

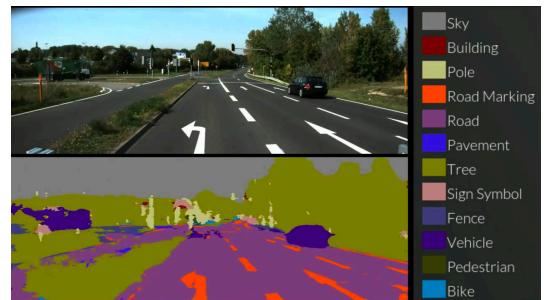
Machine Learning
is
the science (and art) of
programming computers so they
can *learn from data*.



Machine learning makes drug discovery and patient diagnosis quicker, cheaper and more effective. E.g. Dermatologist-level classification of skin cancer with deep neural networks. [Nature](#)



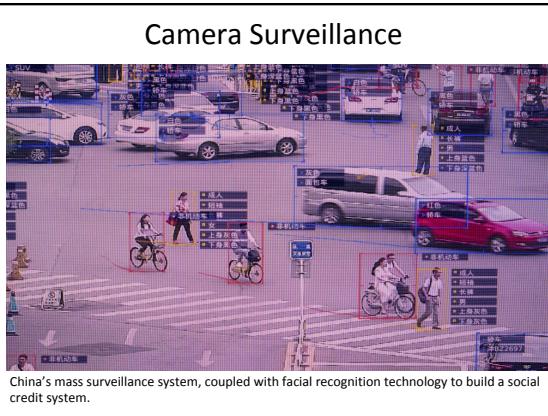
More than 80% of TV shows on Netflix are found through its recommendation engine based on machine learning, as the platform caters to more than 100 million subscribers



Deep learning enables self-driving cars



Machine learning is a fundamental part of the Uber model, which uses machine learning algorithms to determine arrival times, pick-up locations, and UberEATS' delivery estimations.

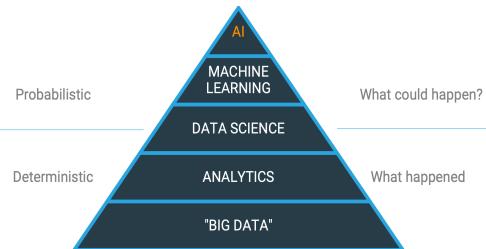


ML is Everywhere

- How many ML-enabled solutions you have encountered today using your smart phone?

What is ML

ML IS AT THE HEART OF TRANSFORMATION



Machine Learning & AI

- **Artificial Intelligence:** Design an intelligent agent that perceives its environment and make decisions to maximize chances of achieving its goal.
Subfields: Robotics, Machine Learning, Natural Language processing
- **Machine Learning:** Give “computer the ability to learn without being explicitly programmed.”
(Author Samuel, 1959)

Why Learning

Learning is used when:

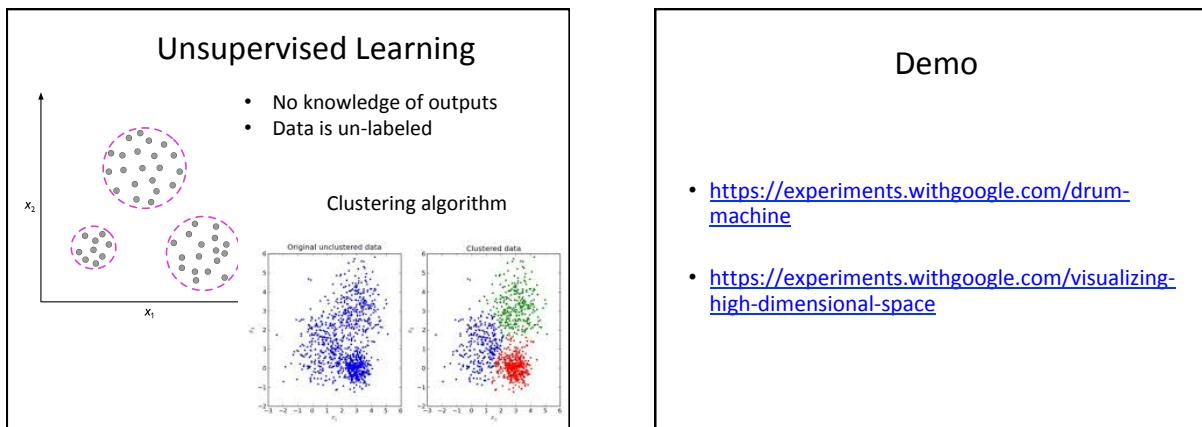
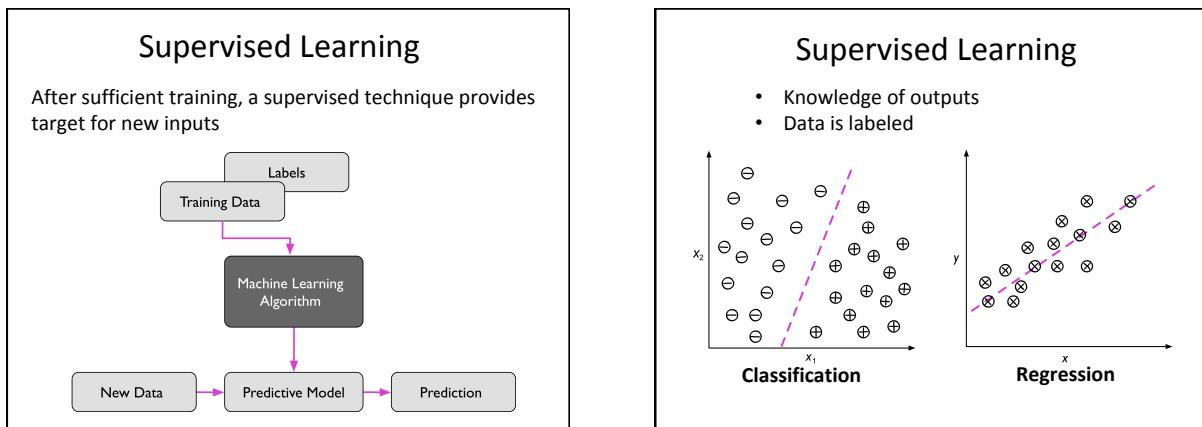
- Human expertise does not exist
(e.g. navigating on Mars)
- Humans are unable to explain their expertise
(e.g. speech recognition)
- Solution changes in time
(e.g. routing on a computer network. self-driving cars)

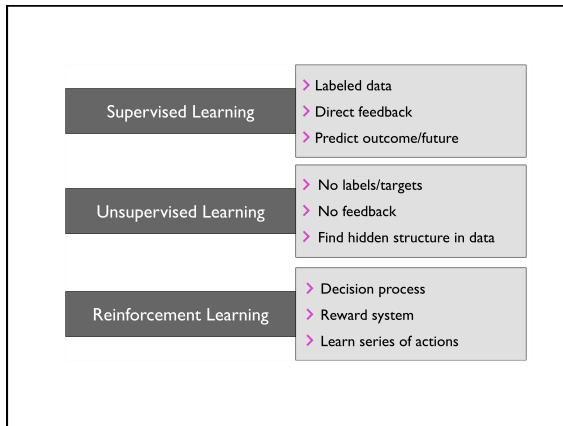


IBM Deep Blue defeated the world chess champion Garry Kasparov May 1997. Rule-based system.



Professional Go player Lee Sedol reviewing his math with AlphaGo after defeat March 2016. [Atlantic](#)

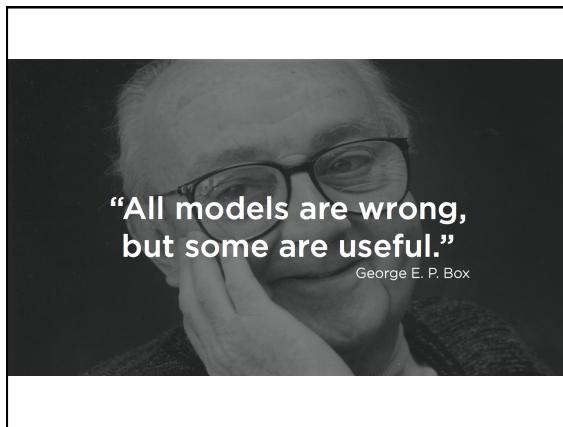




Exercise 1: Supervised or Unsupervised Learning

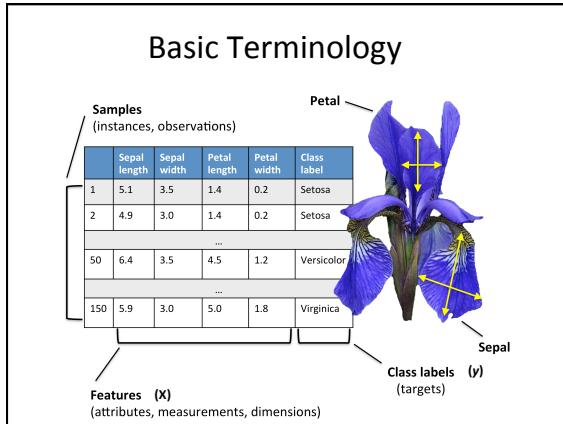
- Customer relationship management analysis(CMA) in a grocery chain e.g. Market Basket.
- Fraud detection and fraud prevention in a credit card company e.g. Visa Card.

CMA: Processing data of customers and improve sales.



"It's not who has the best algorithm that wins; It's who has the most data."

- by Andrew Ng (co-founder of Google Brain)



Terminology & Notation

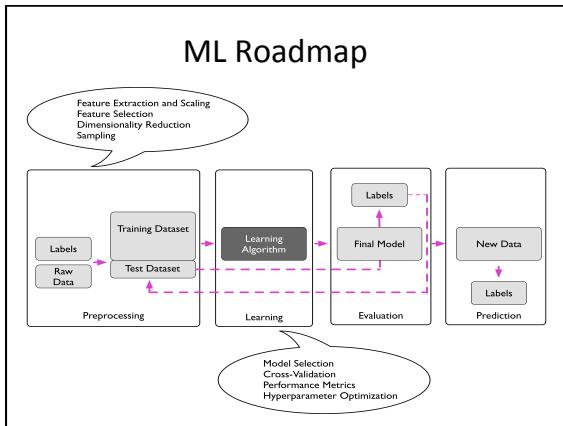
Feature Table
Feature # 1 Sepal Length
Feature # 2 Sepal Width
Feature # 3 Petal Length
Feature # 4 Petal Width

INPUT: The features of the 150 samples can be written as a 150*4 matrix.

Target: [Setosa, Versicolor, Virginica]

TARGET: The targets/labels of the 150 samples can be written as a 150-dimensional column vector.

Sample ID	Sepal Length	Sepal Width	Petal Length	Petal Width	Class
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	3.6	1.4	0.2	Iris-setosa
6	5.4	4.5	1.7	0.4	Iris-setosa
7	4.6	3.4	1.4	0.3	Iris-setosa
8	5	3.4	1.4	0.2	Iris-setosa
9	4.8	3.0	1.4	0.2	Iris-setosa
10	4.4	2.9	1.3	0.2	Iris-setosa
11	4.9	3.1	1.5	0.2	Iris-setosa
12	5.4	3.7	1.5	0.2	Iris-setosa
13	4.8	3.4	1.6	0.2	Iris-setosa
14	4.8	3.0	1.4	0.2	Iris-setosa
15	4.3	3	1.1	0.1	Iris-setosa
16	5.8	4	1.2	0.2	Iris-setosa
17	5.7	4.4	1.5	0.2	Iris-setosa
18	5.4	3.9	1.3	0.4	Iris-setosa
19	5.1	3.5	1.4	0.3	Iris-setosa
20	5.9	3.7	1.7	0.3	Iris-setosa
21	5.1	3.8	1.5	0.3	Iris-setosa
22	5.4	3.4	1.7	0.2	Iris-setosa
23	5.1	3.3	1.5	0.2	Iris-setosa
24	4.6	3.6	1	0.2	Iris-setosa
25	5.1	3.3	1.7	0.5	Iris-versicolor



Exercise 2

- Identify the four ML steps: 1) Pre-processing
2) Learning 3) Evaluation and 4) Prediction, in the first_ml_iris example.