_*\right) \neq y_*.$$

6: Update the weight:

$$\mathbf{w}(t+1) = \mathbf{w}(t) + y_* \mathbf{x}_*.$$

7: $t \leftarrow t + 1$

PLA Convergence

- Theorem: If the data can be fit by a linear separator, then after some finite number of steps, PLA will find one.

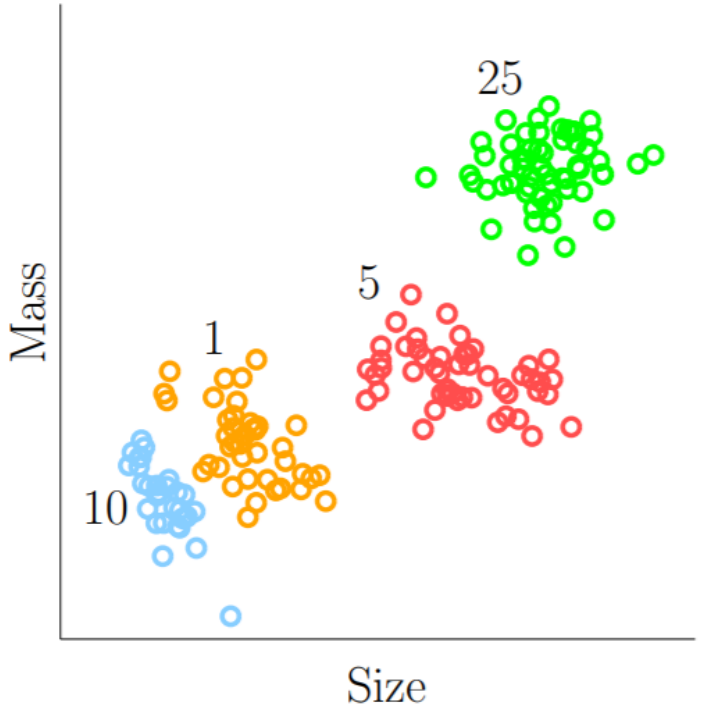
Other Views of Learning

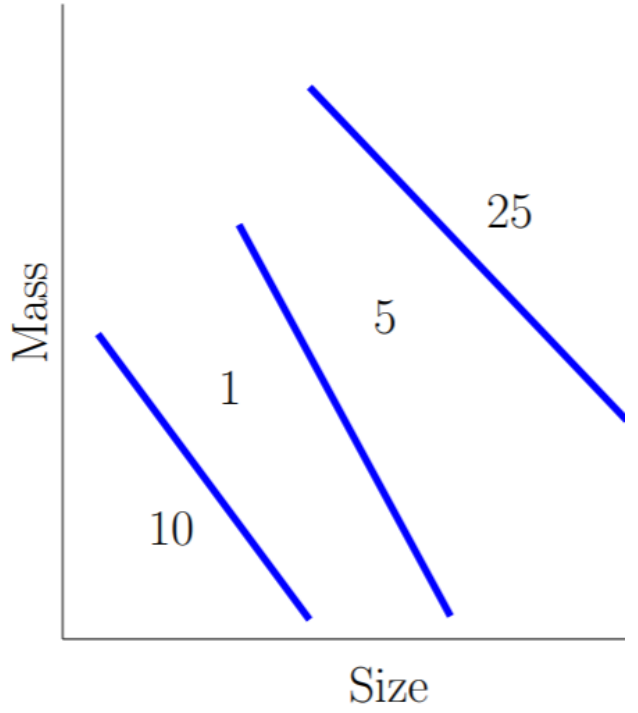
- Design: learning is from data, design is from specs and a model.
- Statistics
- Data Mining

Types of Learning

- Supervised
- Unsupervised
- Reinforcement Learning

Unsupervised Learning





Categorizing Coins