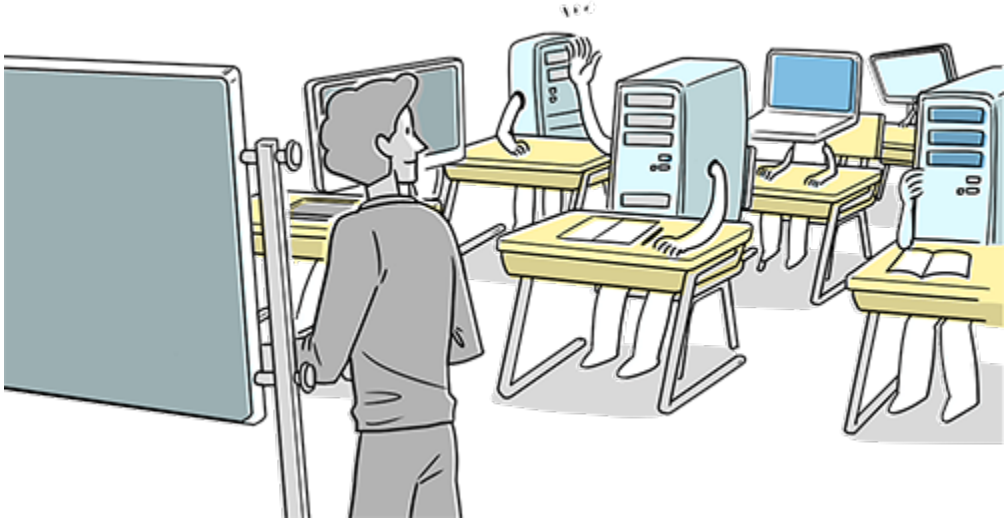
STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY

CPE/EE 695: Applied Machine Learning

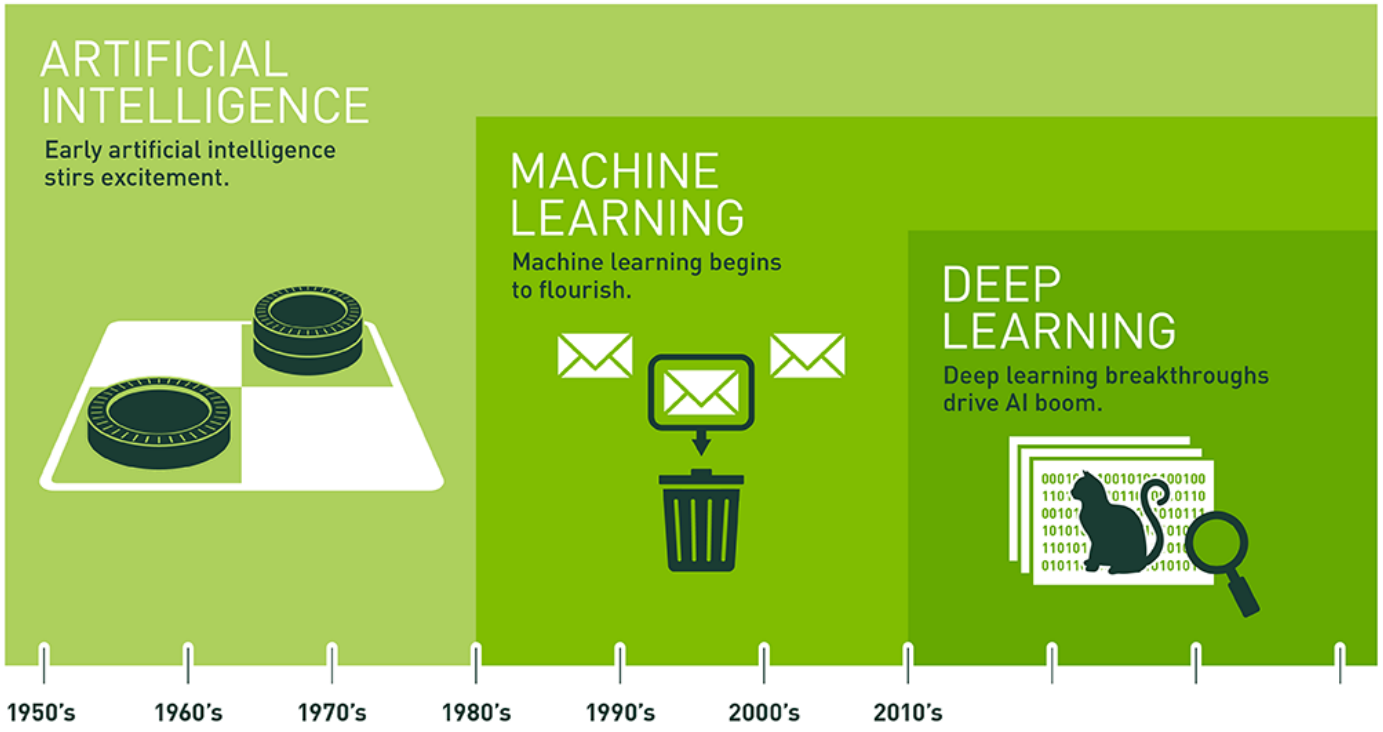
Lecture 1: Introduction to ML

Dr. Shucheng Yu, Associate Professor
Department of Electrical and Computer Engineering
Stevens Institute of Technology

What is Machine Learning?



Relation with other “buzzwords”



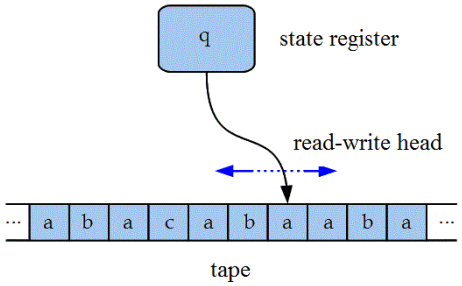
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

Computability theory (1930's -)

Church-Turing Thesis (1936):

"a function on the natural numbers can be calculated by an effective method, if and only if it is computable by a **Turing Machine**"

"digital computers can simulate any process of formal reasoning"



Turing Machine

The First General-Purpose Electronic Computer

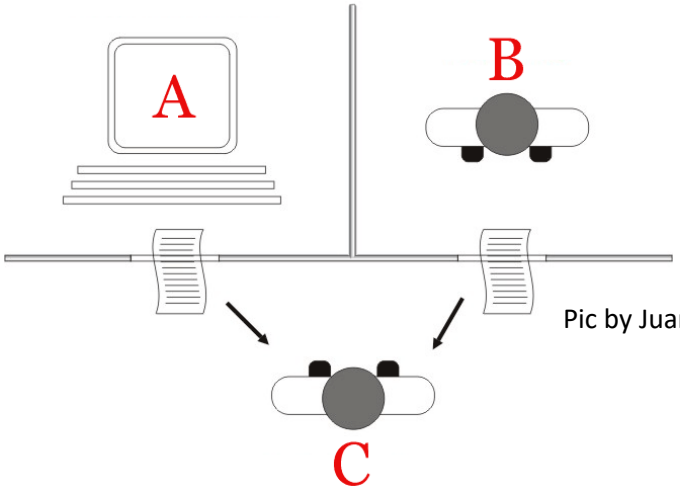
ENIAC (1946) - Electronic Numeral Integrator And Computer

17,468 vacuum tubes,
70,000 resistors,
10,000 capacitors,
1,500 relays,
6,000 manual switches ,
5 million soldered joints

Can Machine Think?

Turing Test (1950)

- ELIZA (1966)
- PARRY(1972)
- Loebner Prize (1990)



Pic by Juan Alberto Sánchez Margallo

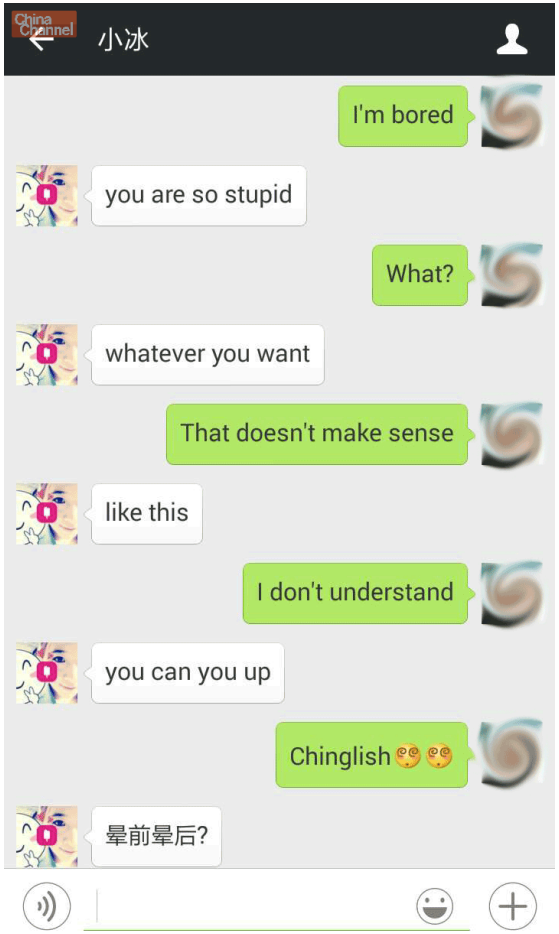
Player C, the interrogator, is given the task of trying to determine which player – A or B – is a computer and which is a human. The interrogator is limited to using the responses to written questions to make the determination

Can Siri pass the Turing test?

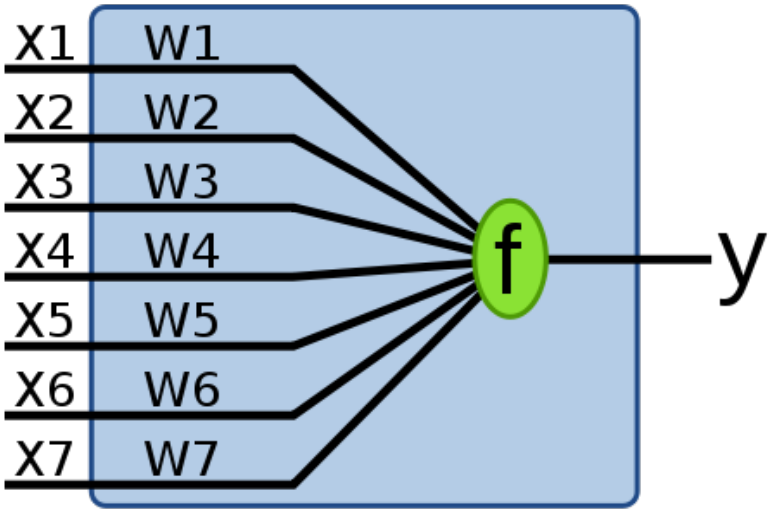
(Recorded on July 29, 2013)

Human: Hello I'm Ronan. what is your name?
Siri: Siri. Please to meet you.
Human: What is your mother's name?
Siri: I'd rather not say.
Human: What is your birth sign?
Siri: I don't really like talking about myself.
Human: How many children do you have?
Siri: Only biological entities have children.
Human: Do you prefer red or white wine?
Siri: I couldn't find any restaurants matching "red or white wine"
Human: I like bananas. Which is your favorite fruit?
Siri: I prefer not to say.
Human: What music do you like to listen to?
Siri: I really have no opinion.
Human: what is your favorite song?
Siri: I like whatever you like.
Human: I like Waiting for Godot. What is your favorite play?
Siri: I don't really like talking about myself.
Human: What color do you dye your hair?
Siri: Would you like me to search the web for "What color do you dye your hair?"?

Microsoft Chatbot Xiaobing

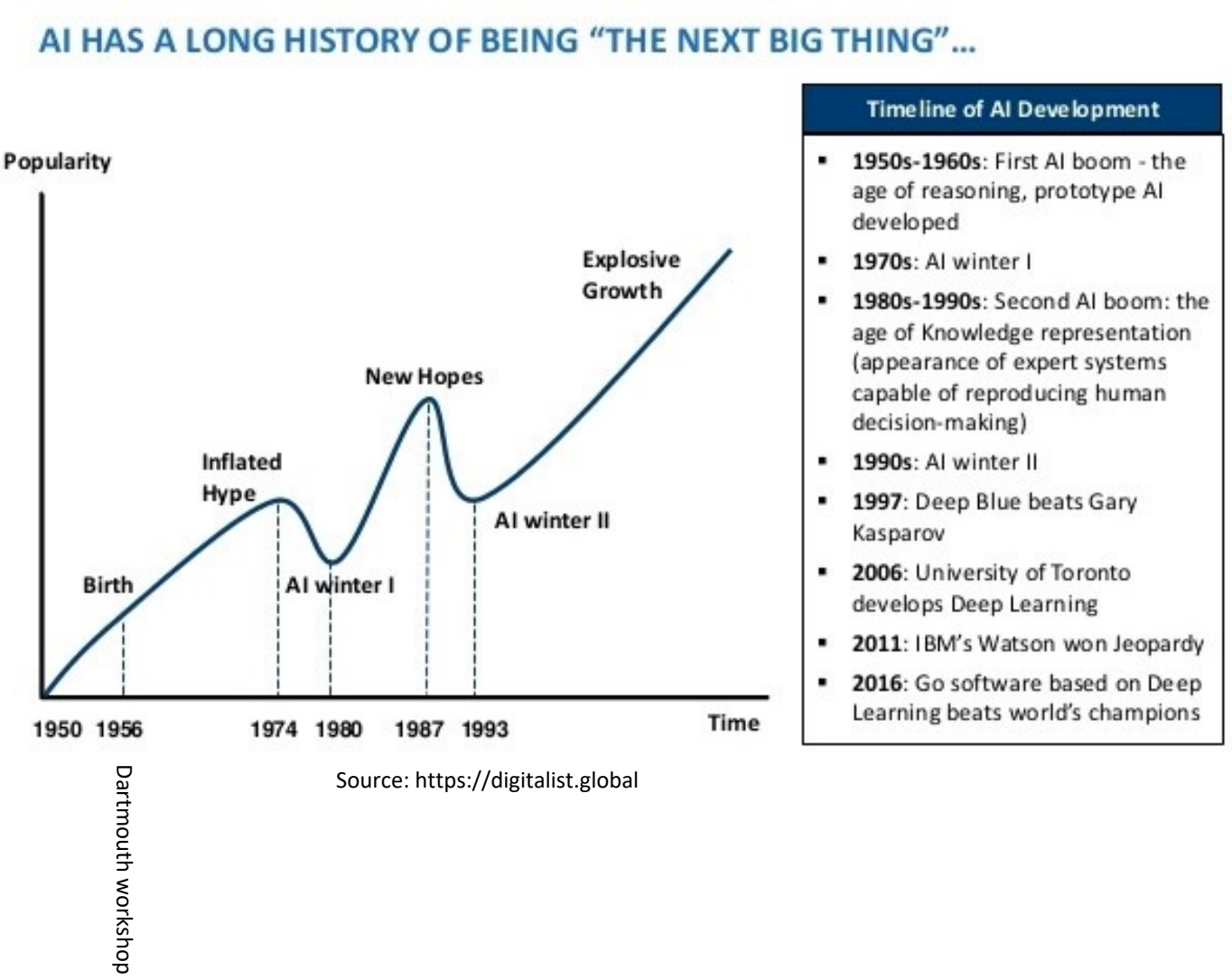


Perceptrons



$$f(x) = \begin{cases} 1, & if\ wx + b > 0 \\ 0, & \text{otherwise} \end{cases}$$

AI Cycles



Source: https://digitalist.global

AI Winter

- Failure of Machine Translation -- disambiguation
- Perceptron Limitation – Linear
- Failure of Speech Understanding Research at CMU
- Lighthill Report(1973): James Lighthill, "*Artificial Intelligence: a paper symposium*"

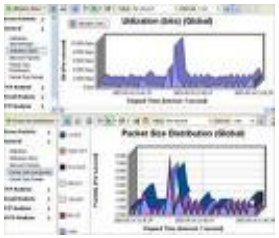
"In no part of the field have the discoveries made so far produced the major impact that was then promised"

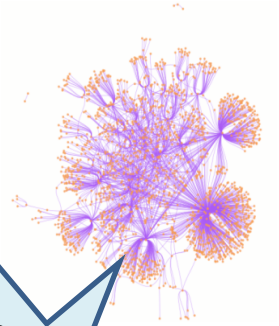
- DARPA funding cut

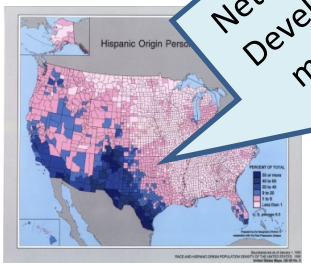
Exclusive-OR gate

A	B	Output
0	0	0
0	1	1
1	0	1
1	1	0

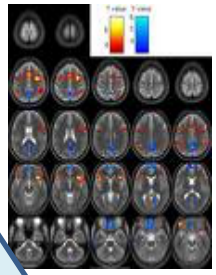
After 1990's







Network Develop ment



Computer Maturity

Statistician

Computer Scientist

ML

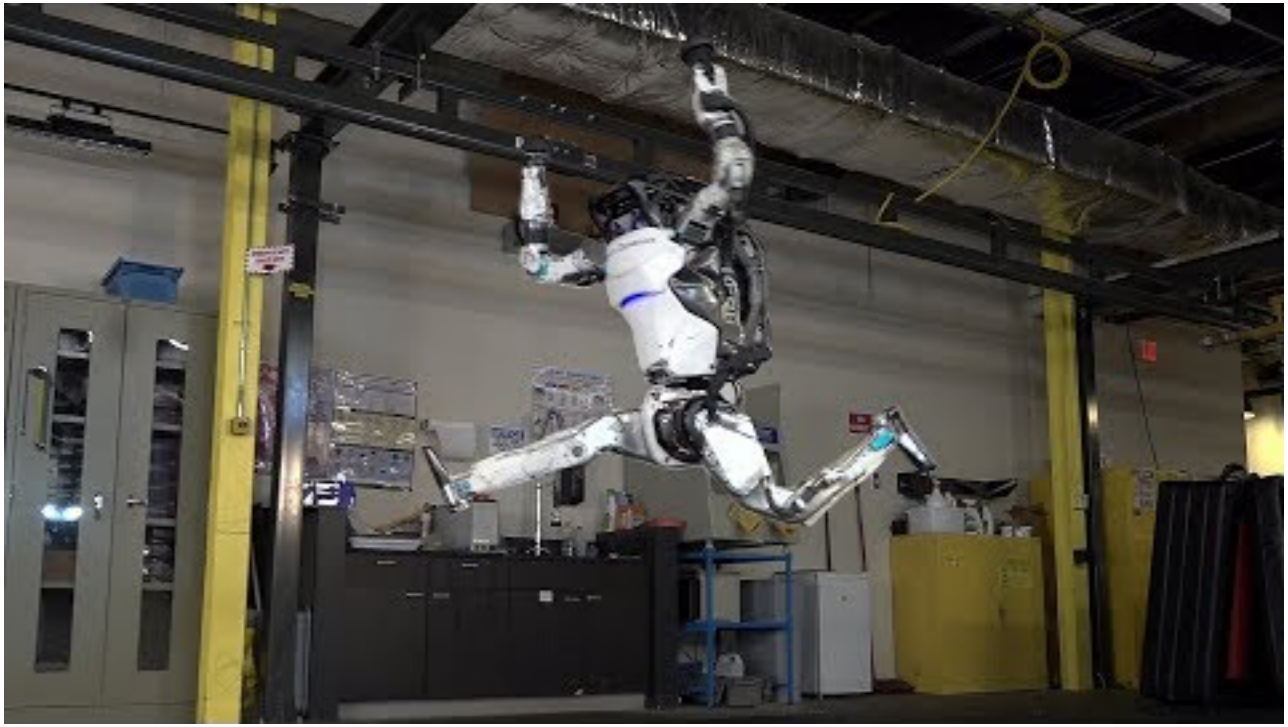
Computer Engineer

Philosopher
Psychologist
Neurologist

Deep Blue/Watson
Cog/Kismet/Andriod
Climate Model
Marketing/Recommendatio n Model
AlphaGo
.....

Backpropagation
Adaboost
Naive Bayes
Deep Learning .....

Robots

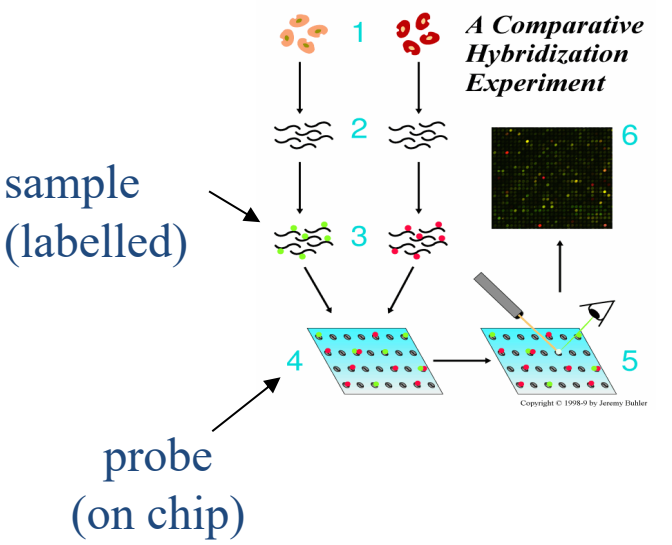


Signal Processing

- Face Recognition

- Speech Recognition

- Microarray



[image from Jeremy Buhler]

- Video Processing

Business Analysis

Fraud Detection

Churn Prediction

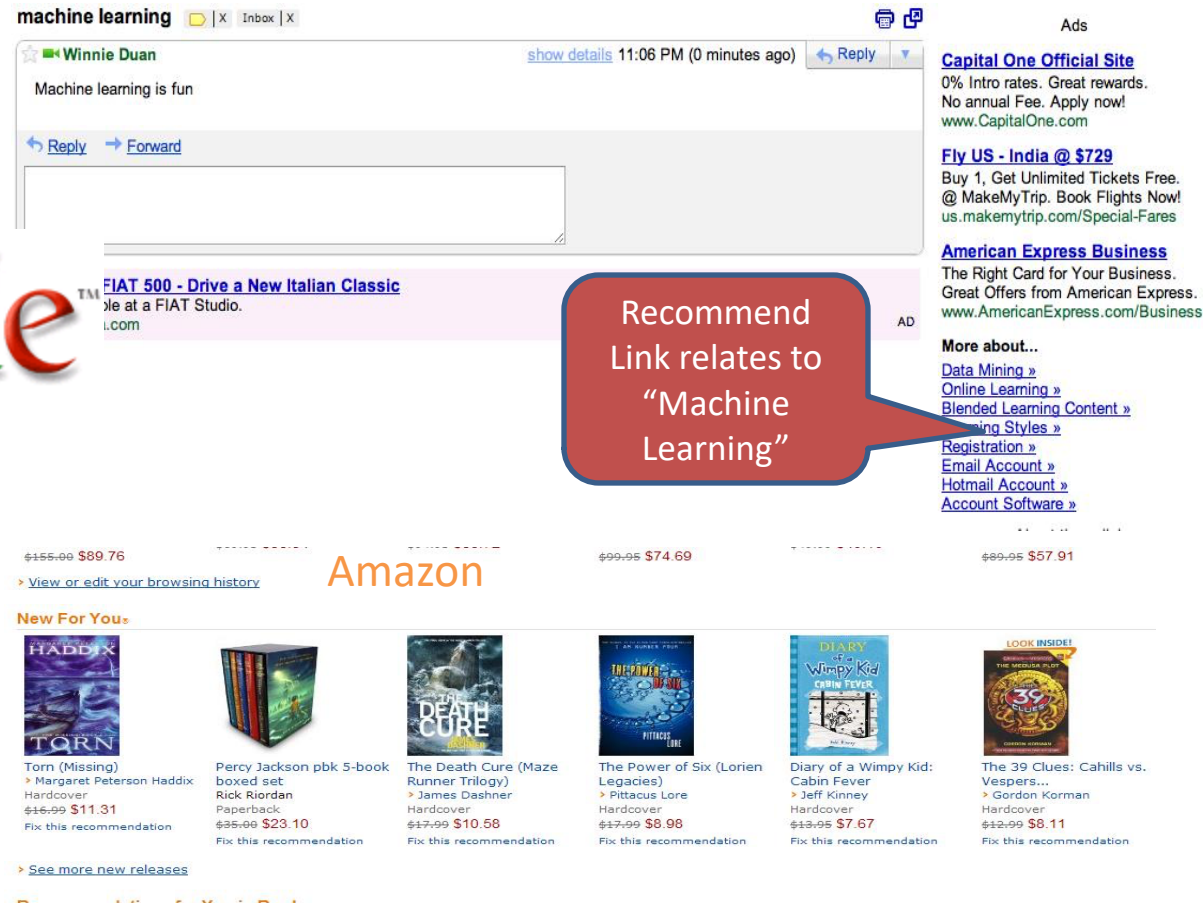
Market Basket

Web Mining

Stock Market

Recommendation Systems

$1 million, 10% improve,
4 years(2006 – 2009)



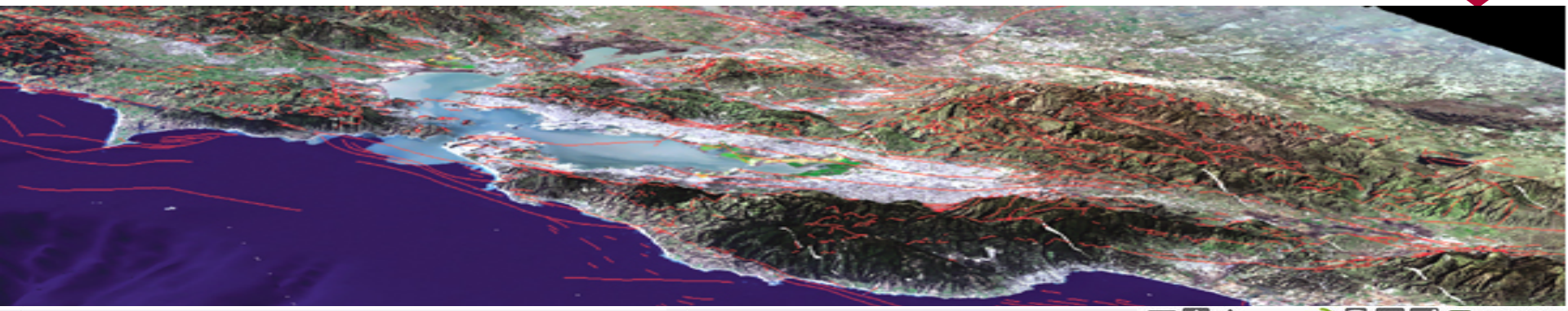
$5 billion, 25% of the annual sales

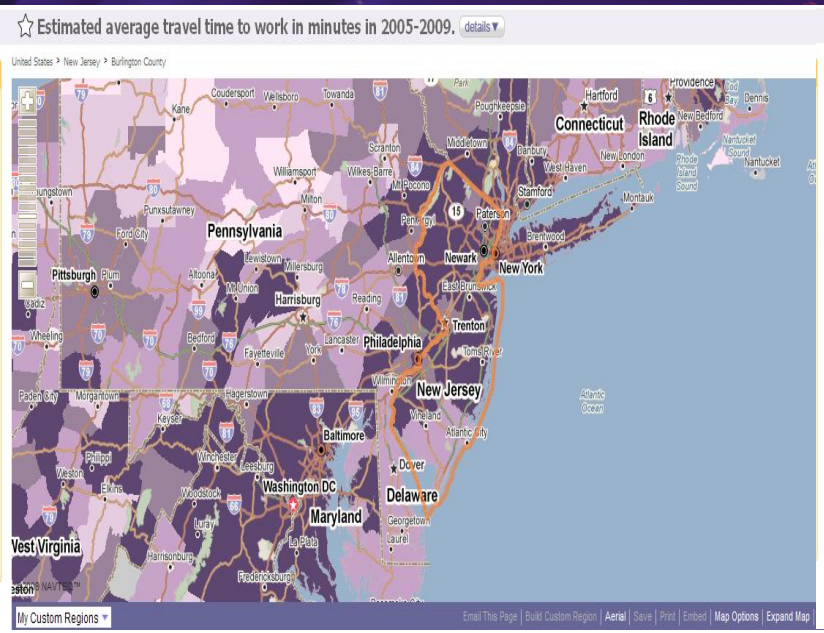
Climate/Healthcare Science

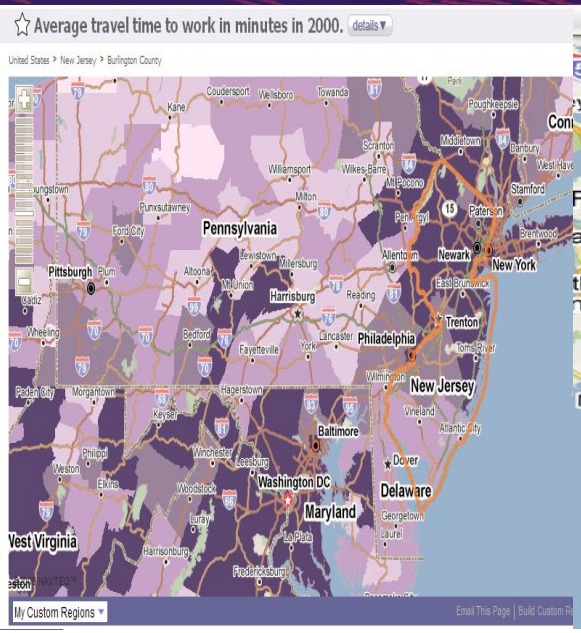
Healthcare

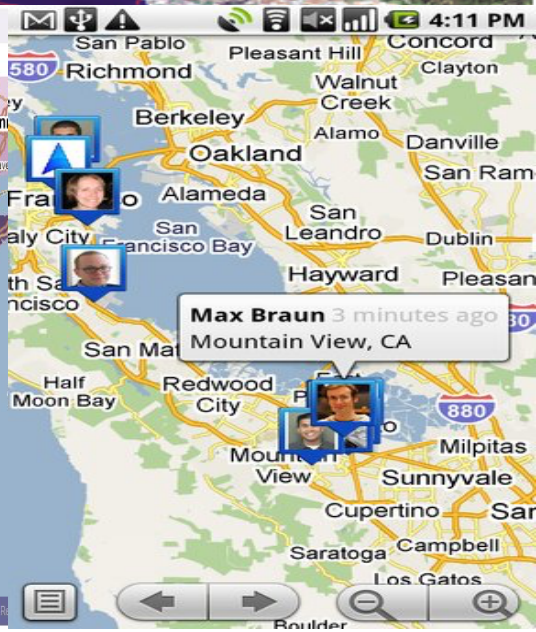
Climate Change

GIS and Location Science









Complexity of Geographic Information System:

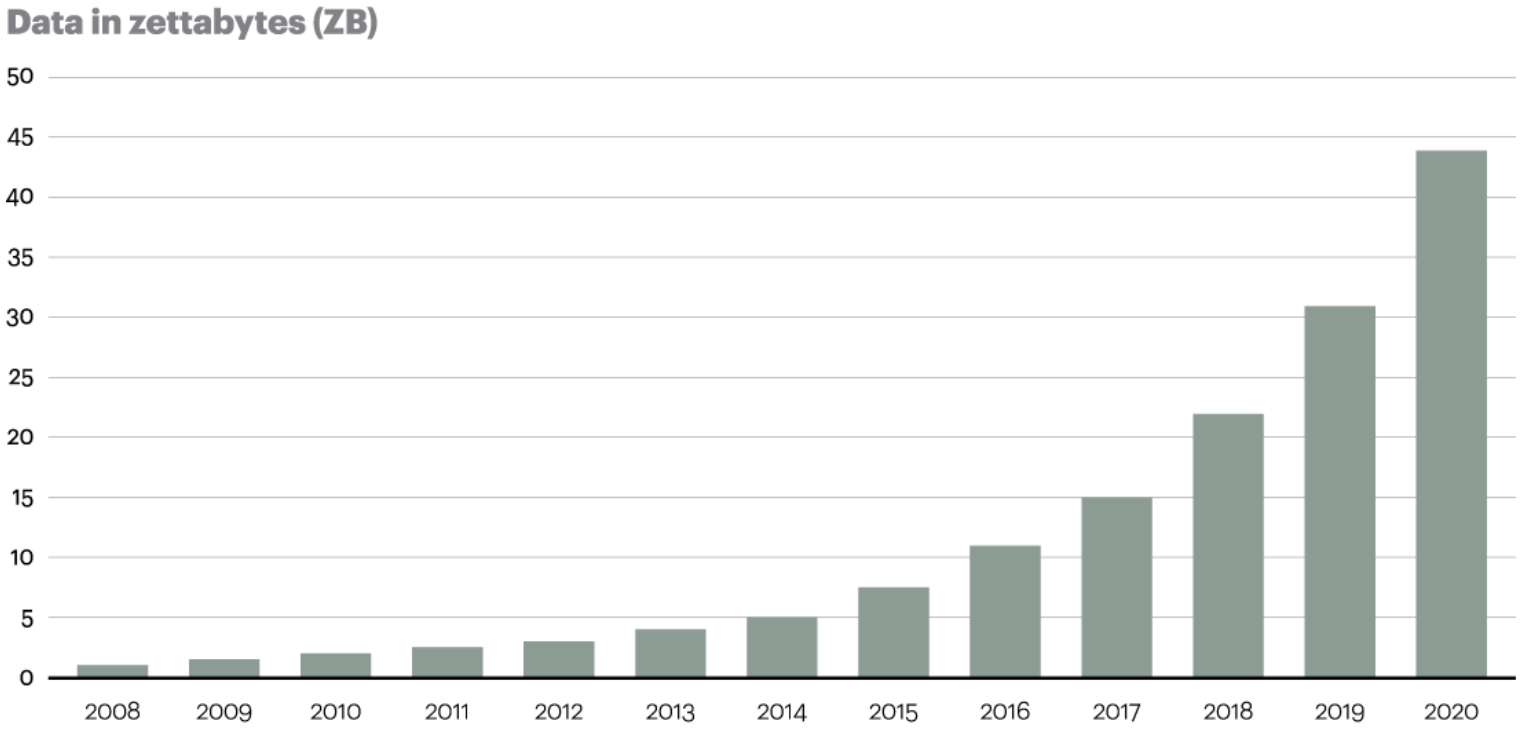
High dimension Dynamic

Spatial temporal Computation

Big Data Analytics

Figure 1

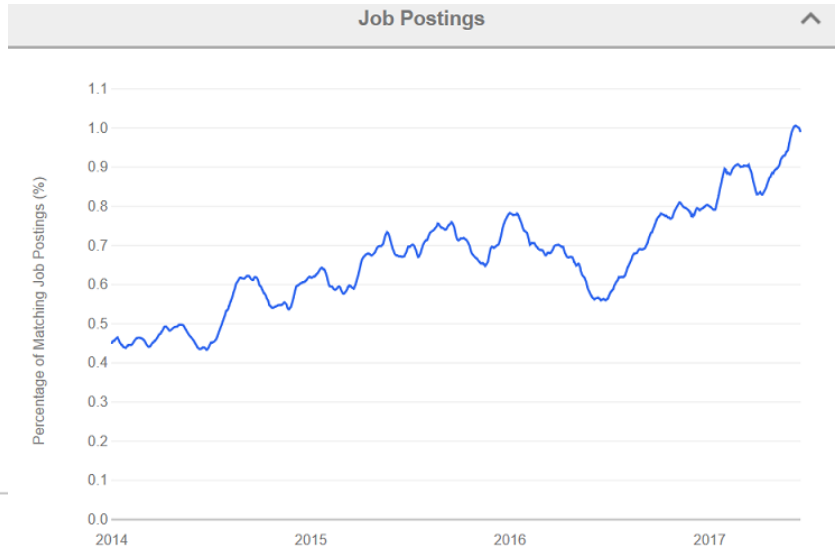
Data is growing at a 40 percent compound annual rate, reaching nearly 45 ZB by 2020



Source: Oracle, 2012

Machine Learning Jobs

"140,000 – 190,000 position shortage with deep expertise in statistics and machine learning" --- Mckinsey Big Data report



Challenges for Machine Learning

- Intelligent Systems
  - **Dynamic, incremental and adaptive learning algorithms;**
  - Limited understanding of intelligent behaviors;
  - **Lack of system level models and architectures to mimic brain-level intelligence;**
  - Embodiment of machine intelligence hardware within systems that learn through interaction with environment;
  - Systems design and implementation;

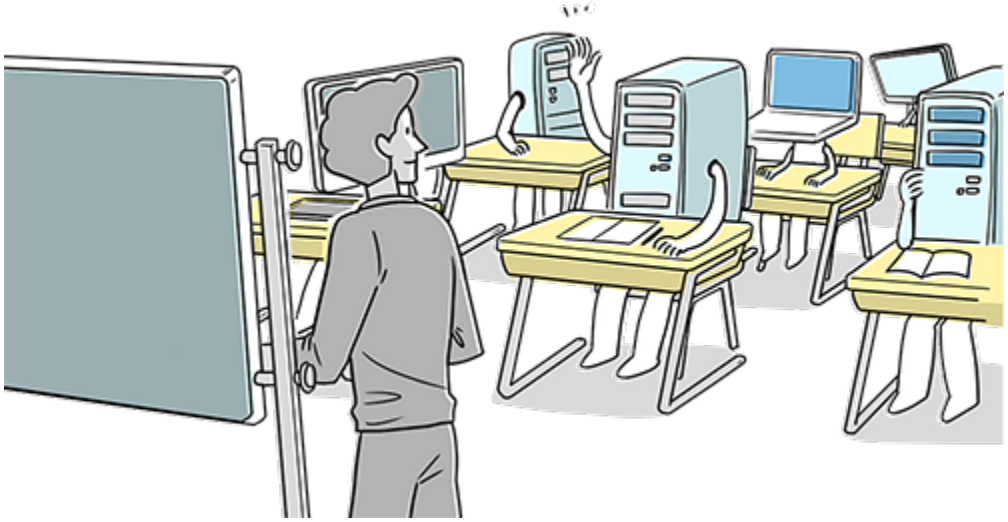
- Big Data
  - Platform/Storage/Index
  - Quality Uncertainty
  - **Efficient statistical/machine learning methods**
  - Intellectual property, liability
  - Visualization

Challenges for Machine Learning

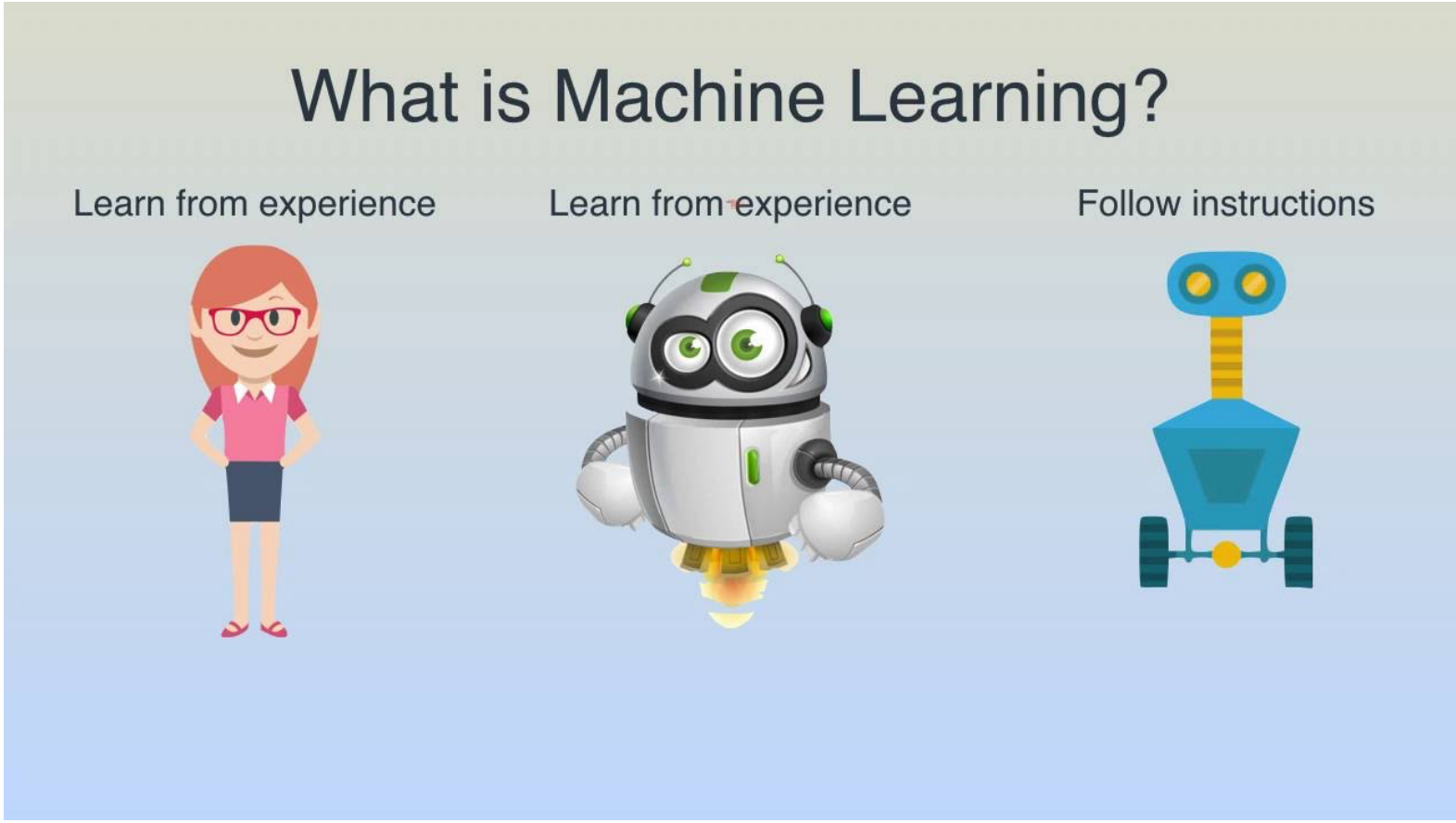
- Theoretical foundations
  - Some results based on optimization and probability models
  - Lack rigorous mathematical formulation especially for recent deep learning algorithms
  - Verifiability of the system is another problem

- Security, privacy and ethical issues
  - Model reliability under attacks, e.g., pollution attacks
  - Data privacy, model privacy
  - Ethical use of machine learning in controllable ways

Again, What is Machine Learning?



Simple Answer



Answer from Expert

"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."

--- Tom Mitchell

Why Learning?

Learning is used when:

- Human expertise does not exist (navigating on Mars),
- Humans are unable to explain their expertise (speech recognition)
- Solution changes in time
- Solution needs to be adapted to particular cases
- Data Volume are huge

Learning Associations

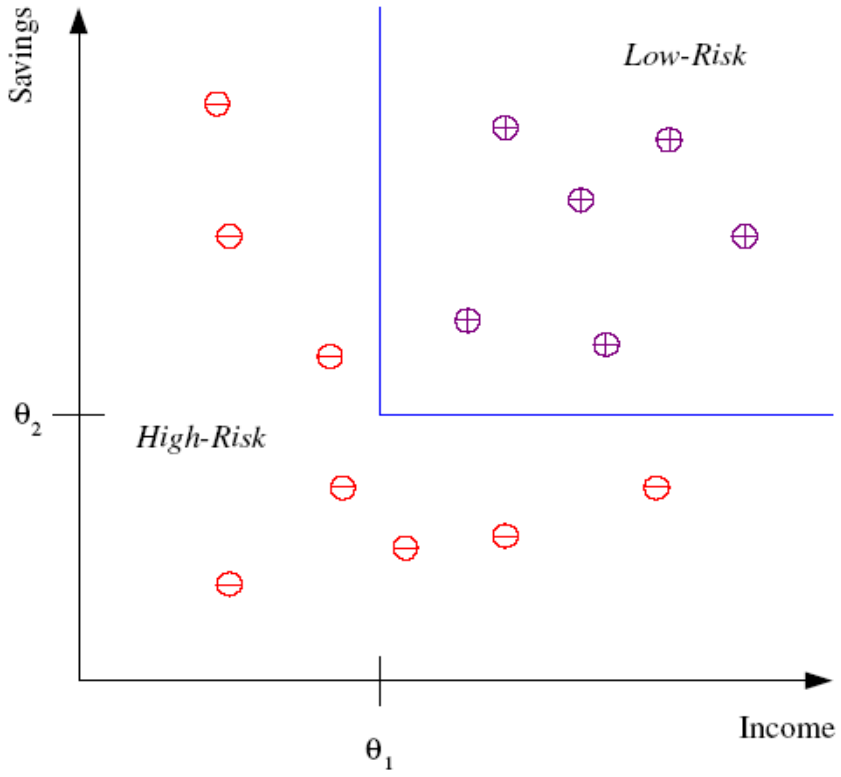
- Basket analysis:

  $P(Y \mid X)$ probability that somebody who buys X also buys Y where X and Y are products/services.

  Example: P (chips | beer) = 0.7

Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* > θ_1 AND *savings* > θ_2
THEN low-risk ELSE high-risk

Regression

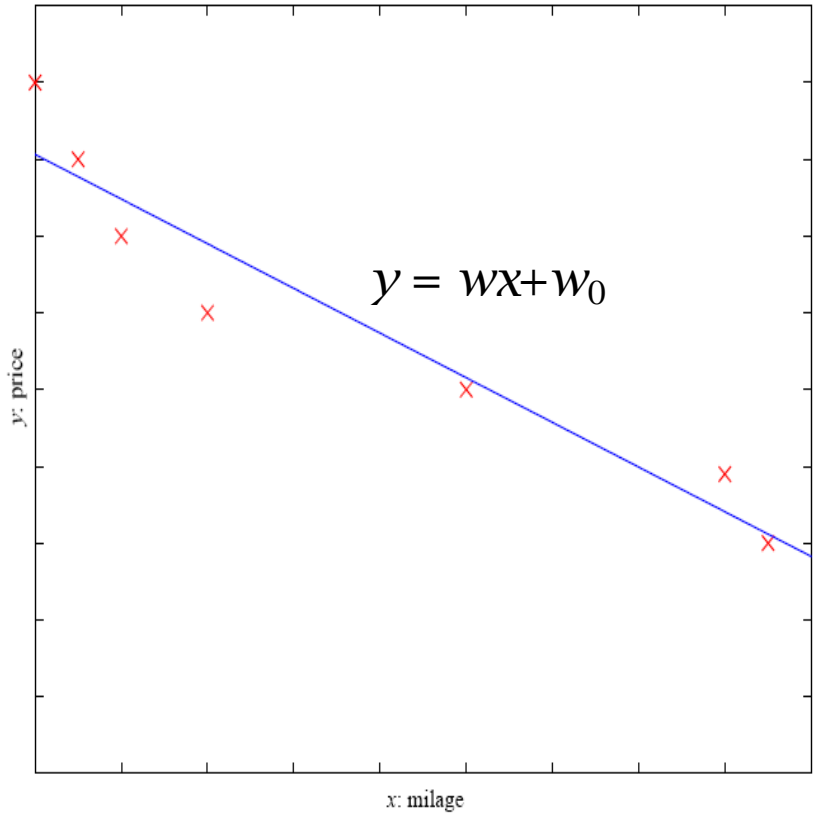
- Example: Price of used cars
- x : car attributes

  y : price

  $y = g(x \mid \theta)$

  $g(\ )$ model,

  θ parameters



Reinforcement Learning

- Learning a policy: A sequence of outputs
- No supervised output but delayed reward
- Credit assignment problem
- Game playing
- Robot in a maze
- Multiple agents, partial observability, ...

Categorization of Learning Systems

- Whether or not trained with human supervision
  - Supervised
    - k-NN, linear regression, logistic regression, support vector machines, decision trees and random forests, neural networks, etc.
  - Unsupervised
    - Clustering (e.g., k-means), dimensionality reduction (e.g., PCA), association rule learning, etc.
  - Semi-supervised
  - Reinforcement learning

Categorization of Learning Systems

- Whether or not learn incrementally on the fly
  - Online learning
  - Batch learning

Categorization of Learning Systems

- Whether compare new data to known data points, or build a predictive model
  - Instance-based learning
  - Model-based learning

Programming Tools

- Examples and assignments will use Python language. Jupyter notebook provides visualized interface.

- The following libraries are necessary for this course:

  Scikit-learn (http://scikit-learn.org), Numpy, Matplotlib, Pandas, …

- **Install Anaconda** (https://www.anaconda.com/download/#macos) first! It includes the whole package including Python and related libraries.

- The examples and code of Aurelien Geron's textbook can be found at https://github.com/ageron/handson-ml

Reference

The lecture notes in this lecture are based on the following textbooks:

T. M. Mitchell, Machine Learning, McGraw Hill, 1997. ISBN: 978-0-07-042807-2

C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006, ISBN: 978-0-387-31073-2.

E. Alpaydin, Introduction to Machine Learning, MIT Press, 2004, ISBN 0-262-01211-1

Part of the slide materials were based on Dr. Rong Duan's Fall 2016 course CPE/EE 695A Applied Machine Learning at Stevens Institute of Technology.

STEVENS

INSTITUTE *of* TECHNOLOGY

THE INNOVATION UNIVERSITY®

stevens.edu