Overview

Kun He

Data Mining and Machine Learning Lab
(John Hopcroft Lab)
Huazhong University of Science & Technology

brooklet60@hust.edu.cn

2022 04 29

Table of contents

Table of Contents

Evaluation

Good news

- No exams!
- No homework!

- Course participation (10%)
- Two small projects (15%*2)
- Final project (60%)

Table of Contents

Source Materials

- No textbook required. Readings will come from freely available online material.
- Machine Learning Lecture Book (Kun He)
- K. Murphy, Machine Learning: a Probabilistic Perspective, MIT Press, 2012
- C. Bishop, Pattern Recognition and Machine Learning, Springer, 2007
- CS 4780 @Cornell University
- Machine Learning (Zhihua Zhou)
- Statistics Learning Method (Hang Li)

Table of Contents

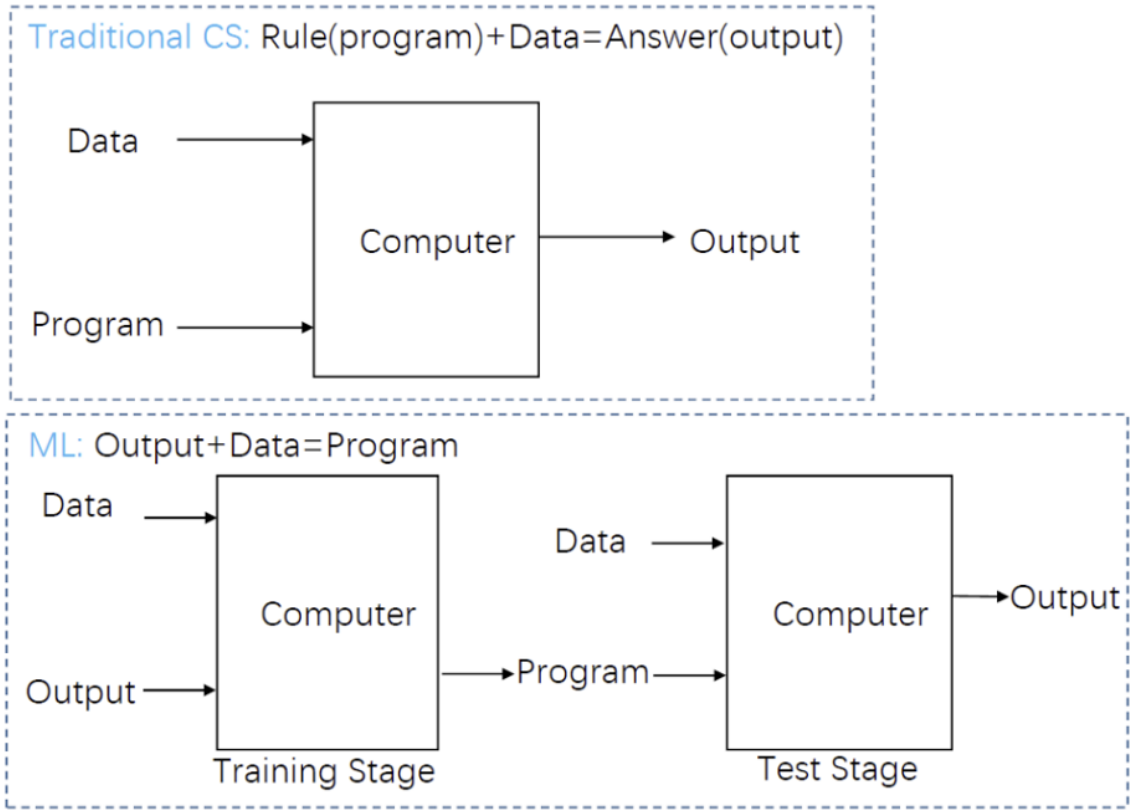
What is machine learning?

Machine Learning

- A branch of artificial intelligence, concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data.
- As intelligence requires knowledge, it is necessary for the computers to acquire knowledge.

- Study of algorithms that improve their performance at some task with experience
- Optimize a performance criterion using example data or past experience
- Role of Statistics: Inference from a sample
- Role of Computer Science: Efficient algorithms to solve the optimization problem, Representing and evaluating the model for inference

ML vs traditional CS



Take the spam email as an example. The core is program but the final goal is the answer.

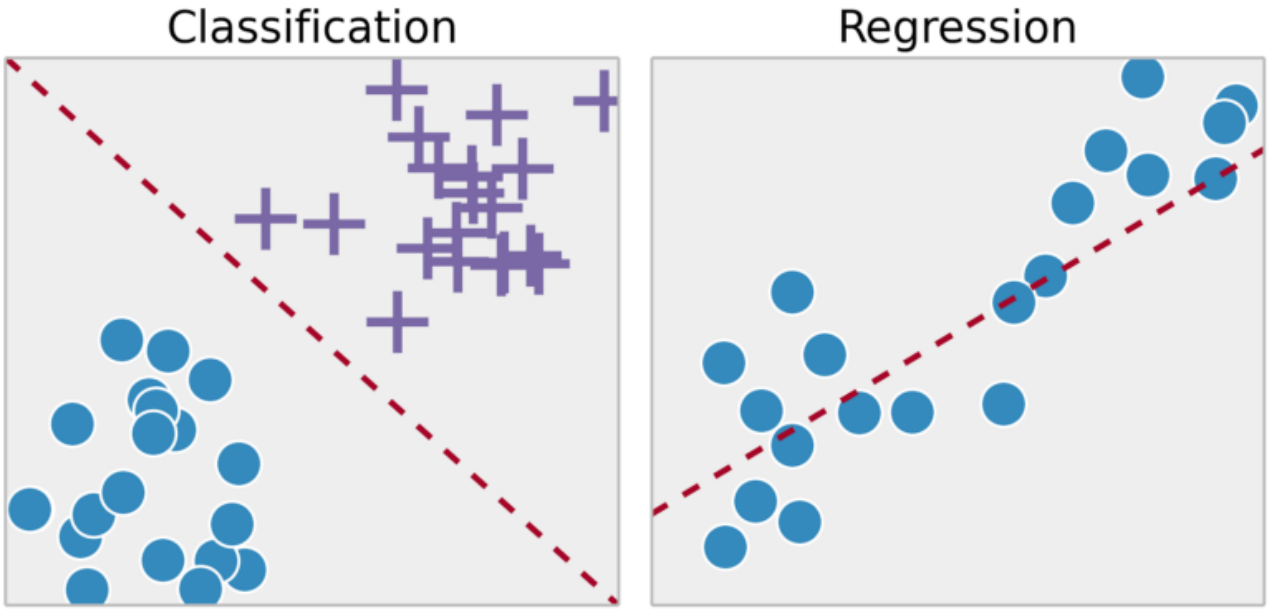
What is machine learning?

Mtichell, 1997

- A computer program A, learn from experience E, w.r.t some tasks T, and performance measure P.
- If E helps improve the performance.

- Based on statistics and optimization, Not on logic.
- Data: i.i.d., independent, identically distributed

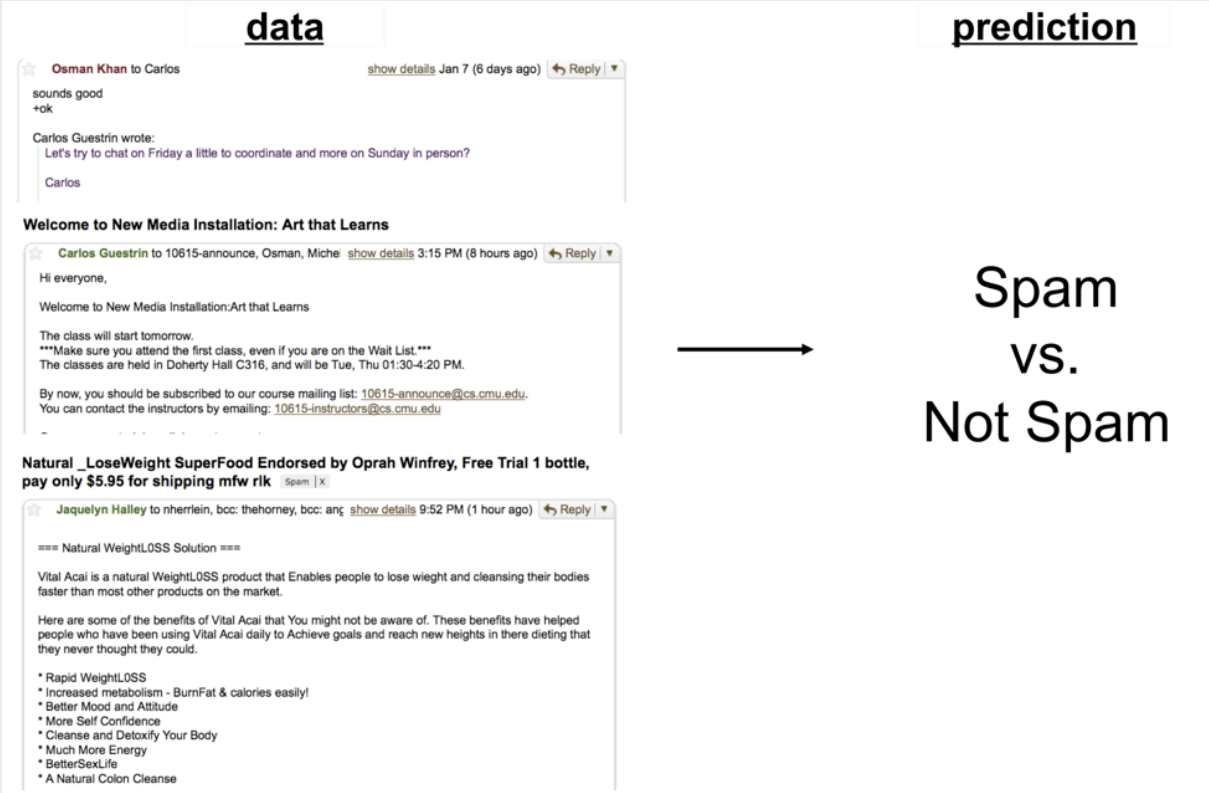
Supervised Learning



Supervised learning, an algorithm learns from a training dataset. We know the correct answers or desired output, the algorithm makes predictions using the given dataset and is corrected by the **supervisor**. The learning stops as and when the algorithm achieves a level of performance which is acceptable. There are two types of supervised learning - **classification** and **regression**.

Supervised Learning-Classification

Spam filtering



Face Recognition

Example training images
for each orientation

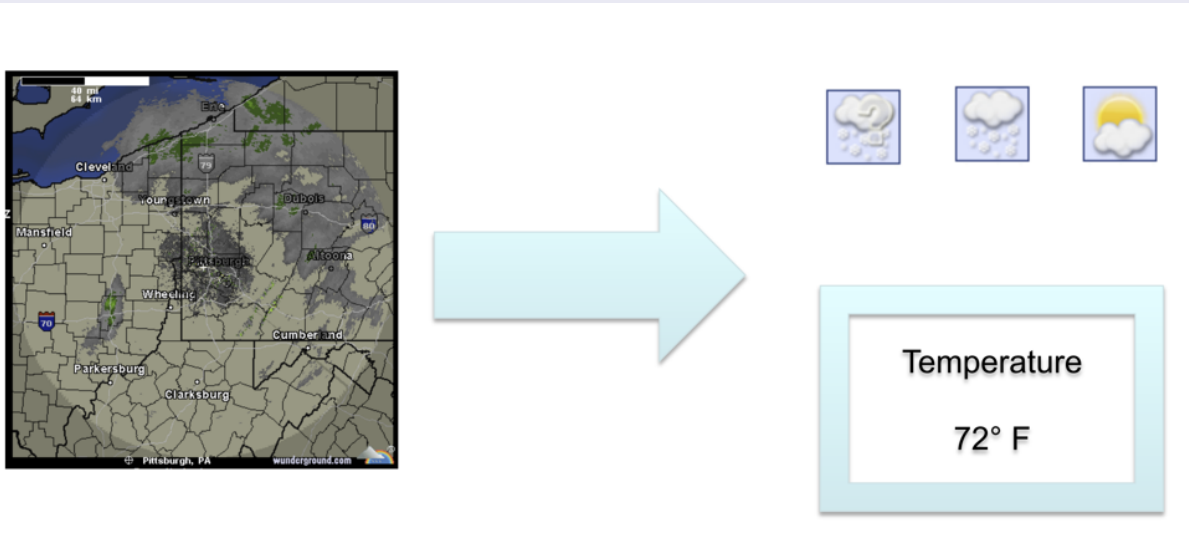
Supervised Learning-Classification

Weather Prediction

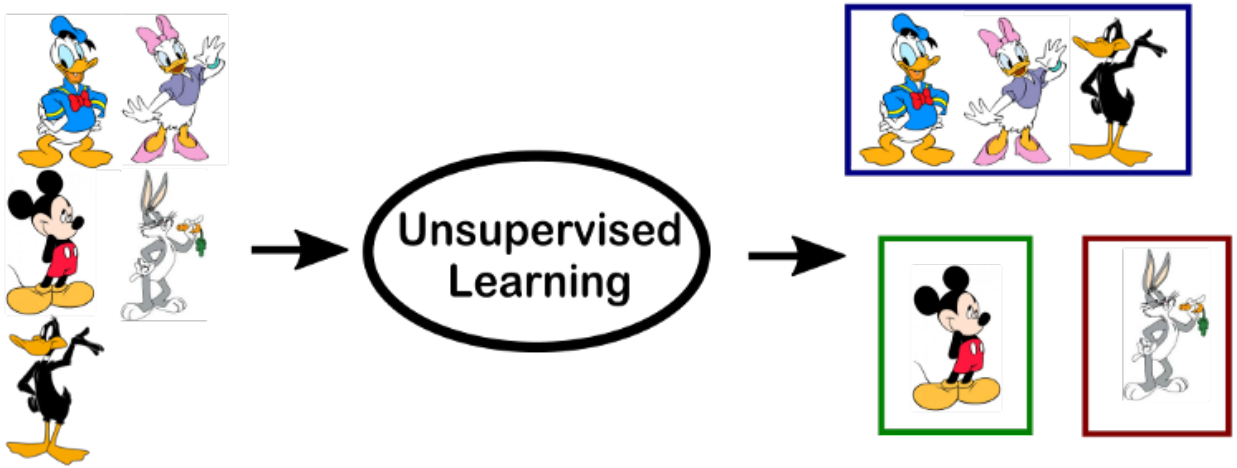
Stock Market Prediction

Supervised Learning-Regression

Weather Prediction



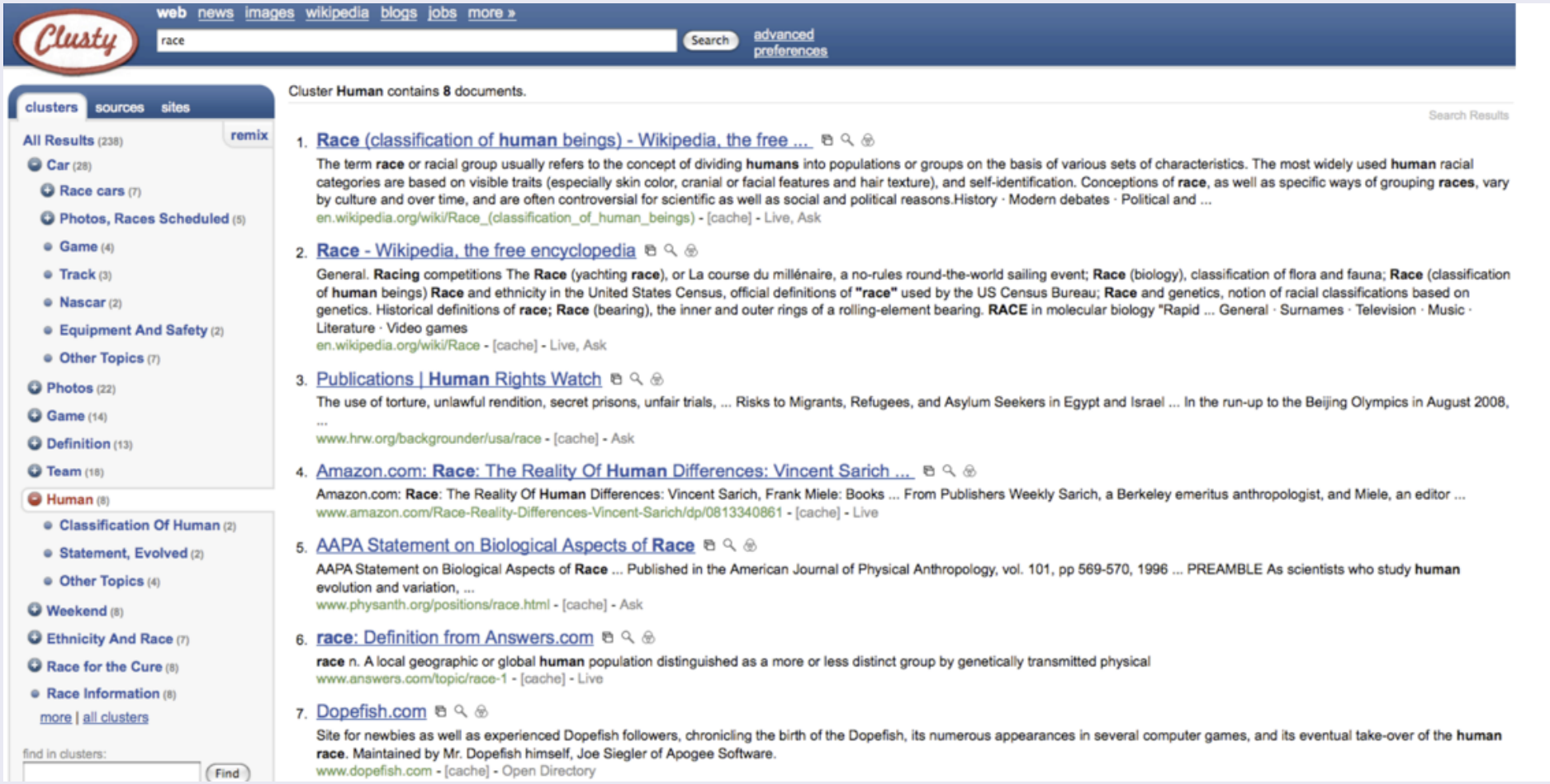
Unsupervised Learning

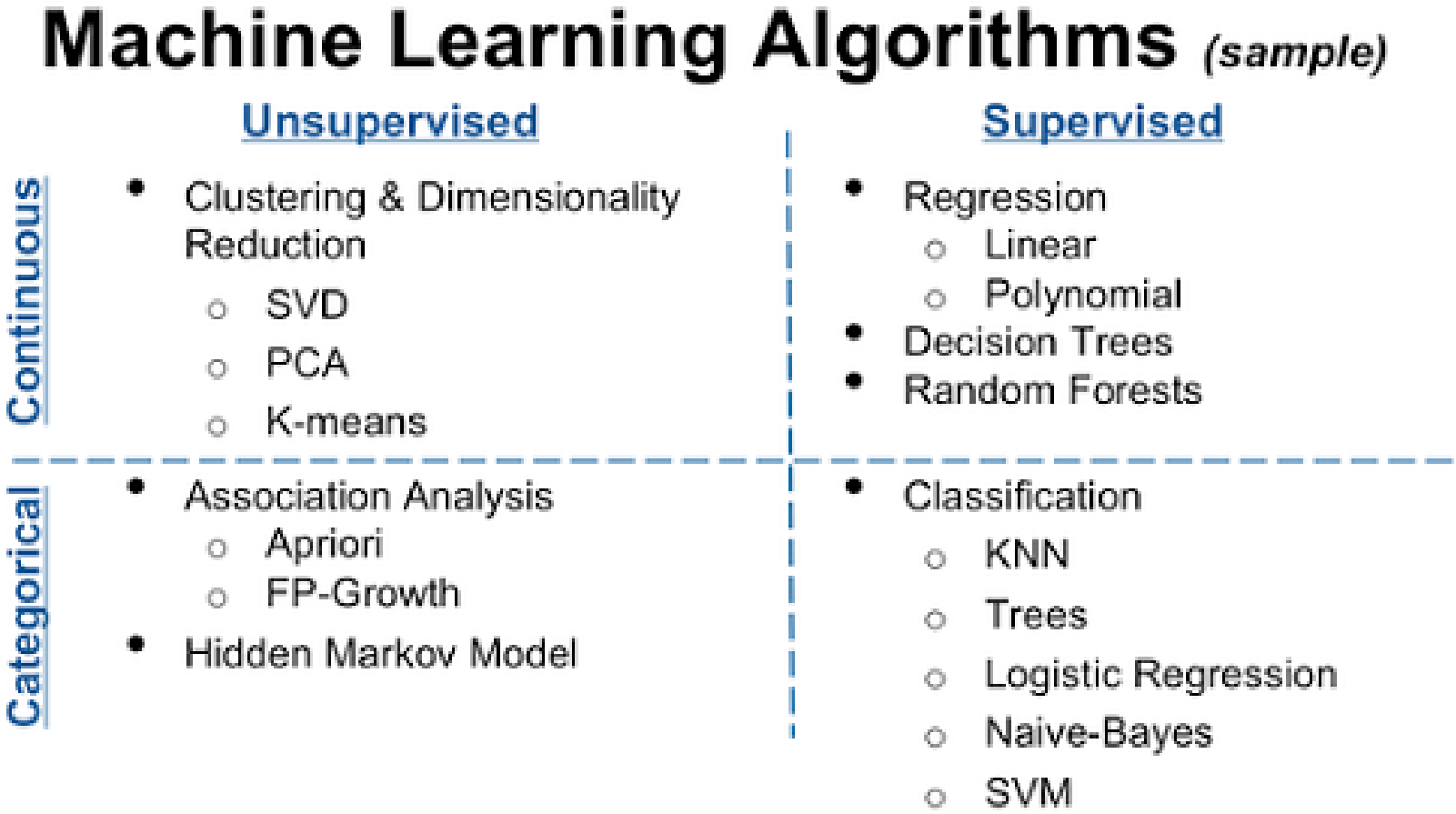


In unsupervised learning there is no trainer or supervisor as the name suggests. We do not show the output, or the specific input required to achieve specific output. The machine learns based on its own capability, it decides what should be the input and output - clustering is a method of unsupervised learning, where the machine learns on its own. A well known algorithm in clustering is the k-means clustering.

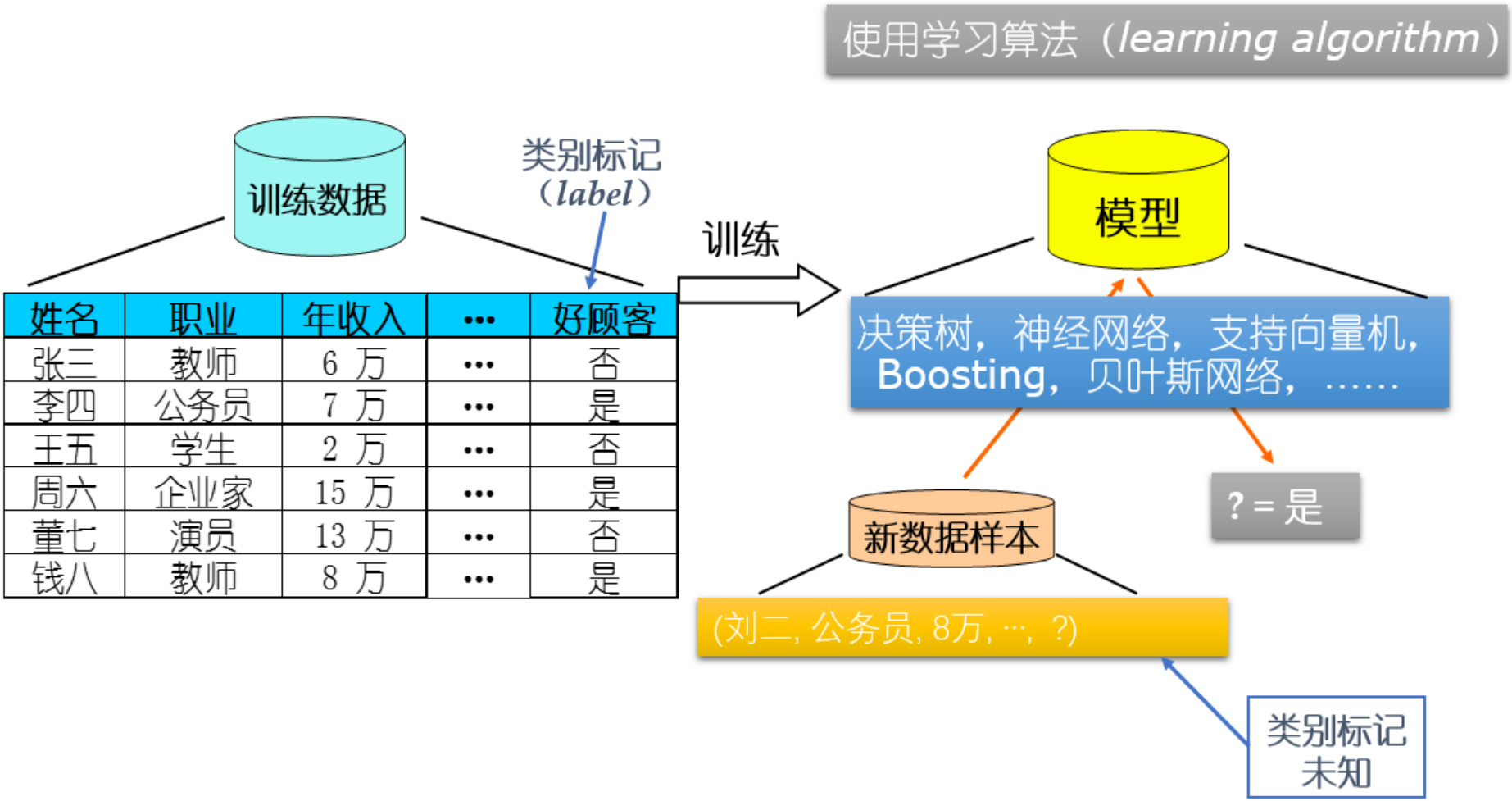
Clustering Images

Clustering Web Search Results

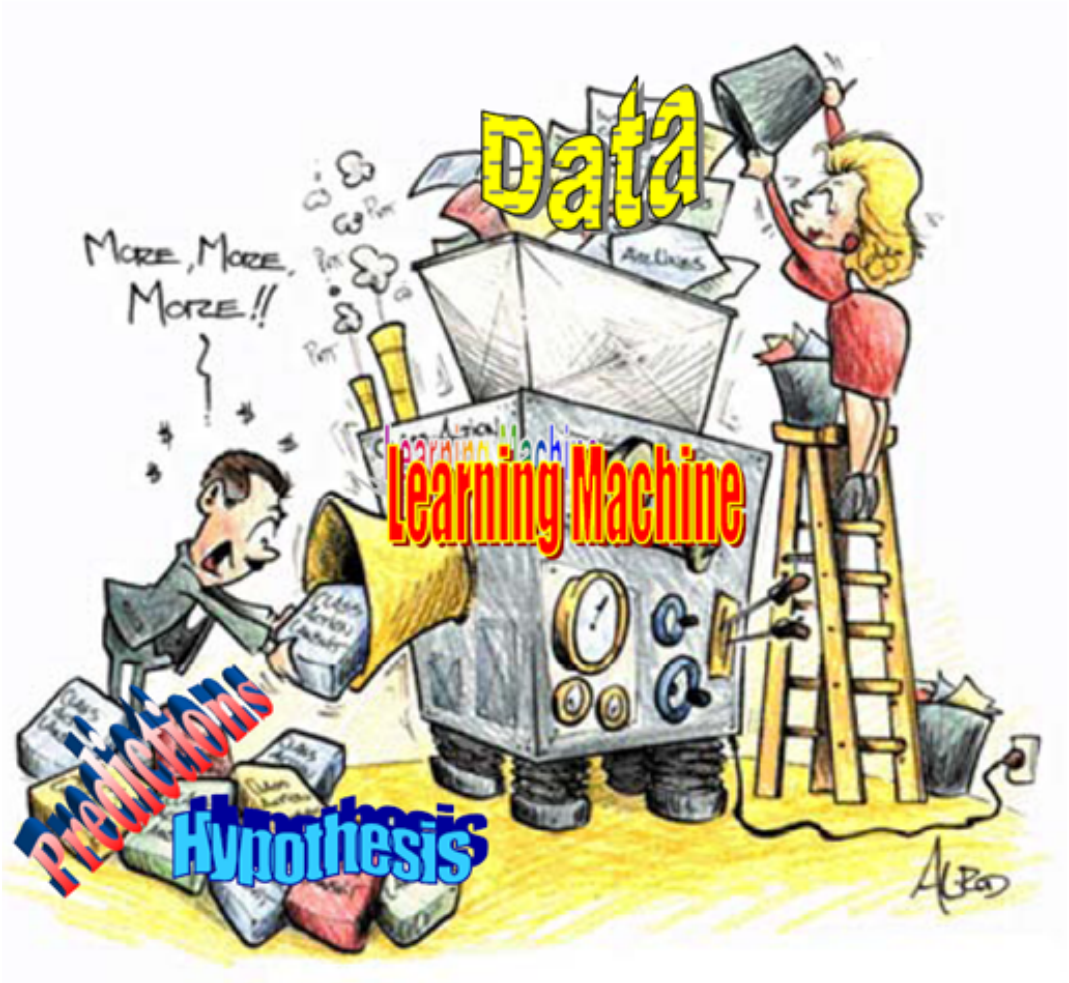



Machine Learning Algorithms (sample)
Unsupervised
Supervised
Continuous
Categorical
Clustering & Dimensionality Reduction
SVD
PCA
K-means
Association Analysis
Apriori
FP-Growth
Hidden Markov Model
Regression
Linear
Polynomial
Decision Trees
Random Forests
Classification
KNN
Trees
Logistic Regression
Naive-Bayes
SVM

Typical Machine Learning Process



Machine Learning Illustration



Rapid growth of machine learning

Machine Learning is perferred approach to

- Natural Language Processing (NLP)
- Computer Vision (CV)
- Medical outcomes analysis
- Robot control
- Computational biology
- Sensor networks
- ...

This trend is accelerating

- Big data
- Improved machine learning algorithms
- Faster computers
- Good open-source software

Table of Contents

History of AI

(1)1952 AI landmark: Samuel Checkers-playing Program-IBM

Herbert A. Simon

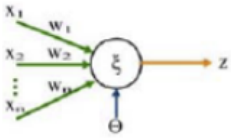
Allen Newell

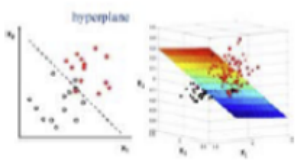
Logical Theory

Principia Mathematica

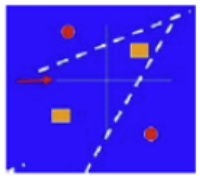
(2) perceptron (1957)

Frank Rosenblatt
(1928-1971)





(3) MLP (1969) Minsky & Papert (1969)



AI winter: AI-logic(Top Down)→
ML: bottom up/Statistical optimization

(4) Edward Albert Feigenbaum(1970)

father of expert systems

(5)Enforcement learning(1994) Deep Blue

(6)Nowadays
Deep learning, Knowledge base, reinforcement learning

History of AI

(1)1952 AI landmark: Samuel Checkers-playing Program-IBM

Herbert A. Simon

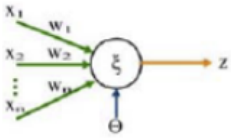
Allen Newell

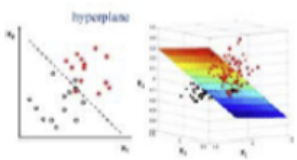
Logical Theory

Principia Mathematica

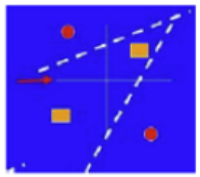
(2) perceptron (1957)

Frank Rosenblatt
(1928-1971)





(3) MLP (1969) Minsky & Papert (1969)



AI winter: AI-logic(Top Down)→
ML: bottom up/Statistical optimization

(4) Edward Albert Feigenbaum(1970)

father of expert systems

(5)Enforcement learning(1994) Deep Blue

(6)Nowadays
Deep learning, Knowledge base, reinforcement learning

The End