CS6510
Applied Machine Learning

Course Introduction

5 Aug 2017

Vineeth N Balasubramanian

आई आई टी हैदराबाद
IIT Hyderabad

A few (not so, yet) recent quotes

- “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates, Chairman, Microsoft)
- “Machine learning is the next Internet” (Tony Tether, Director, DARPA)
- Machine learning is the hot new thing” (John Hennessy, President, Stanford)
- “Web rankings today are mostly a matter of machine learning” (Prabhakar Raghavan, ex-Dir. Research, Yahoo)
- “Machine learning is going to result in a real revolution” (Greg Papadopoulos, CTO, Sun)
- “Machine learning is today’s discontinuity” (Jerry Yang, ex-CEO, Yahoo)

What is Machine Learning?

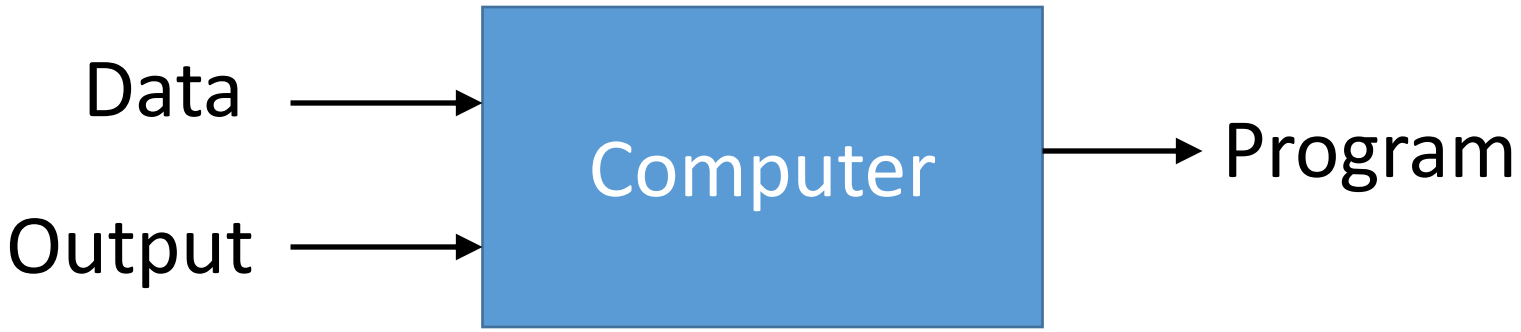
What is Machine Learning?

- Making predictions or decisions from data
- “Programming computers to optimize a performance criterion using example data or past experience” (Ethem Alpaydin, Machine Learning, 2010)
- “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, Machine Learning, 1997)
- “Learning general models from a data of particular examples”
- “Build a model that is *a good and useful approximation* to the data.“

Today

Traditional Programming

Machine Learning



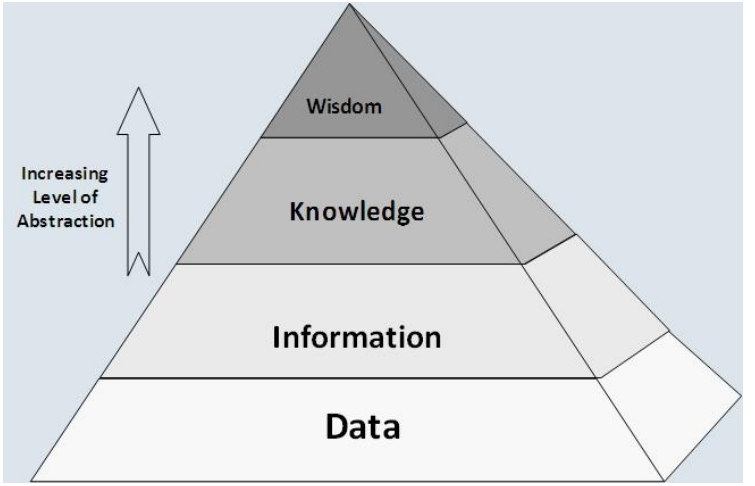
Source: Domingos

Related Terms

Machine Learning, Data Mining, Knowledge Discovery, Artificial Intelligence, Statistical Learning, Pattern Recognition, Computational Learning

When is Machine Learning Used?

- 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
- Solution needs to be adapted to particular cases
  - E.g. user biometrics
- Data is cheap and abundant; knowledge is expensive and scarce



Applications of Machine Learning

From: cheapsales@buystufffromme.com
To: ang@cs.stanford.edu
Subject: Buy now!

Deal of the week! Buy now!
Rolex w4tchs - $100
Med1cine (any kind) - $50
Also low cost M0rgages
available.

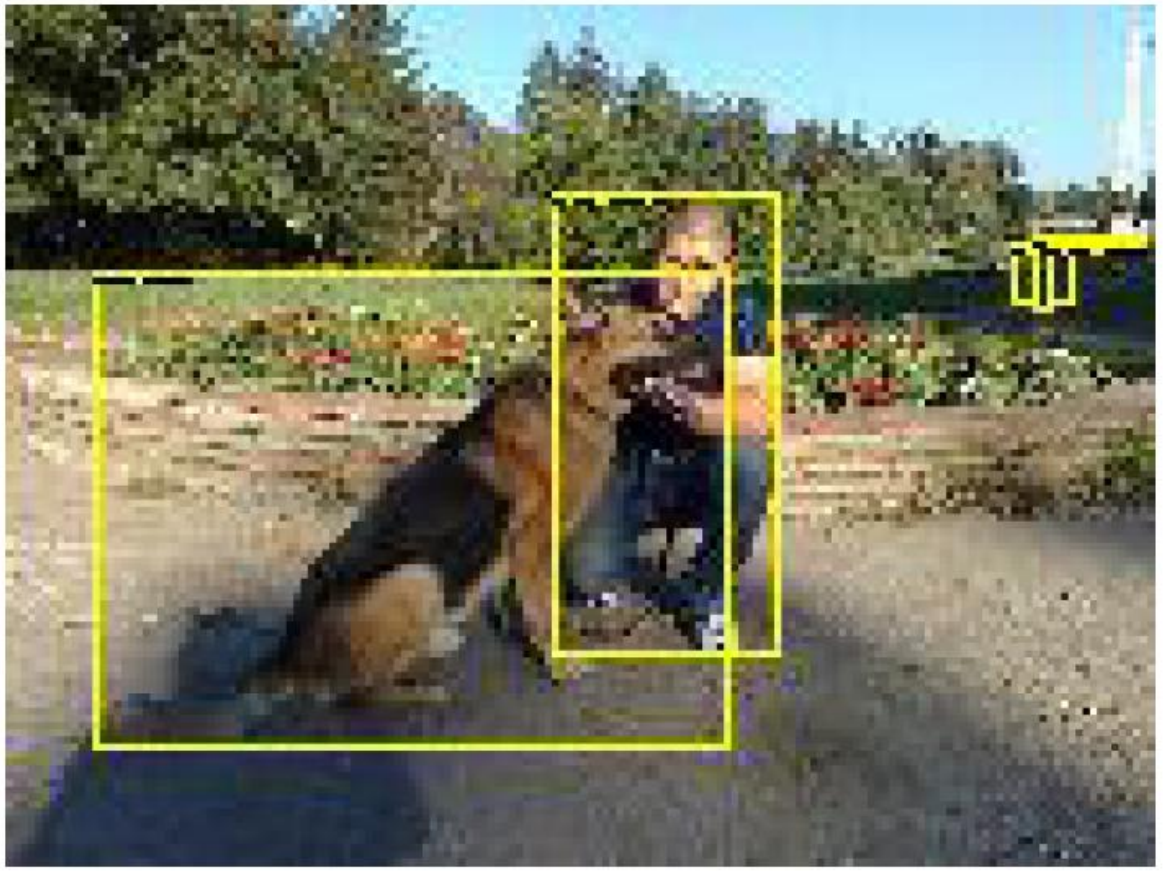
Spam

From: Alfred Ng
To: ang@cs.stanford.edu
Subject: Christmas dates?

Hey Andrew,
Was talking to Mom about plans
for Xmas. When do you get off
work. Meet Dec 22?
Alf

Non-spam

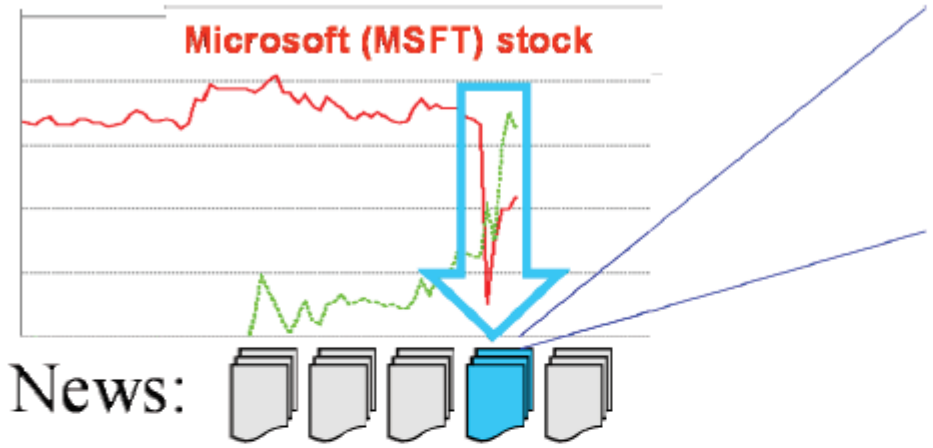
Applications of Machine Learning



Applications of Machine Learning

```
CCAGCTGCATCACAGGAGGCCAGCGAGCAGGTCTGTTCCAAGGGCCTTCGAGCCAGTCTG EI
GAGGTGAAGGACGTCCTTCCCCAGGAGCCGGTGAGAAGCGCAGTCGGGGGCACGGGGATG EI
TAAATTCTTCTGTTTGTTAACACCTTTCAGACTTATGTGTATGAAGGAGTAGAAGCCAAA IE
AAACTAAAGAATTATTCTTTTACATTTCAGTTTTTCTTGATCATGAAAACGCCAACAAAA IE
AAAGCAGATCAGCTGTATAAACAGAAAATTATTCGTGGTTTCTGTCACTTGTGTGATGGT N
TTGCCCTCAGCATCACCATGAACGGAGAGGCCATCGCCTGCGCTGAGGGCTGCCAGGCCA N
```

Applications of Machine Learning



Software giant Microsoft saw its shares dip a few percentage points this morning after U.S. District Judge Thomas Penfield Jackson issued his "findings of fact" in the government's ongoing antitrust case against the Seattle wealth-creation machine...

Words like **Jackson** and **antitrust** are more likely in the stories preceding the plunge.

P (**shares**) = 0.074
P (**antitrust**) = 0.009
P (**judge**) = 0.006
P (**trading**) = 0.032
P (**against**) = 0.025
P (**Jackson**) = 0.001

P (**shares** | MSFT↓) = 0.071
P (**antitrust** | MSFT↓) = 0.044
P (**judge** | MSFT↓) = 0.039
P (**trading** | MSFT↓) = 0.029
P (**against** | MSFT↓) = 0.027
P (**Jackson** | MSFT↓) = 0.025

$$P(\text{MSFT}\downarrow \mid \textbf{Jackson}) = P(\textbf{Jackson} \mid \text{MSFT}\downarrow)\ P(\text{MSFT}\downarrow) / P(\textbf{Jackson})$$

Applications of Machine Learning

More ML Applications

- Science (Astronomy, neuroscience, medical imaging, bio-informatics)
- Environment (energy, climate, weather, resources)
- Retail (Intelligent stock control, demographic store placement)
- Manufacturing (Intelligent control, automated monitoring, detection methods)
- Security (Intelligent smoke alarms, fraud detection)
- Marketing (promotions, ...)
- Management (Scheduling, timetabling)
- Finance (credit scoring, risk analysis...)
- Web data (information retrieval, information extraction, ...)

More Recent ML Applications

- AlphaGo!
- Automating Employee Access Control
- Identifying whales in ocean based on audio recordings
- Predict wait times for patients in emergency rooms
- Extract heart failure diagnosis criteria from free-text physician notes
- Predicting hospital readmissions
- Is (s)he a pyschopath?

Source: http://www.forbes.com/sites/85broads/2014/01/06/six-novel-machine-learning-applications/#6b6f9a9e67bf

When are ML algorithms not needed?

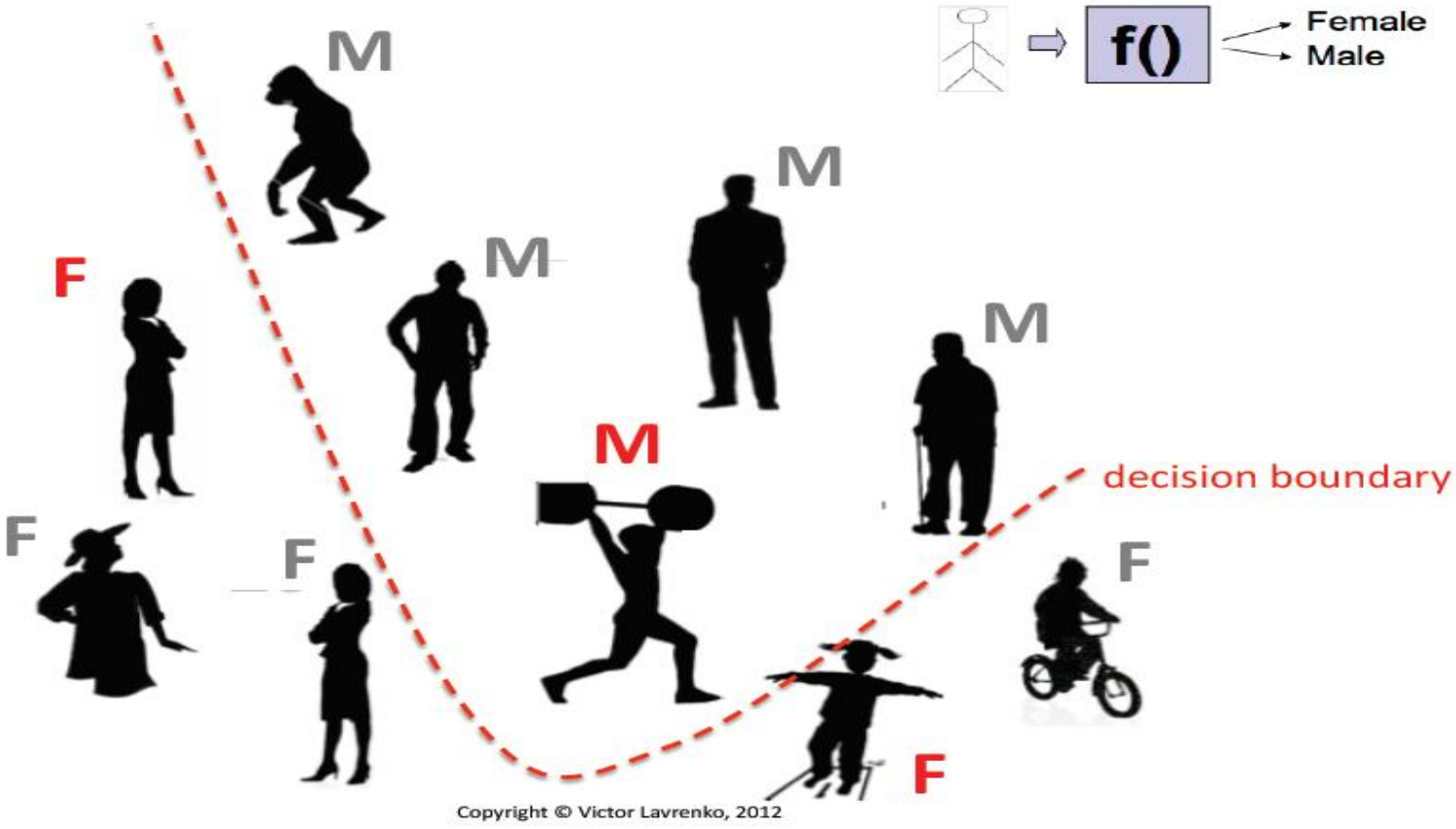
- When the relationships between all system variables (input, output, and hidden) is completely understood!

- This is NOT the case for almost any real system!

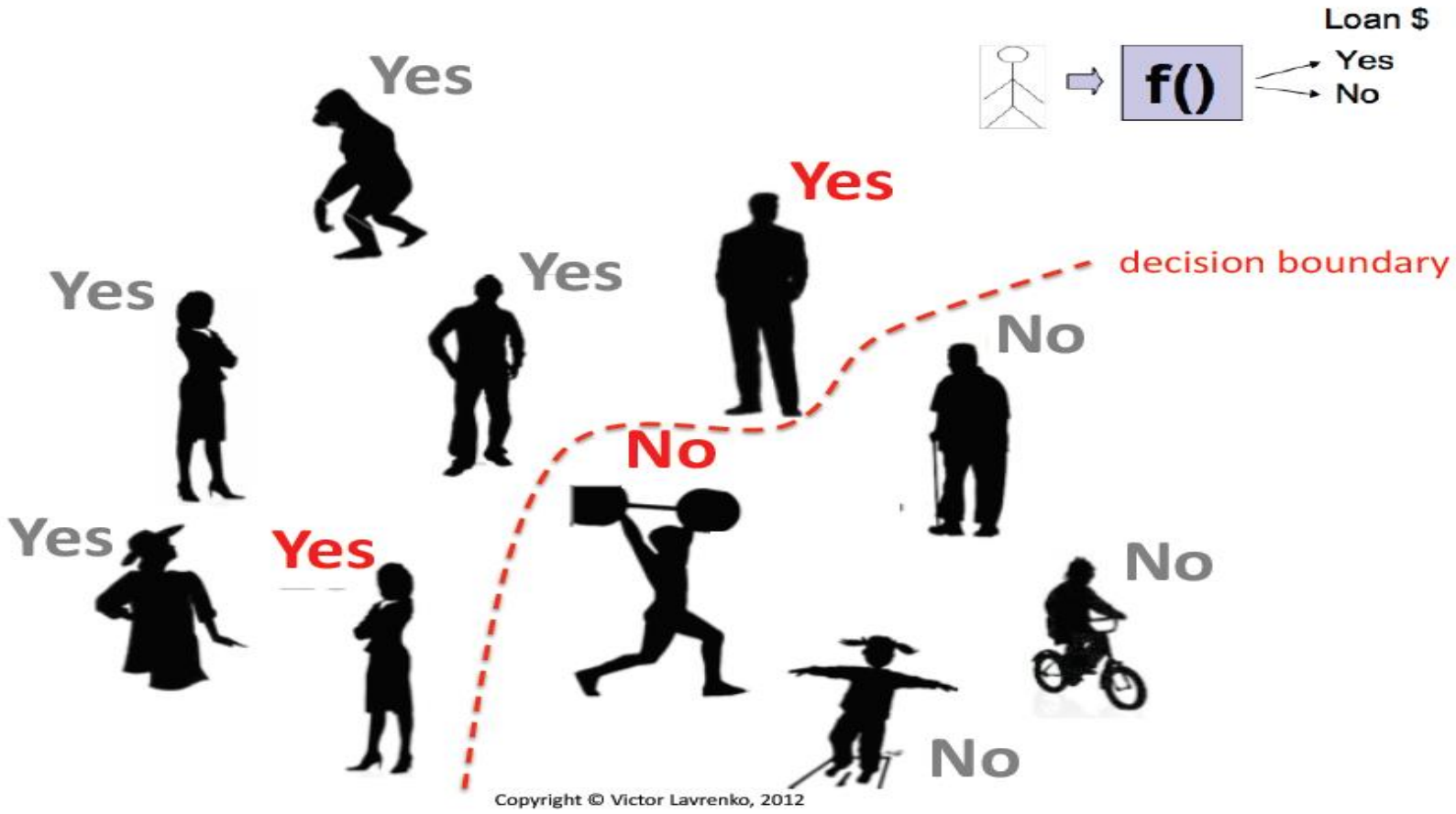
Overview of ML

- Supervised learning
  - Predict an output y when given an input **x**
  - For categorical y : classification.
  - For real-valued y : regression.
- Unsupervised learning
  - Create an internal representation of the input, e.g. clustering, dimensionality
  - This is important in machine learning as getting labels is often difficult and expensive
- Other settings of ML
  - Reinforcement learning (learning from “rewards”)
  - Semi-supervised learning (combines supervised + unsupervised)
  - Active learning, Transfer learning, Structured prediction

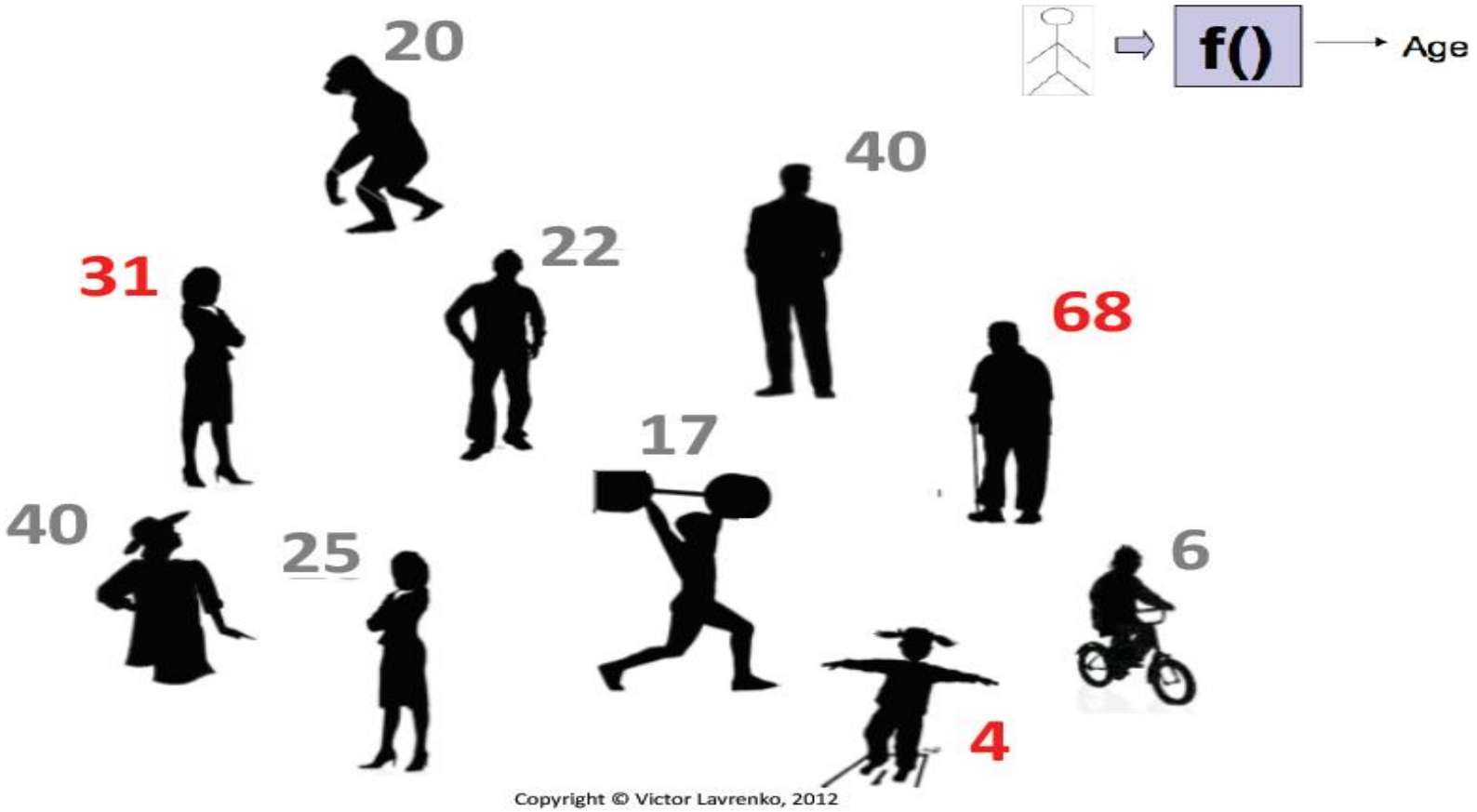
Classification (Supervised Learning)



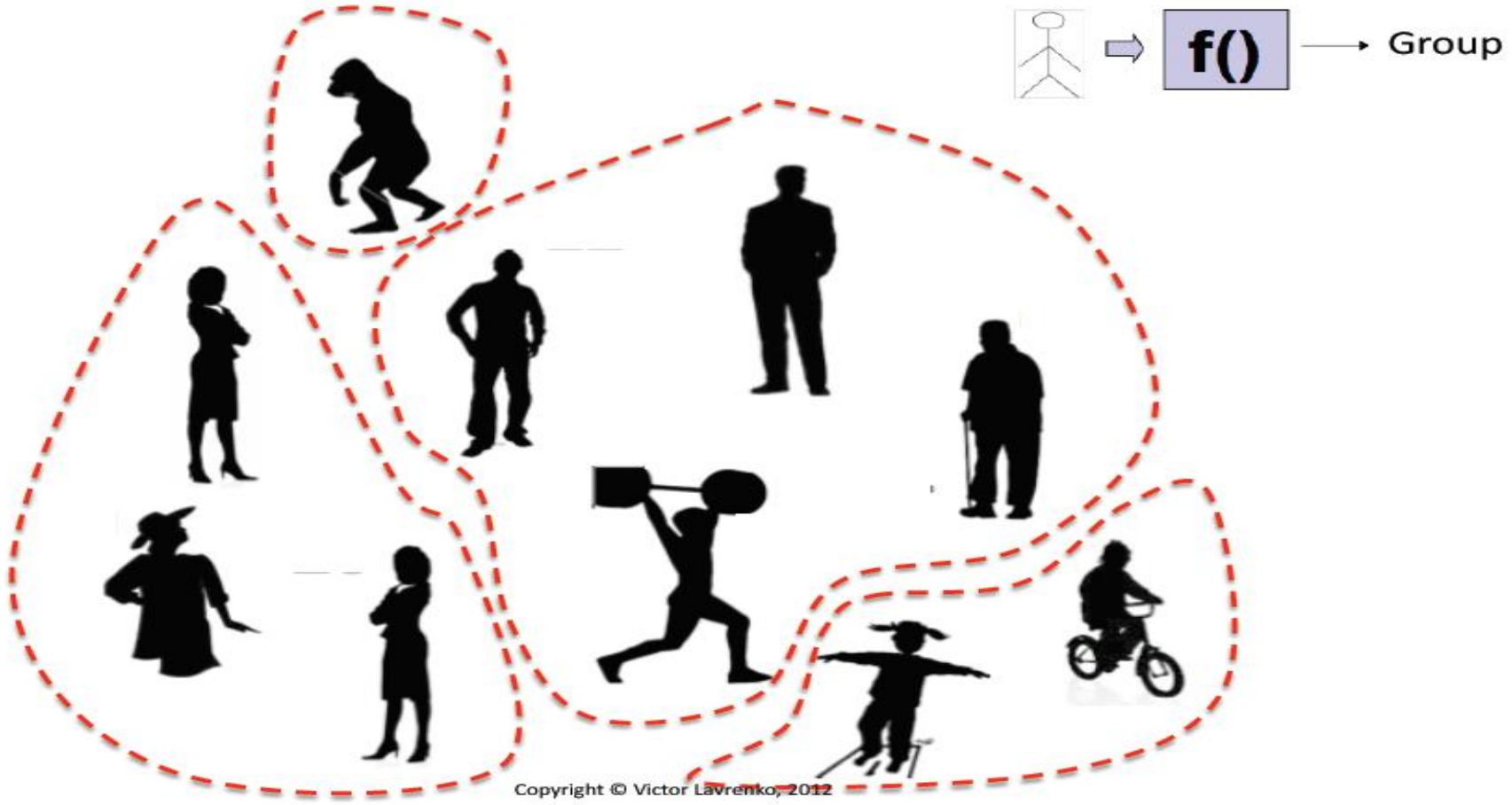
Classification (Supervised Learning)



Regression (Supervised Learning)

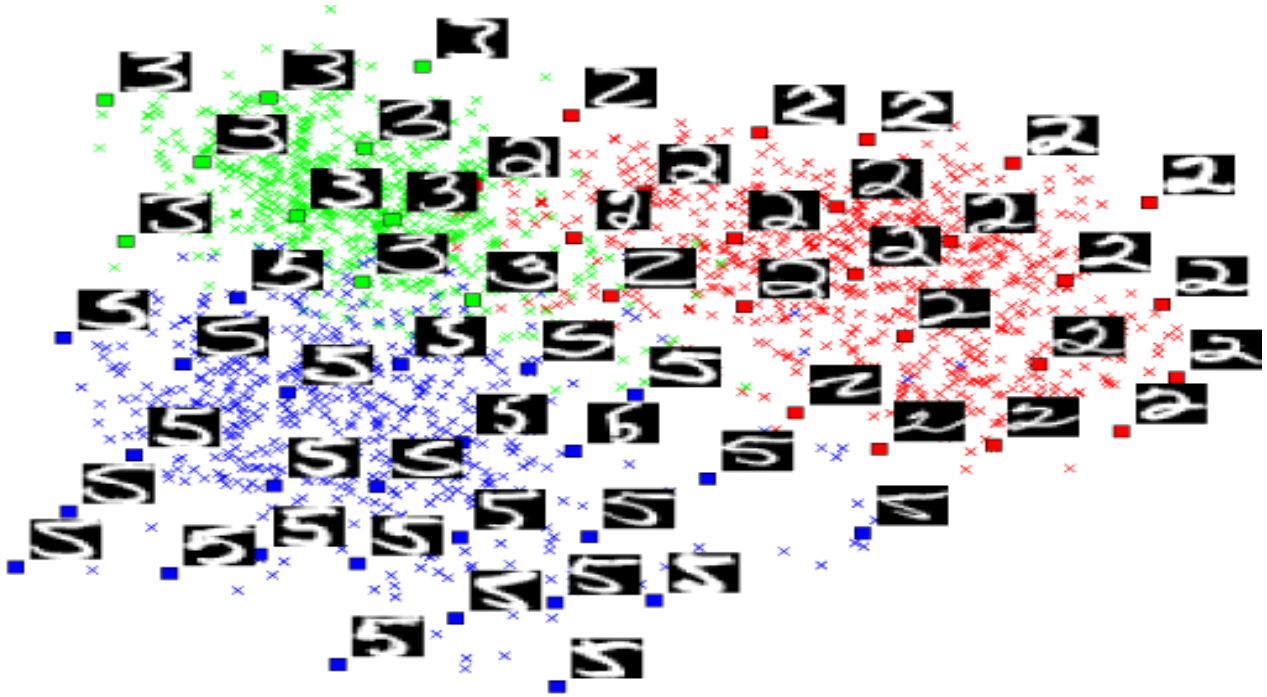


Clustering (Unsupervised Learning)



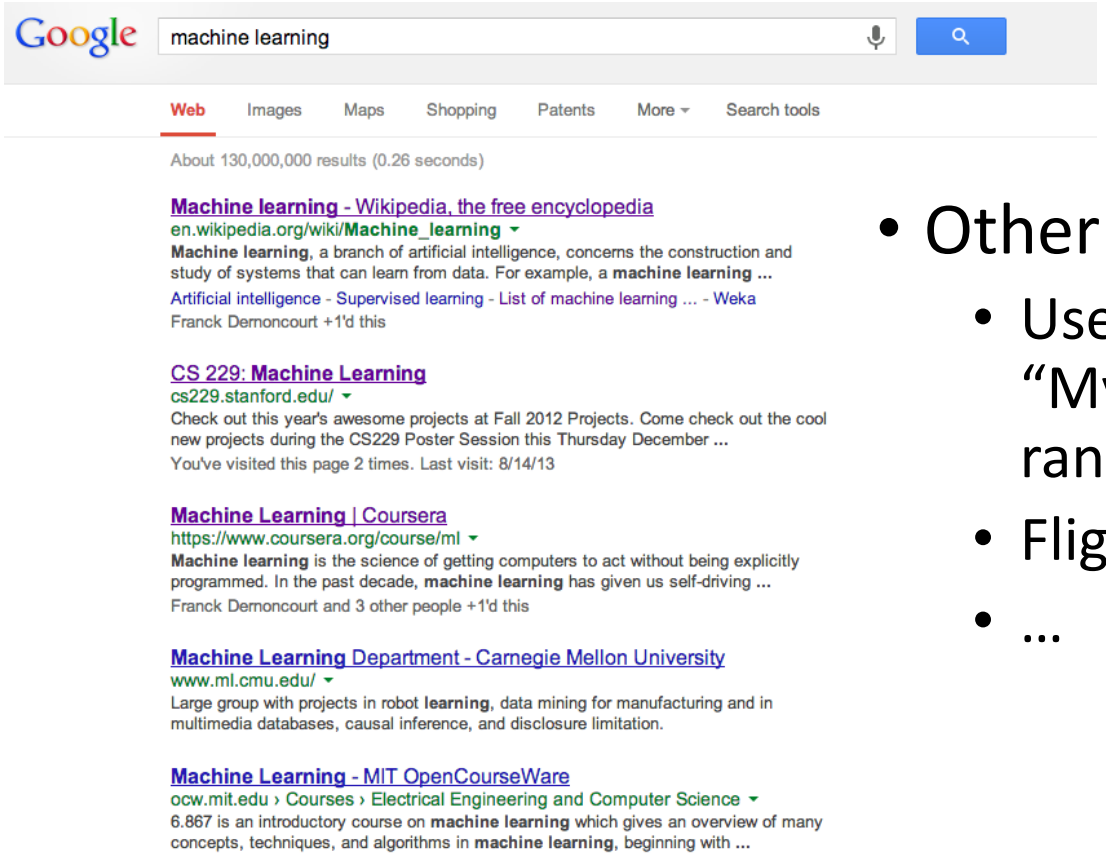
Dimensionality Reduction (Unsupervised Learning)

- Large sample size is required for high-dimensional data
- Query accuracy and efficiency degrade rapidly as the dimension increases
- Strategies
  - Feature reduction
  - Feature selection
  - Manifold learning
  - Kernel learning



Other Settings: Ranking (Supervised Learning)

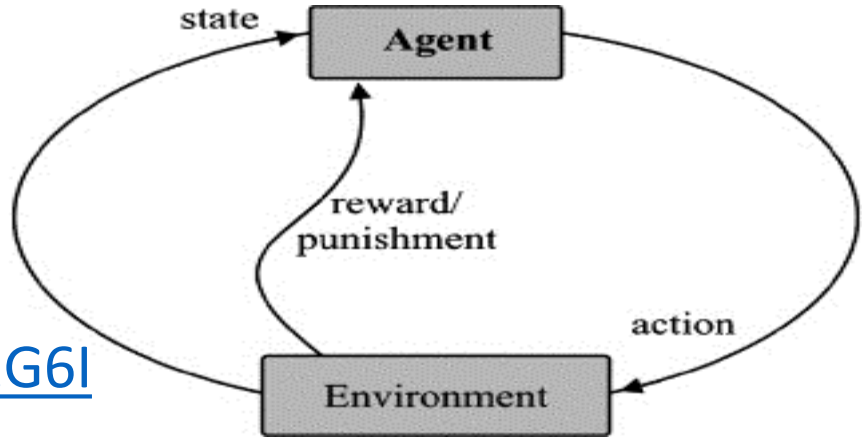
Given a query and a set of web pages, rank them according to relevance



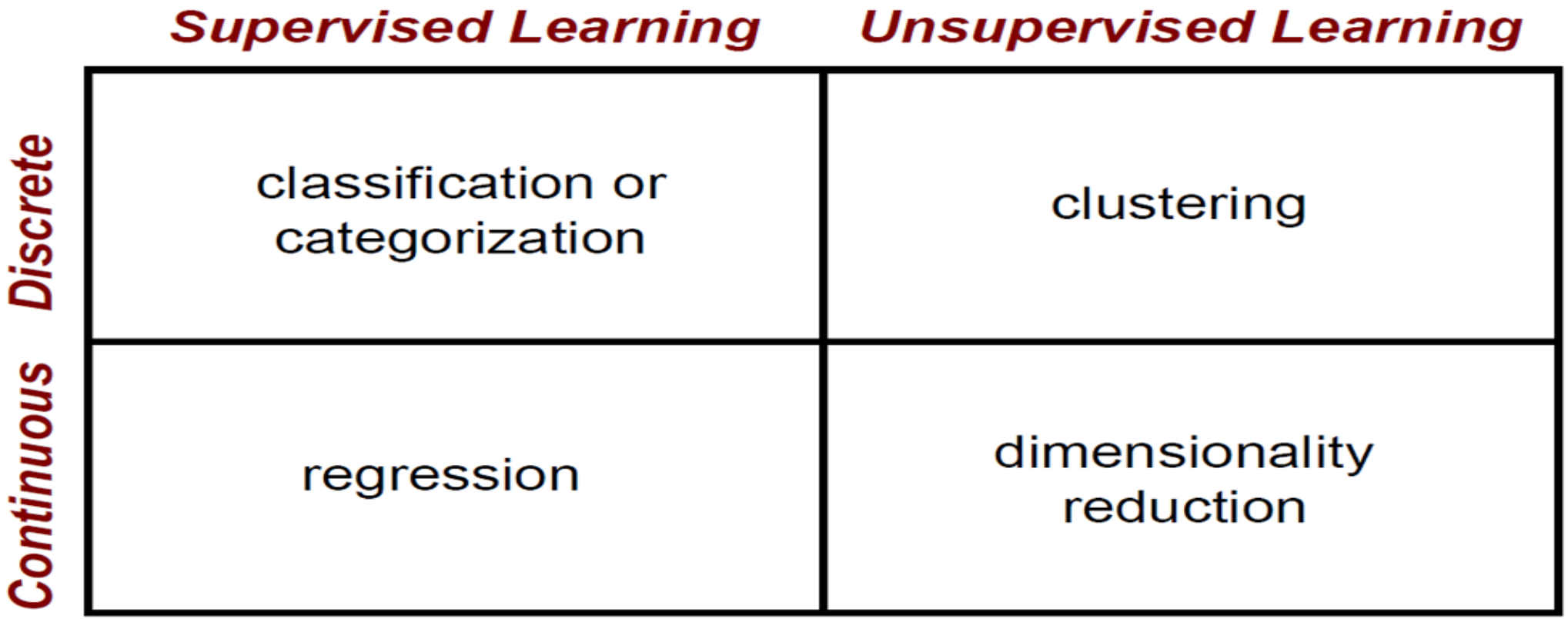
- Other applications
  - User preference, e.g. Netflix "My List" -- movie queue ranking
  - Flight search (search in general)
  - …

Other Settings: Reinforcement Learning

- Learning a policy: A sequence of outputs
- No supervised output but delayed reward
  - E.g. Game playing
  - E.g. Robot in a maze
- Multiple agents, partial observability, ...
- Example (Simple Demo):
  - https://www.youtube.com/watch?v=DCjbk4m1G6I



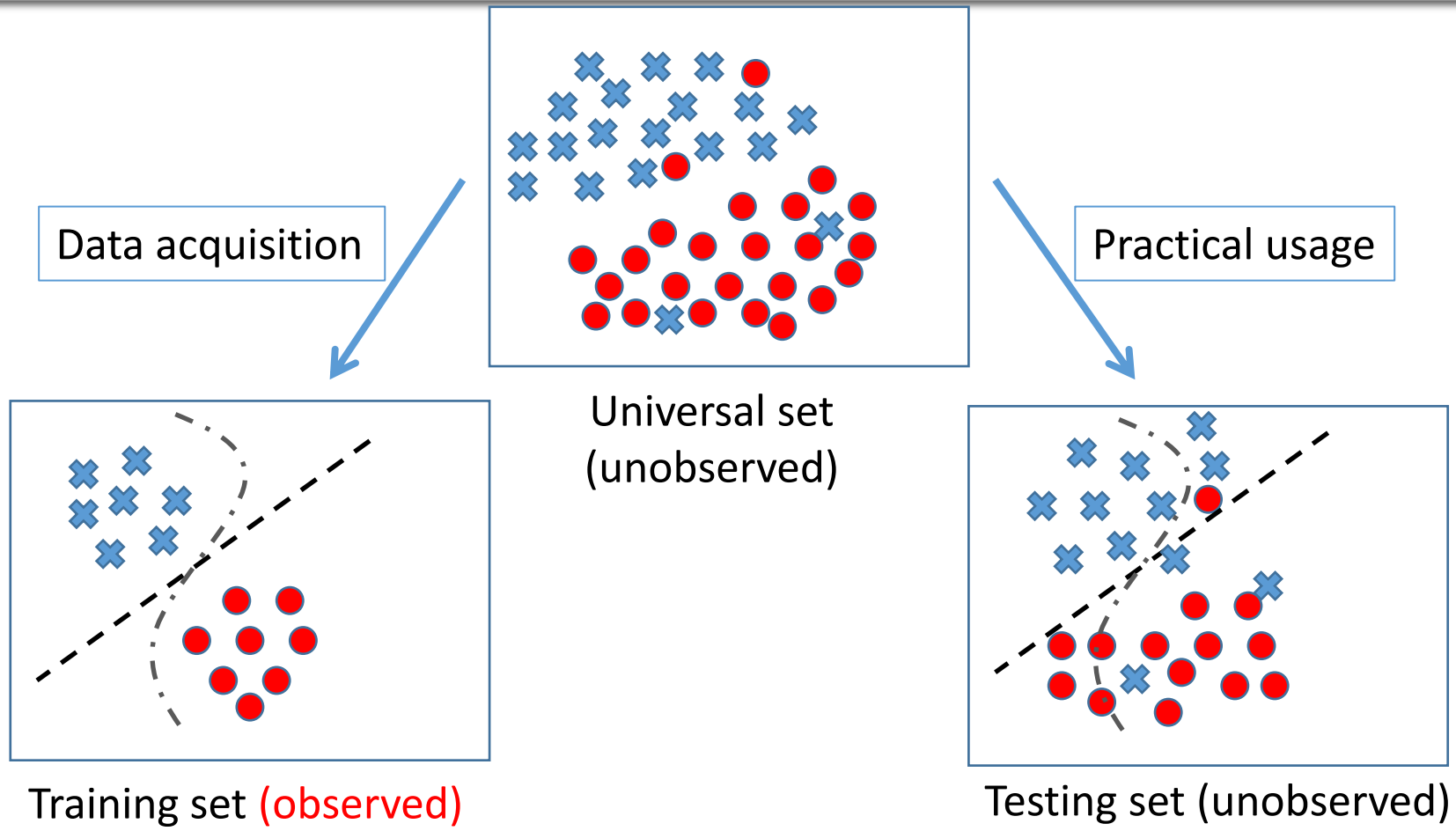
ML Problems


Supervised Learning
Unsupervised Learning
Discrete
Continuous
classification or categorization
clustering
regression
dimensionality reduction

ML in Practice

- Understanding domain, prior knowledge, and goals
- Data integration, selection, cleaning, pre-processing, etc.
- Learning models
- Interpreting results
- Consolidating and deploying discovered knowledge
- Loop

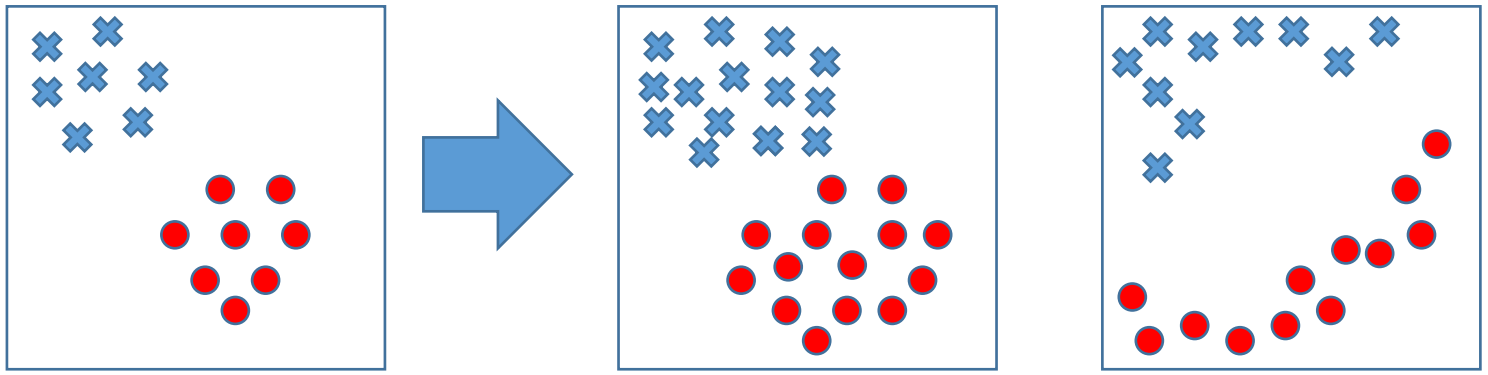
Training and Testing ML Models



Training and Testing ML Models

- Training is the process of making the system able to learn.

- **No free lunch** rule:
  - Training set and testing set may not come from the same distribution
  - Need to make some assumptions or bias



Types of Models

- Inductive vs Transductive Learning
- Online vs Offline Learning
- Generative vs Discriminative Models
- Parametric vs Non-Parametric Models

ML Datasets

- UCI Repository: http://www.ics.uci.edu/~mlearn/MLRepository.html
- Statlib: http://lib.stat.cmu.edu/
- Kaggle
- Many more…

ML Resources

- MOOCs
  - Coursera, EdX, Udacity
- Conferences/Journals
  - JMLR, Machine Learning, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Pattern Analysis and Machine Intelligence, Annals of Statistics
  - ICML, NIPS, KDD, IJCAI, AAAI, ICDM

Mathematical Basis

- Functions, Logarithms and Exponentials
- Vectors, Dot Products, Orthogonality
- Matrices, Matrix Operations, Linear Transformations, Eigendecomposition
- Calculus, Differentiation, Integration
- Probability and Statistics
- Functional Analysis, Hilbert Spaces

Course Details

- Available at: https://goo.gl/q6dNoP

Topics not to be covered (likely)

- Deep Learning (not in its “depth”, at least)
- Bayesian Networks (not in depth)
- Reinforcement Learning
- Learning Theory

Programming

- Python
- Libraries
  - Numpy, Scipy – numerical/scientific computing, linear algebra
  - Matplotlib – for plotting
  - Scikitlearn – for machine learning

Foundational Reading/Follow-up

- Math
  - Part 1 of Deep Learning book: http://www.deeplearningbook.org/
  - Essence of linear algebra: http://youtu.be/kjBOesZCoqc
  - Essence of calculus: https://goo.gl/Hnk1jA
- Programming
  - Learn Python
    - https://try.jupyter.org/
    - https://docs.python.org/3/tutorial/
    - Video Tutorials: https://www.youtube.com/watch?v=cpPG0bKHYKc
  - Play with Numpy