RUTGERS

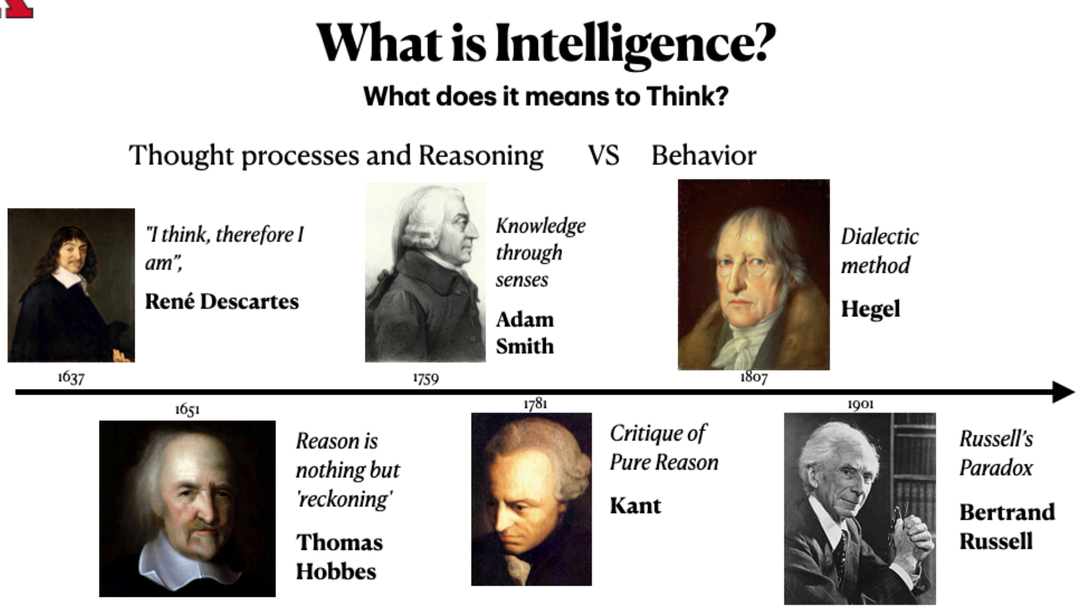
THE STATE UNIVERSITY
OF NEW JERSEY

Machine Learning

Introduction

Edgar Granados

Previously...



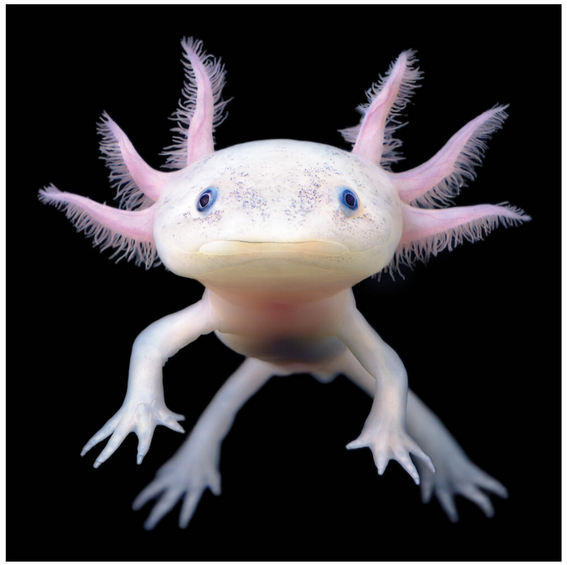
Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Intelligence

Assume...

- Intelligence is a **goal-directed adaptive** behavior

Adaptive Behavior



Learn

Definition

- To gain knowledge or understanding of or skill in by study, instruction, or experience

Learn

Definitions

To gain knowledge or understanding of or skill in by study, instruction, or experience

- **Kupfermann (1985):** Learning is the acquisition of knowledge about the world

Learn

Definitions

To gain knowledge or understanding of or skill in by study, instruction, or experience

- **Kupfermann (1985):** Learning is the acquisition of knowledge about the world
- **Shepherd (1988):** Learning is an adaptive change in behavior caused by experience

Learn

Definitions

To gain knowledge or understanding of or skill in by study, instruction, or experience

- **Kupfermann (1985):** Learning is the acquisition of knowledge about the world
- **Shepherd (1988):** Learning is an adaptive change in behavior caused by experience
- **Einstein?:** *If you can't explain it simply, you don't understand it well enough*

Machine Learning

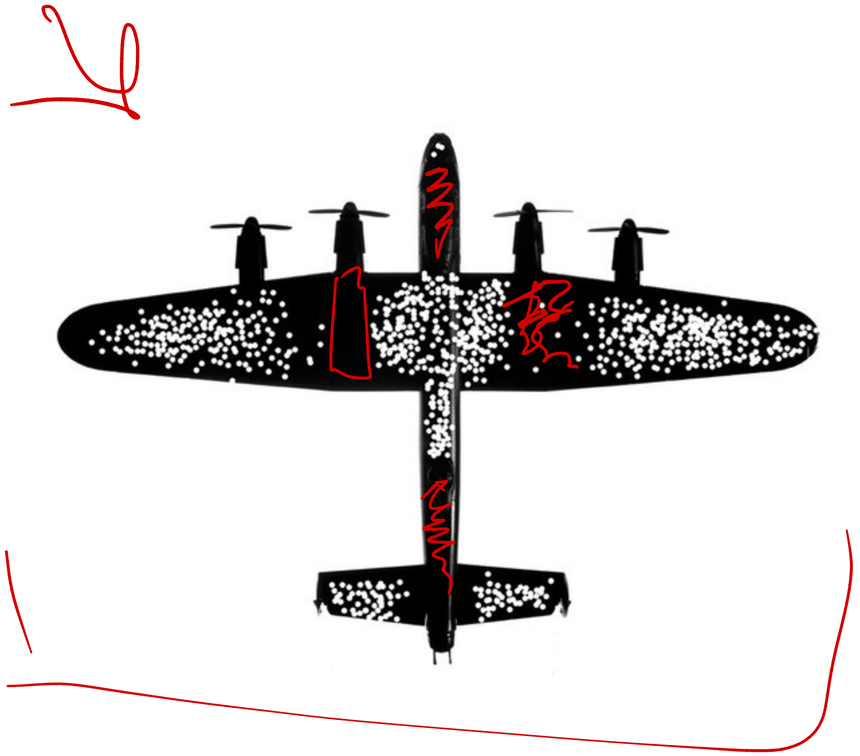
What is it?

- **Arthur Samuel:** a field of study that gives computers the ability to learn without being explicitly programmed.
- **Tom M Mitchell**: a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.
- **Ron Kohavi; Foster Provost (1988):** subfield of computer science that explores the study and construction of algorithms that can learn from and make predictions on data.

Machine Learning

What is it?

- Search for patterns in data
  - Predict

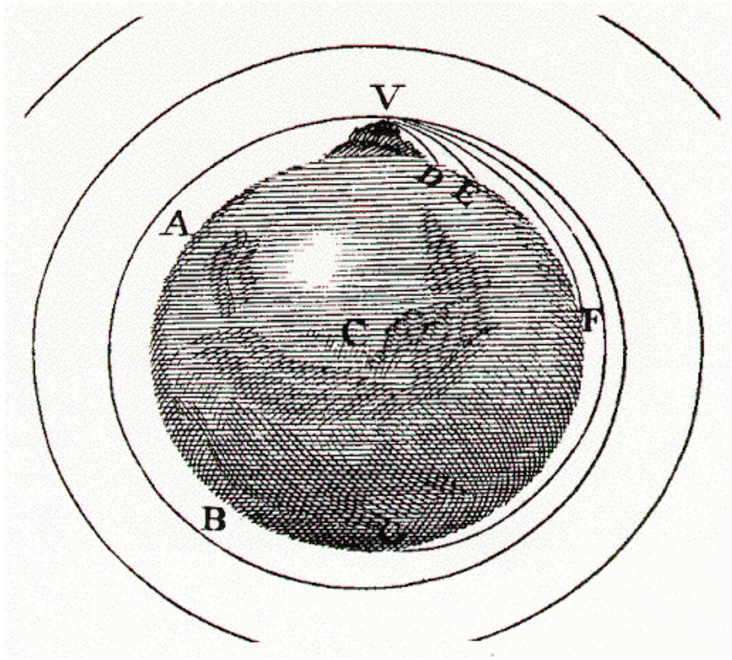


ML vs Statistics

- Both: Summarize data
- Statistics: Data analysis - hypothesis testing
- ML: Efficient algorithms —> Fast, low data requirements, make predictions

Empirical Inference

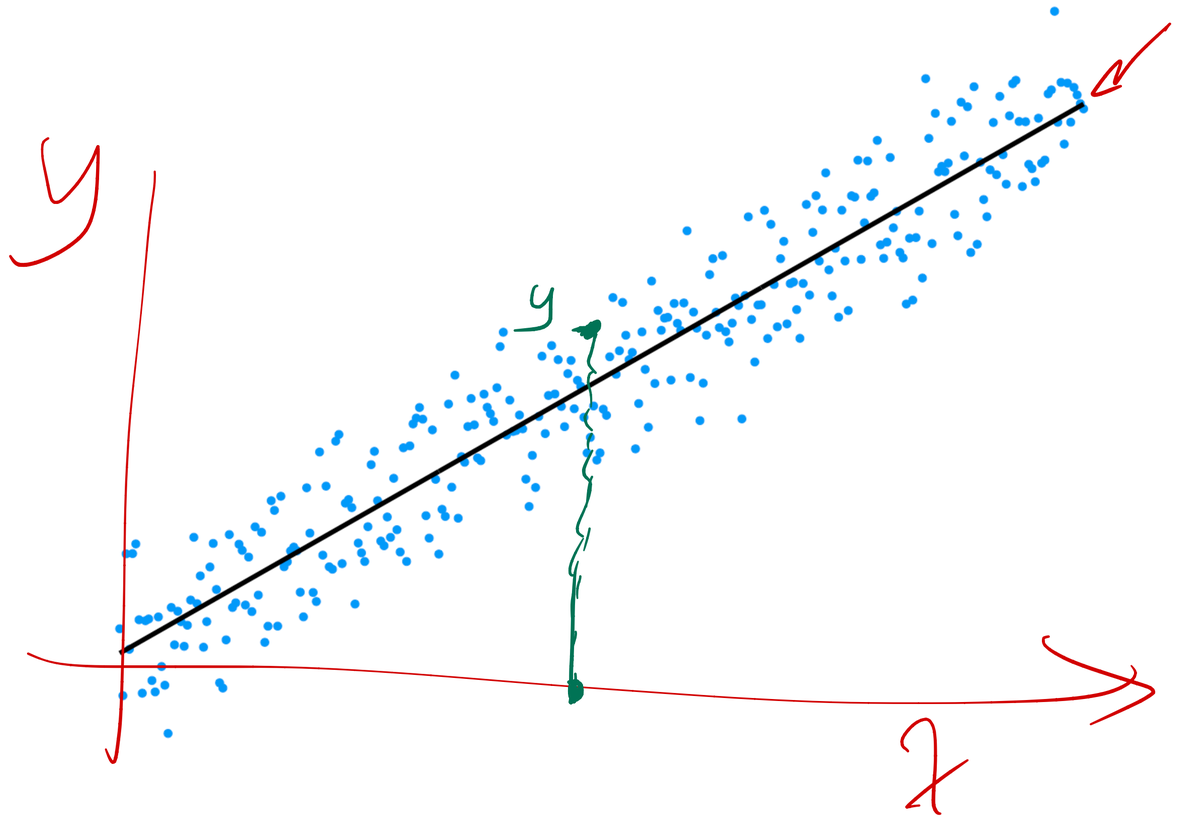
- (Data, observations) => (Rules, models)



$$F = \frac{GM_1M_2}{r^2}$$

Empirical Inference

- (Data, observations) => (Rules, models)



$$Y = aX + b$$

Law

Generalization

- The rule/model should be used to predict new observations
- Example: 1,2,4,7,...
  - What is next?

- 1, 2, 4, 7, 11, ... $a_{n+1} = a_n + n$
- 1, 2, 4, 7, 12, ... $a_{n+2} = a_{n+1} + a_{n+1}$
- 1, 2, 4, 7, 14, 28 divisors of 28
- 1, 2, 4, 7, 1, 1, 5

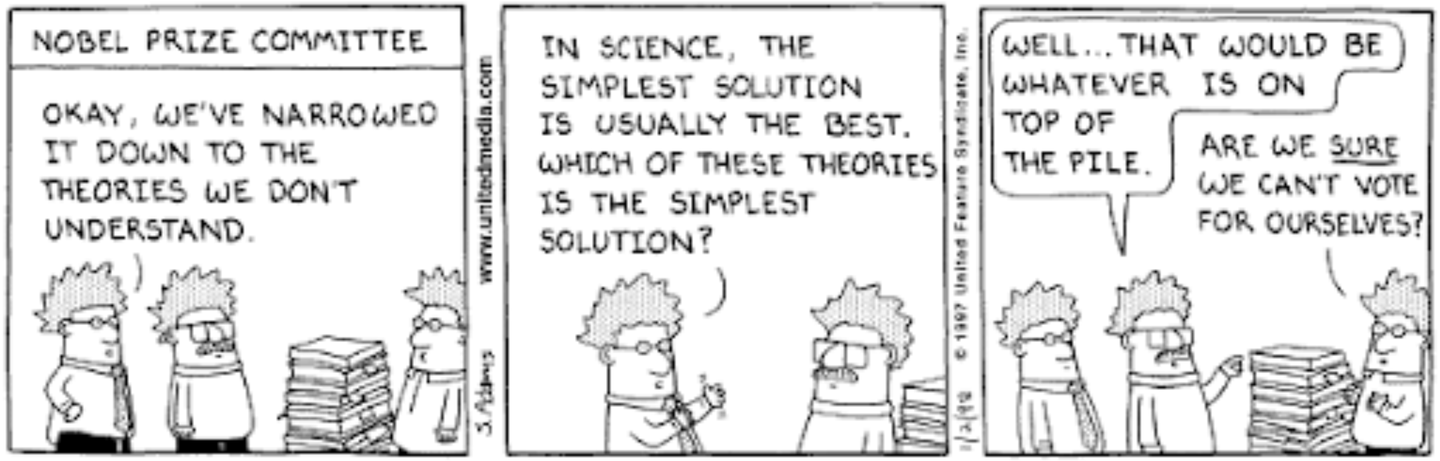
$\pi = 1415$

$e = 2.71$

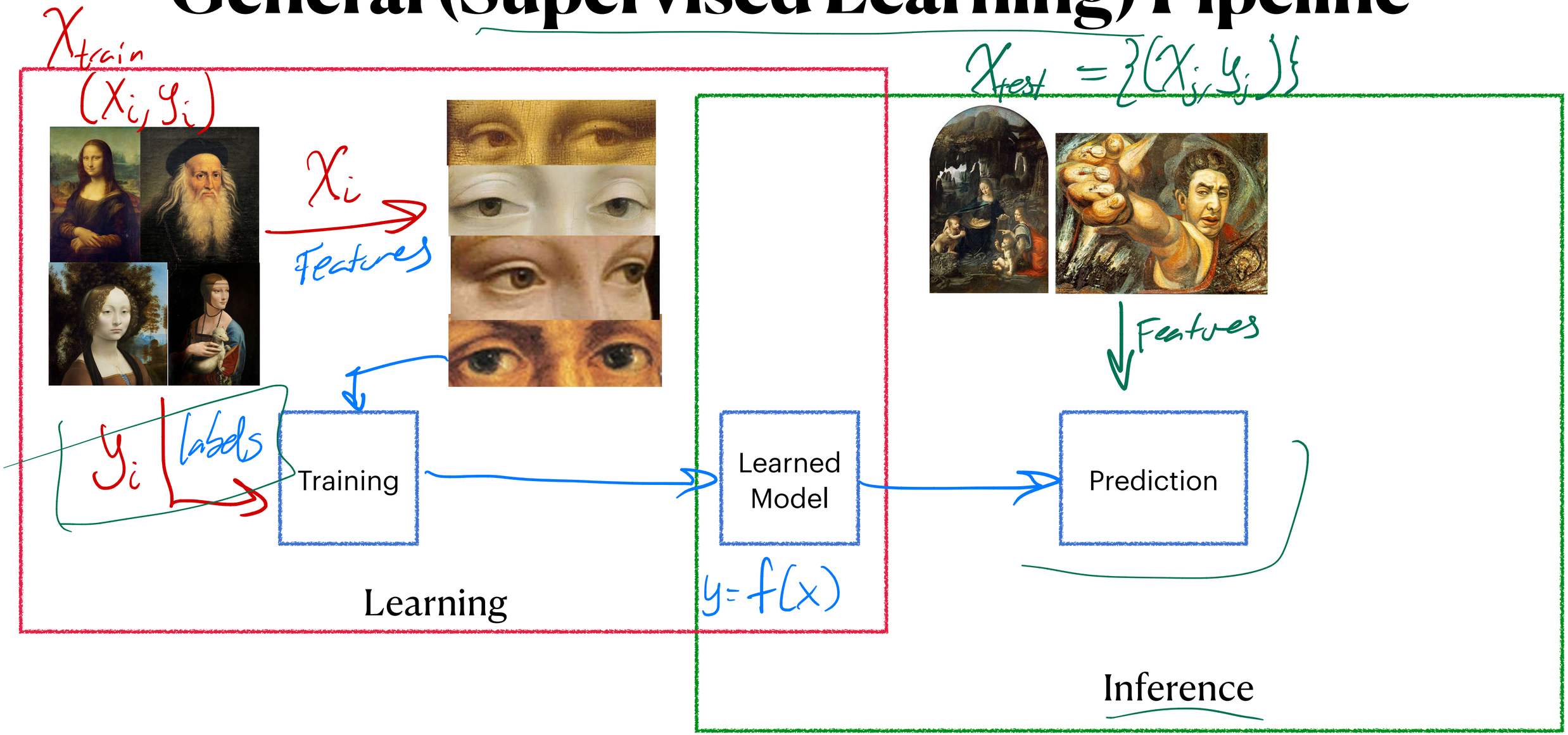
Generalization

Occam's razor

- Among competing hypotheses, the one with the fewest assumptions should be selected.



General (Supervised Learning) Pipeline

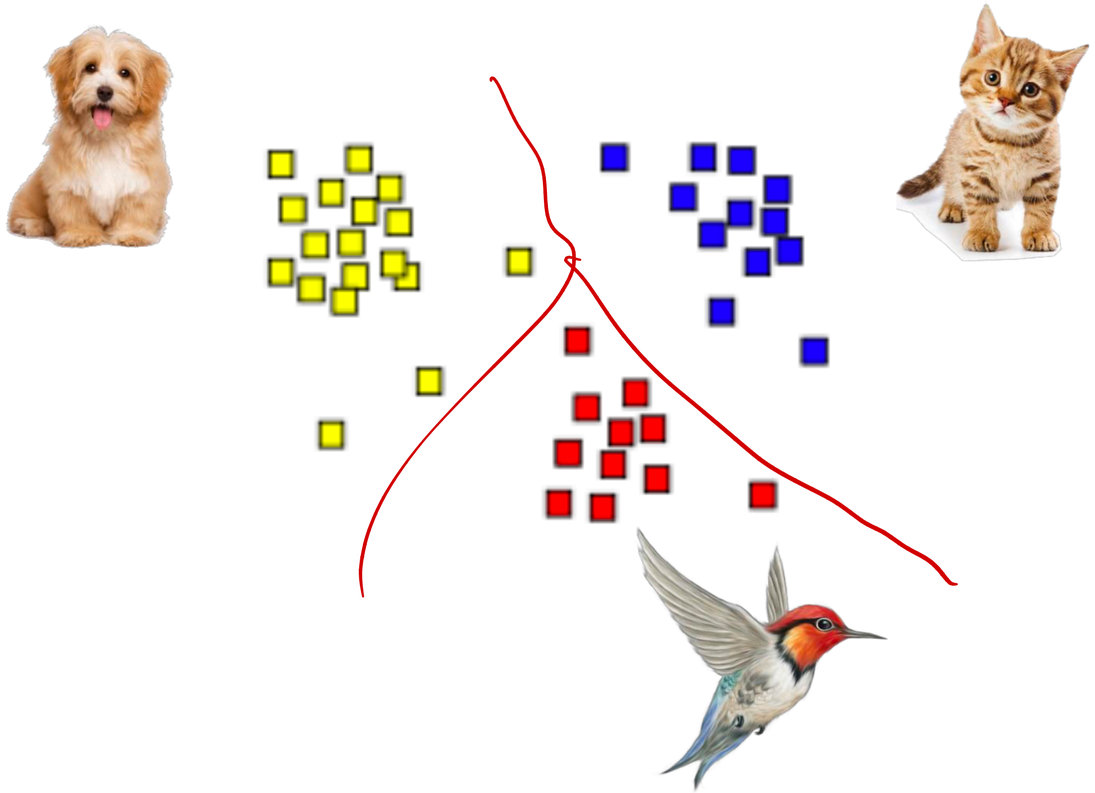


Unsupervised Learning

No labels

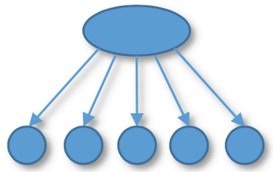
→ clustering

→ dimension Robotics

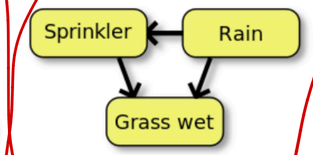


ML

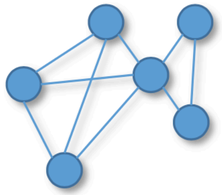
Approaches



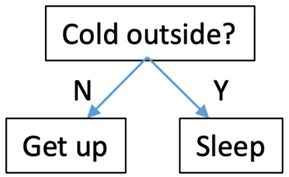
Naïve Bayes



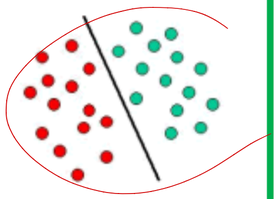
Bayes net



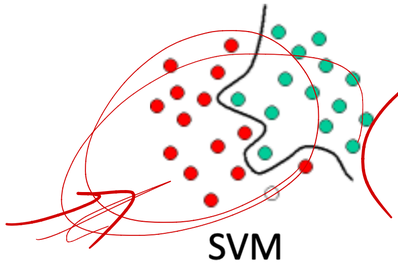
Markov network



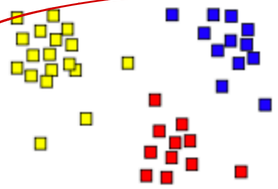
Decision tree



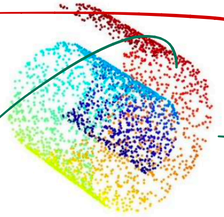
Linear classifier



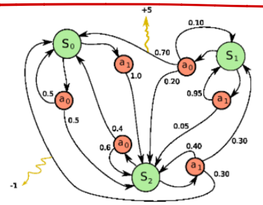
SVM



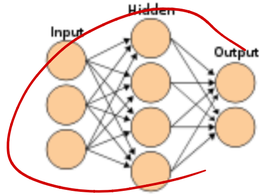
Clustering



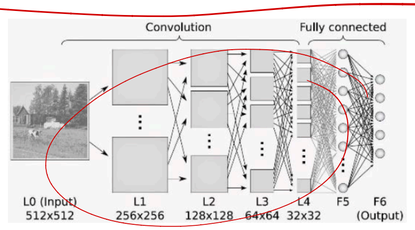
Dimension reduction



Reinforcement learning



Artificial neural network



Deep learning / CNN

Classification

- Set of training examples:
- Learn:
- Evaluate performance:

Naive Bayes Algorithm

- Recall Bayes Rule:

- Likelihood:
  - $L(x) = \frac{P(y = true \mid x)}{P(y = fasle \mid x)}$

Naive Bayes Algorithm

- Estimate $P(y)$ from the training examples
  - $P(y = true) =$
  - $P(y = false) =$

Naive Bayes Algorithm

- From the training examples

1. Define ϕ, a function that maps each input x into a features vector

Naive Bayes Algorithm

- From the training examples

1. Define ϕ, a function that maps each input x into a features vector
2. Estimate the probability distribution of each feature ϕ_j

Naive Bayes Algorithm

- From the training examples

1. Define ϕ, a function that maps each input x into a features vector
2. Estimate the probability distribution of each feature ϕ_j
3. Estimate $P(x \mid y)$ using the naive Bayes assumption

$$p(x|y = true) = \prod_{j=1}^{l} p(\phi_j(x)|y = true)$$

$$p(x|y = false) = \prod_{j=1}^{l} p(\phi_j(x)|y = false)$$