

Overview

Kun He

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

brooklet60@hust.edu.cn

2021年09月14日



1 Evaluation

2 Source Materials

3 Introduction

- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning

1 Evaluation

2 Source Materials

3 Introduction

- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning

Good news

- No exams!
- No homework!
- Course participation (10%)
- Two small projects (15%*2)
- Final project (60%)

Table of Contents

1 Evaluation

2 Source Materials

3 Introduction

- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning

- 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

1 Evaluation

2 Source Materials

3 Introduction

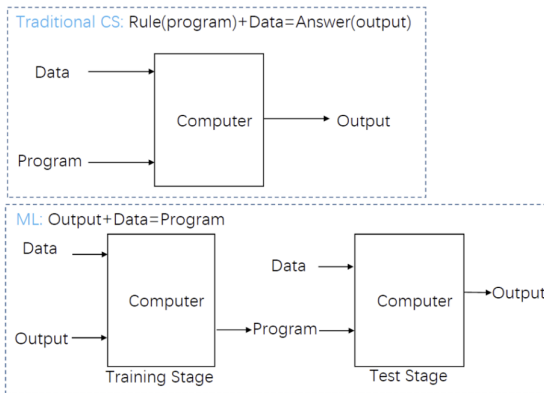
- What is Machine Learning?
- Supervised Learning
- Unsupervised Learning

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



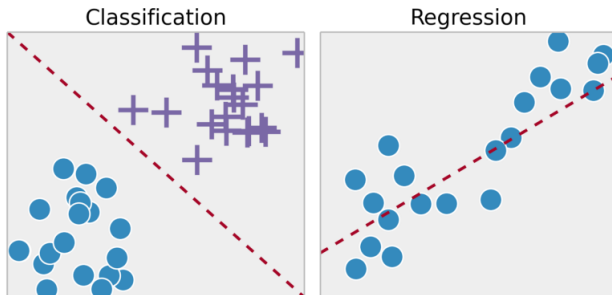
Take the spam email as an example. The core is program but the final goal is the answer.

What is machine learning?

Mitchell, 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**.

Spam filtering

| data | prediction |
|--|----------------------------------|
| <p>Osman Khan to Carlos sounds good rok</p> <p>Carlos Guestrin wrote: Let's try to chat on Friday a little to coordinate and more on Sunday in person?</p> <p>Carlos</p> | <p>Spam vs. Not Spam</p> |
| <p>Welcome to New Media Installation: Art that Learns</p> <p>Carlos Guestrin to 10615-announce, Osman, Miche show details 3:15 PM (8 hours ago) Reply</p> <p>Hi everyone,</p> <p>Welcome to New Media Installation:Art that Learns</p> <p>The class will start tomorrow. ***Make sure you attend the first class, even if you are on the Wait List.*** The classes are held in Doherty Hall C316, and will be Tue, Thu 01:30-4:20 PM.</p> <p>By now, you should be subscribed to our course mailing list: 10615-announce@cs.cmu.edu. You can contact the instructors by emailing: 10615-instructors@cs.cmu.edu</p> | |
| <p>Natural _LoseWeight SuperFood Endorsed by Oprah Winfrey, Free Trial 1 bottle, pay only \$5.95 for shipping mfw rik spam [X]</p> <p>Jaquelyn Halley to mherlein, bcc: thehorney, bcc: anj show details 9:52 PM (1 hour ago) Reply</p> <p>==== Natural WeightLOSS Solution ====</p> <p>Vital Acai is a natural WeightLOSS product that Enables people to lose weight and cleansing their bodies faster than most other products on the market.</p> <p>Here are some of the benefits of Vital Acai that You might not be aware of. These benefits have helped people who have been using Vital Acai daily to Achieve goals and reach new heights in there desing that they never thought they could.</p> <ul style="list-style-type: none">• Rapid WeightLOSS• Increased metabolism - BurnFat & calories easily!• Better Mood and Attitude• More Self Confidence• Cleanse and Detoxify Your Body• Much More Energy• BetterSexLife• A Natural Colon Cleanse | |

Face Recognition



Example training images
for each orientation



Weather Prediction



Stock Market Prediction

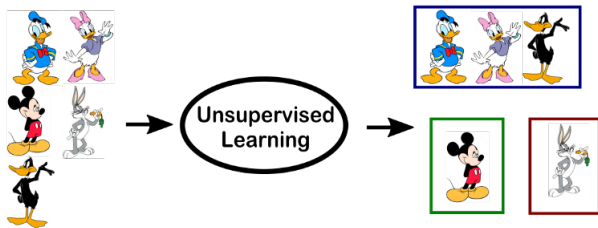


Weather Prediction



Temperature
72° F

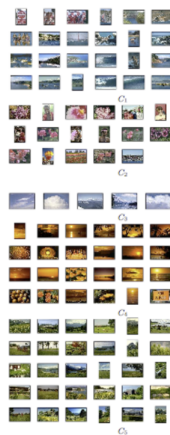
Unsupervised Learning



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 images



[Goldberger et al.]

Clustering Web Search Results



web news images wikipedia blogs jobs more »

race

Search advanced preferences

clusters sources sites remix

All Results (238)




- Car (28)
 - Race cars (7)
 - Photos, Races Scheduled (5)
 - Game (4)
 - Track (3)
 - Nascar (2)
 - Equipment And Safety (2)
 - Other Topics (7)
- Photos (22)
- Game (14)
- Definition (13)
- Team (18)
- Human (8)
 - Classification Of Human (2)
 - Statement, Evolved (2)
 - Other Topics (4)
- Weekend (8)
- Ethnicity And Race (7)
- Race for the Cure (8)
- Race Information (8)




more | all clusters




find in clusters: (Find)




Cluster Human contains 8 documents.




Search Results




- [Race \(classification of human beings\) - Wikipedia, the free ...](#)   




The term **race** or racial group usually refers to the concept of dividing **humans** into populations or groups on the basis of various sets of characteristics. The most widely used **human** racial categories are based on visible traits (especially skin color, cranial or facial features and hair texture), and self-identification. Conceptions of **race**, as well as specific ways of grouping **rac**es, vary by culture and over time, and are often controversial for scientific as well as social and political reasons.History · Modern debates · Political and ...
[en.wikipedia.org/wiki/Race_\(classification_of_human_beings\)](#) - [cache] - Live, Ask
- [Race - Wikipedia, the free encyclopedia](#)   

General. **Racing** competitions The **Race** (yachting **race**), or La Course du millénaire, a no-rules round-the-world sailing event; **Race** (biology), classification of flora and fauna; **Race** (classification of human beings) **Race** and ethnicity in the United States Census, official definitions of "**race**" used by the US Census Bureau; **Race** and genetics, notion of racial classifications based on genetics. Historical definitions of **race**; **Race** (bearing), the inner and outer rings of a rolling-element bearing. **RACE** in molecular biology "Rapid ... General · Surnames · Television · Music · Literature · Video games
[en.wikipedia.org/wiki/Race](#) - [cache] - Live, Ask
- [Publications | Human Rights Watch](#)   

The use of torture, unlawful rendition, secret prisons, unfair trials, ... Risks to Migrants, Refugees, and Asylum Seekers in Egypt and Israel ... In the run-up to the Beijing Olympics in August 2008, ...
[www.hrw.org/backgroundunderusa/race](#) - [cache] - Ask
- [Amazon.com: Race: The Reality Of Human Differences: Vincent Sarich...](#)   

Amazon.com: **Race: The Reality Of Human Differences**: Vincent Sarich, Frank Miele: Books ... From Publishers Weekly Sarich, a Berkeley emeritus anthropologist, and Miele, an editor ...
[www.amazon.com/Race-Reality-Differences-Vincent-Sarich/dp/0813340861](#) - [cache] - Live
- [AAPA Statement on Biological Aspects of Race](#)   

AAPA Statement on Biological Aspects of **Race** ... Published in the American Journal of Physical Anthropology, vol. 101, pp 569-570, 1996 ... PREAMBLE As scientists who study human evolution and variation, ...
[www.physanth.org/positions/race.html](#) - [cache] - Ask
- [race: Definition from Answers.com](#)   

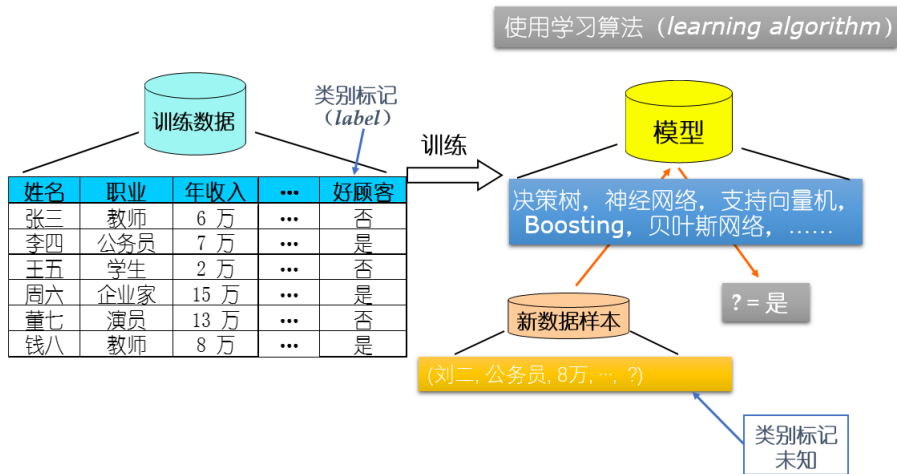
race n. A local geographic or global human population distinguished as a more or less distinct group by genetically transmitted physical
[www.answers.com/topic/race-1](#) - [cache] - Live
- [Dopefish.com](#)   

Site for newbies as well as experienced Dopefish followers, chronicling the birth of the Dopefish, its numerous appearances in several computer games, and its eventual take-over of the **human race**. Maintained by Mr. Dopefish himself, Joe Siegler of Apogee Software.
[www.dopefish.com](#) - [cache] - Open Directory

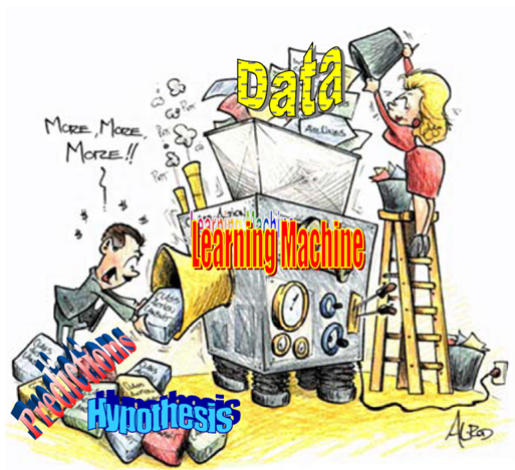
Machine Learning Algorithms *(sample)*

| | <u>Unsupervised</u> | <u>Supervised</u> |
|--------------------|---|---|
| <u>Continuous</u> | <ul style="list-style-type: none">• Clustering & Dimensionality Reduction<ul style="list-style-type: none">◦ SVD◦ PCA◦ K-means | <ul style="list-style-type: none">• Regression<ul style="list-style-type: none">◦ Linear◦ Polynomial• Decision Trees• Random Forests |
| <u>Categorical</u> | <ul style="list-style-type: none">• Association Analysis<ul style="list-style-type: none">◦ Apriori◦ FP-Growth• Hidden Markov Model | <ul style="list-style-type: none">• Classification<ul style="list-style-type: none">◦ KNN◦ Trees◦ Logistic Regression◦ Naive-Bayes◦ SVM |

Typical Machine Learning Process



Machine Learning Illustration



Rapid growth of machine learning

Machine Learning is preferred 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

History of AI

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

Herbert A. Simon



Allen Newell



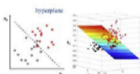
Logical Theory

Principia Mathematica

(2) perceptron (1957)



Frank Rosenblatt
(1918-1970)



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



AI winter: AI-logic(Top Down)→

ML: bottom up/Statistical optimization

(6) Nowadays

Deep learning, Knowledge base, reinforcement learning



(5) Enforcement learning(1994) Deep Blue



(4) Edward Albert Feigenbaum(1970)



father of expert systems



The End