COMS 4030A/7047A Adaptive Computation and Machine Learning

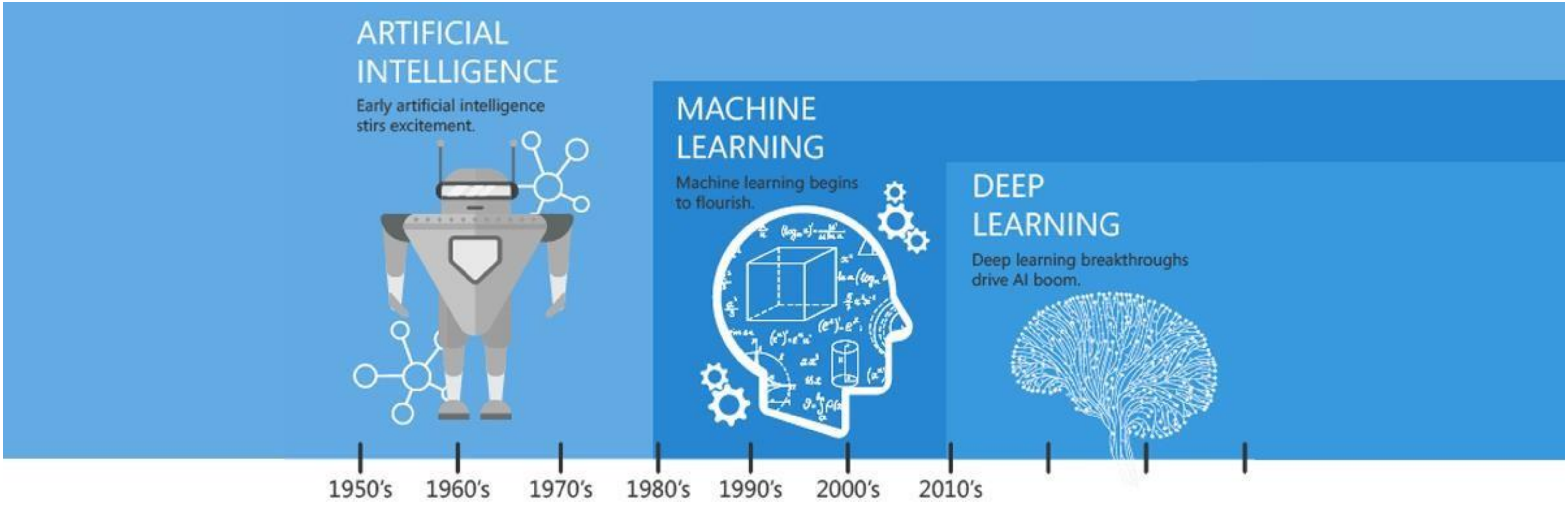
Hima Vadapalli

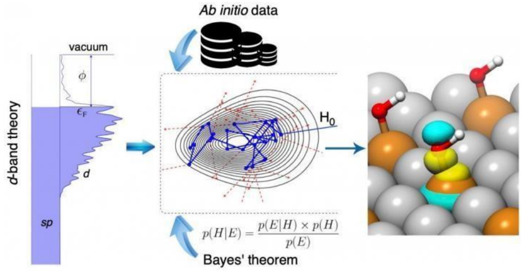
Semester I, 2022

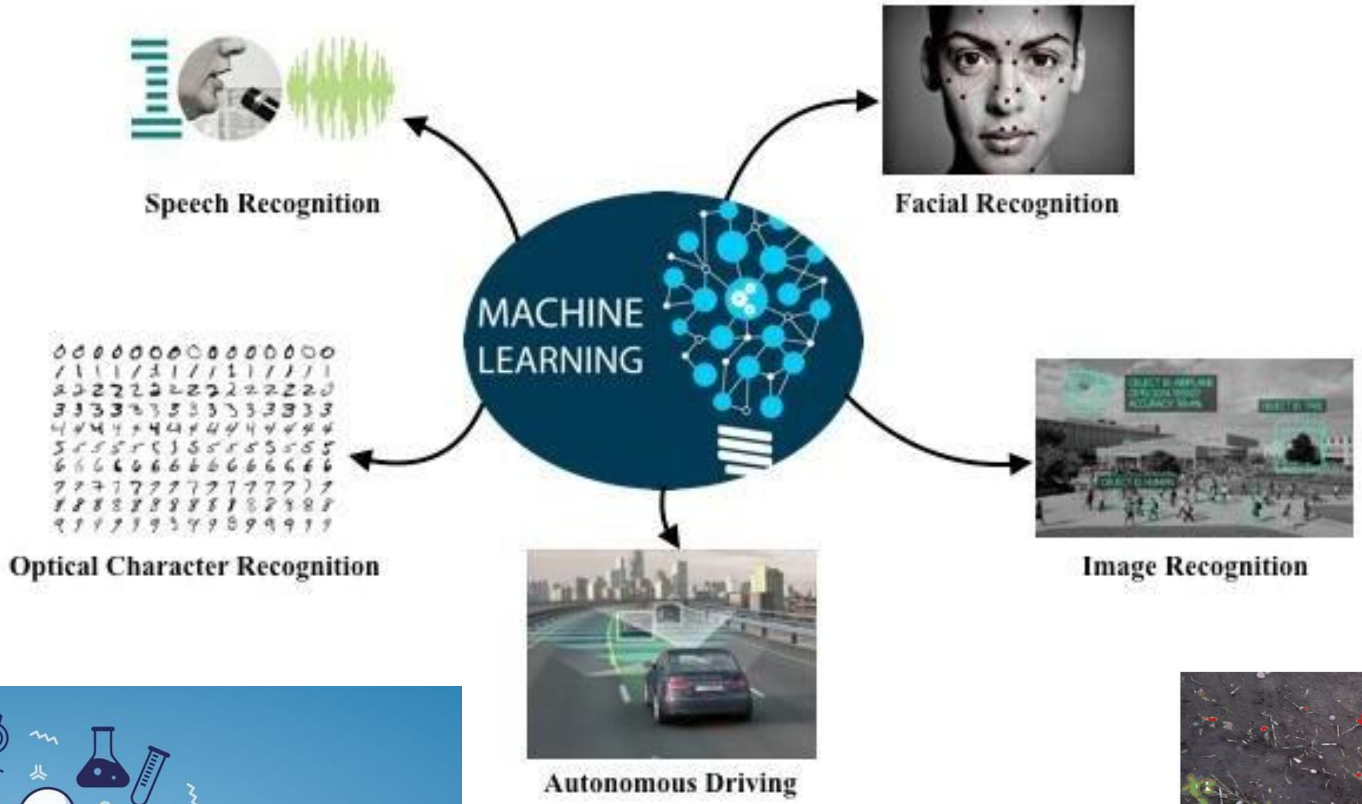
Adaptive Computation and Machine Learning

Course Logistics

Introduction


ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
MACHINE LEARNING
Machine learning begins to flourish.
DEEP LEARNING
Deep learning breakthroughs drive AI boom.
1950's
1960's
1970's
1980's
1990's
2000's
2010's


Ab initio data
vacuum
d-band theory
sp
d
H0
p(H|E) = p(E|H) × p(H) / p(E)
Bayes' theorem


Speech Recognition
Facial Recognition
MACHINE
LEARNING
Optical Character Recognition
Image Recognition
Autonomous Driving


Lesson 1


melissa
weeds

Course Objective

To provide a broad overview of the machine learning paradigms along with an in-depth review of specific learning algorithms in supervised, unsupervised, and reinforcement learning categories.

Learning Outcomes:

- *Good understanding of the fundamental issues and challenges of machine learning*
- *Good understanding of the strengths and weaknesses of many popular machine learning approaches*
- *Appreciation of the underlying mathematical relationships within and across Machine Learning algorithms and the paradigms of supervised and unsupervised learning*
- *Practical experience of designing and implementing various machine learning algorithms in a range of real-world applications*

Course Prerequisites:

- Familiarity with Python programming
- Familiarity with basic linear algebra
- Familiarity with basic calculus
- Familiarity with basic probability theory

Course breakdown:

- Introduction to Machine Learning
- Supervised Learning
- Unsupervised Learning   (Semi-supervised Learning)
- Computational Learning Theory
- Reinforcement Learning
- Introduction to Deep Learning

Course events:

- Lectures
- Labs/tutorials

Course sites:

[COMS4030A - Adaptive Computation and Machine Learning-2021-SM1 (wits.ac.za)]

[COMS7047A - Adaptive Computation and Machine Learning-2021-SM1 (wits.ac.za)]

Assessments:

- Class test (20%)
- Assignments (20%)
- Open ended Project (25%)
- Final exam (35%)

Resources

Textbooks

Machine Learning

Authors: Tom Mitchell
Publisher: McGraw Hill, 1997

Authors repository
http://www.cs.cmu.edu/~tom/mlbook.html

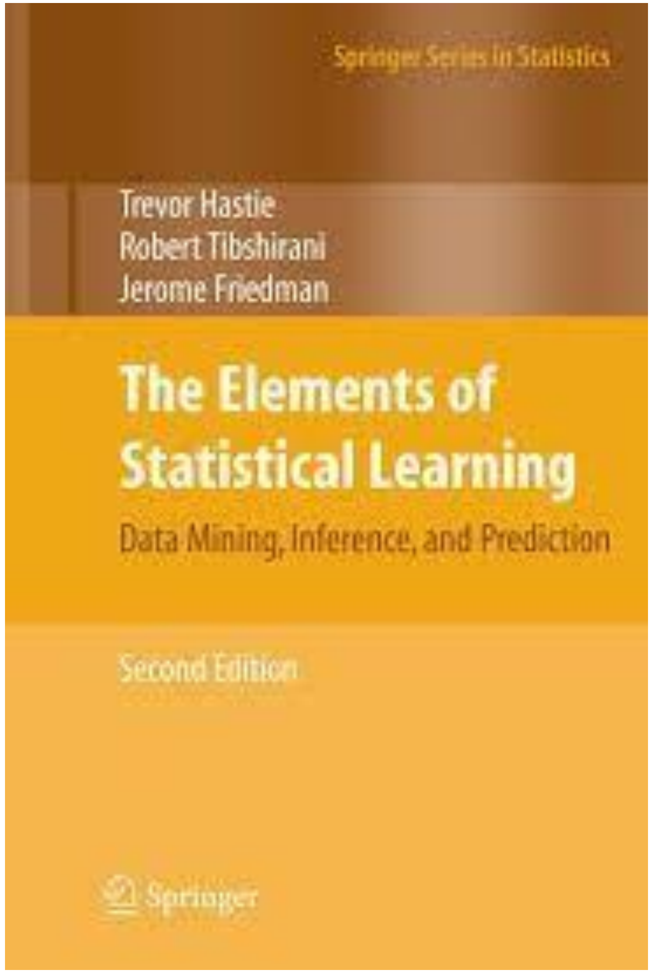
Resources

Textbooks

The Elements of Statistical Learning

Authors: Trevor Hastie, Robert Tibshirani,
Jerome Friedman
Publisher: Springer, 2009

Author's repository
https://web.stanford.edu/~hastie/Papers/ESLII.pdf



Resources

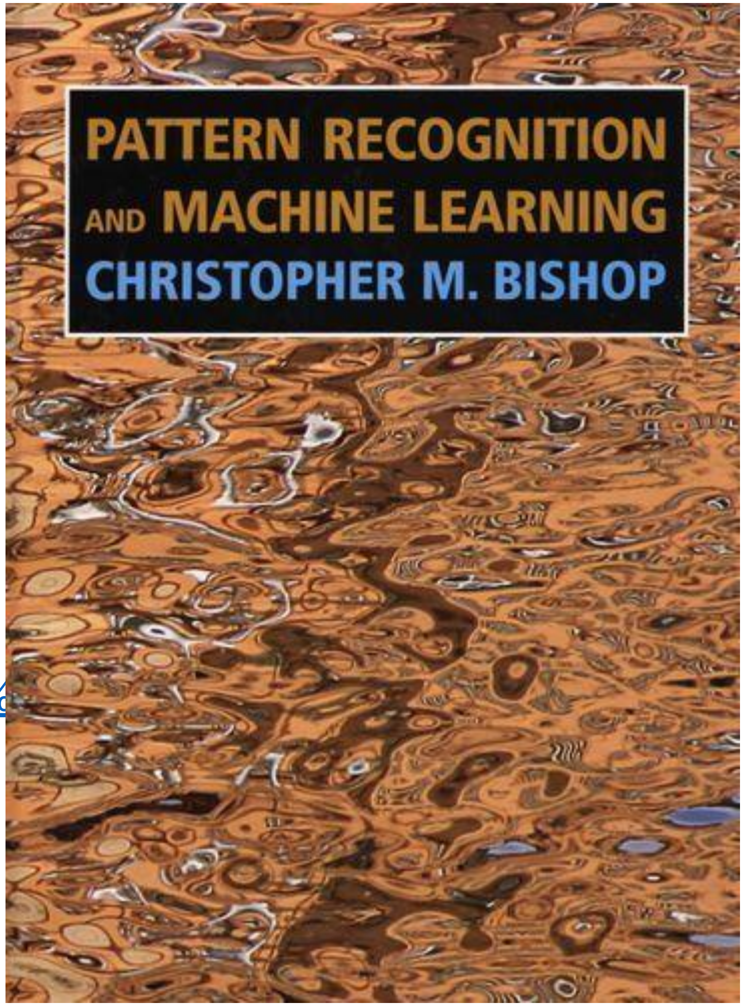
Textbooks

Pattern Recognition and Machine Learning

Authors: Christopher M. Bishop

Publisher: Springer, 2006

Online Copy
http://users.isr.ist.utl.pt/~wurmd/Livros/school/Bishop%20-%20Pattern%20Recognition%20And%20Machine%20Learning%



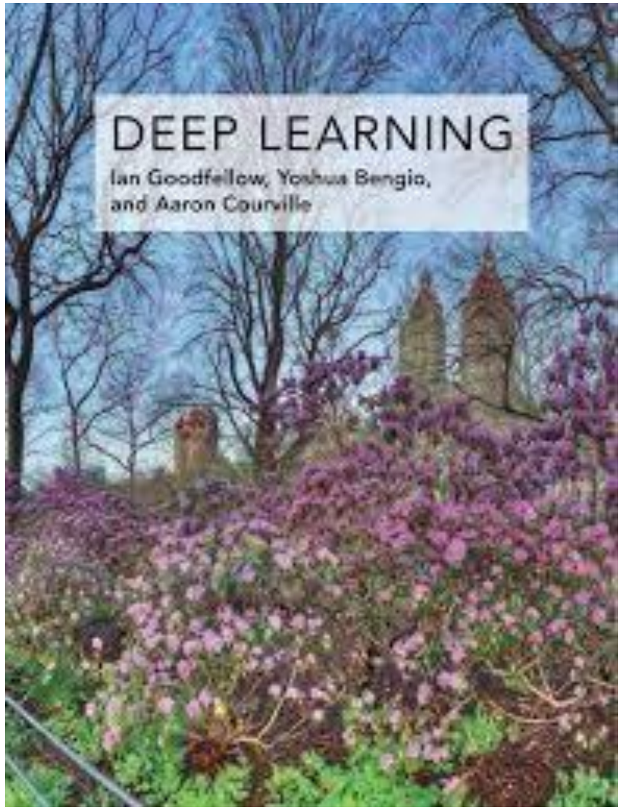
Resources

Textbooks

Deep Learning

Authors: Ian Goodfellow, Yoshua Bengio and Aaron Courville
Publisher: MIT Press, 2016

Author's repository
https://www.deeplearningbook.org/



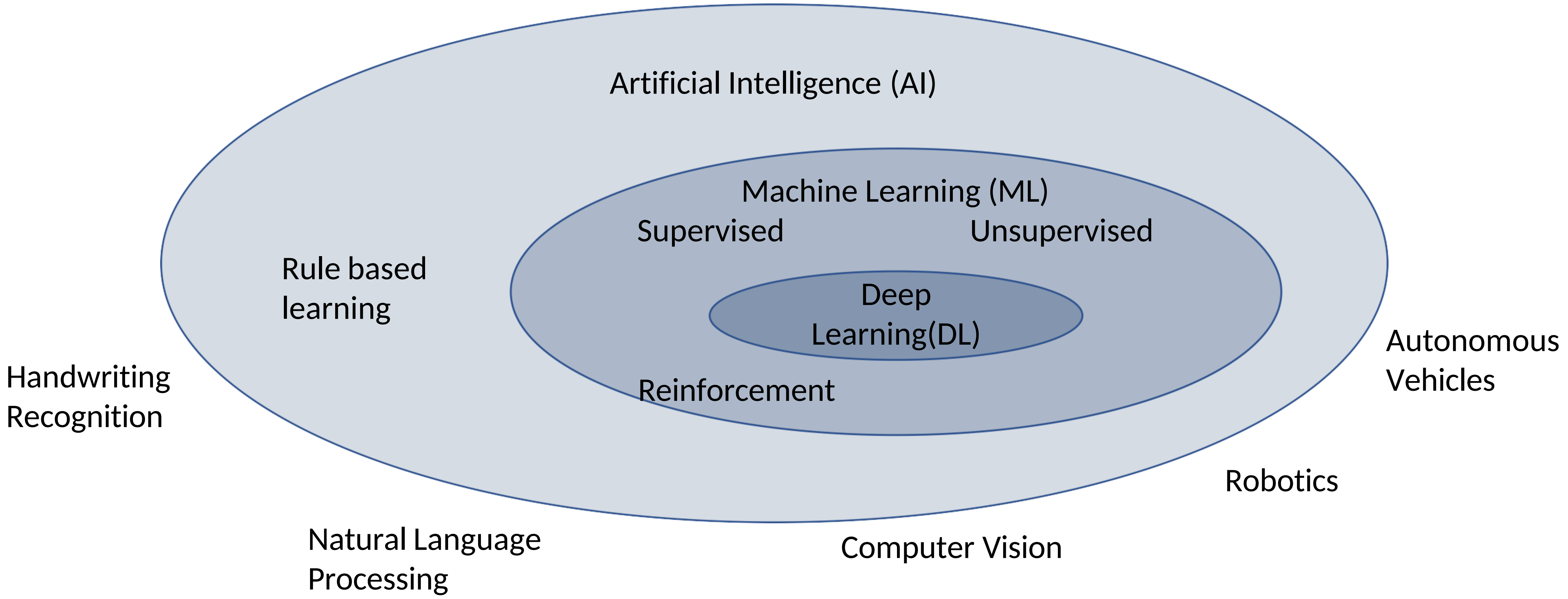
Questions?

Post them on the discussion forum

Machine Learning - Introduction

Slides based heavily on course material by Moses Charikar, Tengyu Ma, Chris Re and Andrew Ng

Machine Learning - Introduction



What is machine learning?

- Arthur Samuel (1959)

***Machine learning*: "**Field of study that gives computers the ability to learn without being explicitly programmed**"**

  - Samuels wrote a checkers playing program
    - Had the program play 10000 games against itself
    - Work out which board positions were good and bad depending on wins/losses

- Tom Michel (1999)

***Well posed learning problem*:** *"*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."**

  - The checkers example,
    - E = 10000s games
    - T is playing checkers
    - P if you win or not

Learning Algorithms

- Several types of learning algorithms
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning

- In this course, we
  - Look at practical advice as to **when** to apply learning algorithms (when?)
  - Learning a set of tools and **how** to apply them (how?)

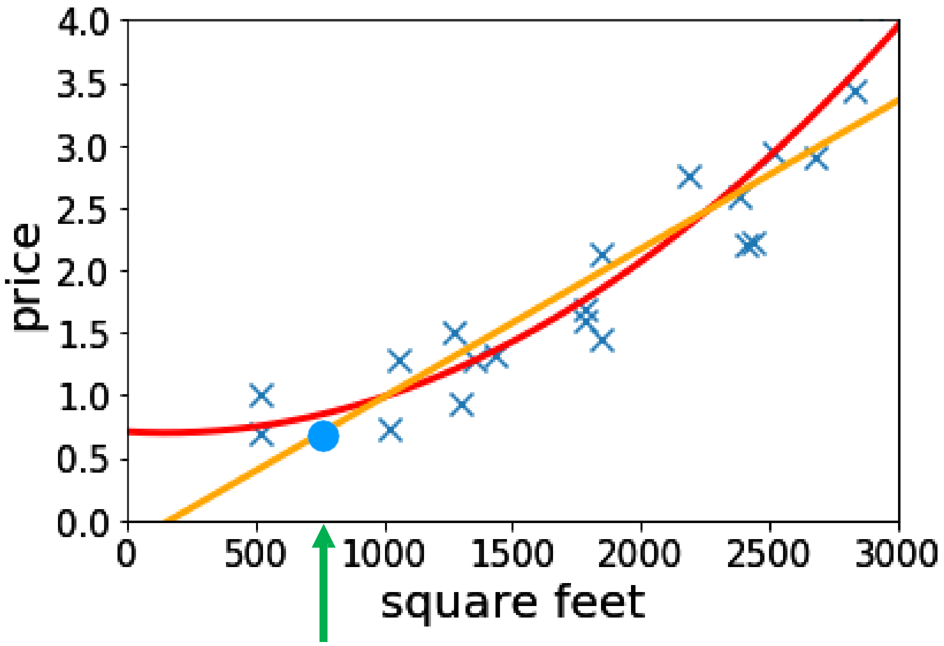
Supervised Learning

Teach the computer how to do something, then let it use it; use new found knowledge to do it

Housing Price Prediction

- Given: a dataset that contains n samples
  $$\left(x^{(1)}, y^{(1)}\right), \ldots \left(x^{(n)}, y^{(n)}\right);; \quad x \rightarrow sq\,ft, y \quad \rightarrow price$$
- Task: if a residence has x square feet, predict its price?

$x = 800$
$y = ?$

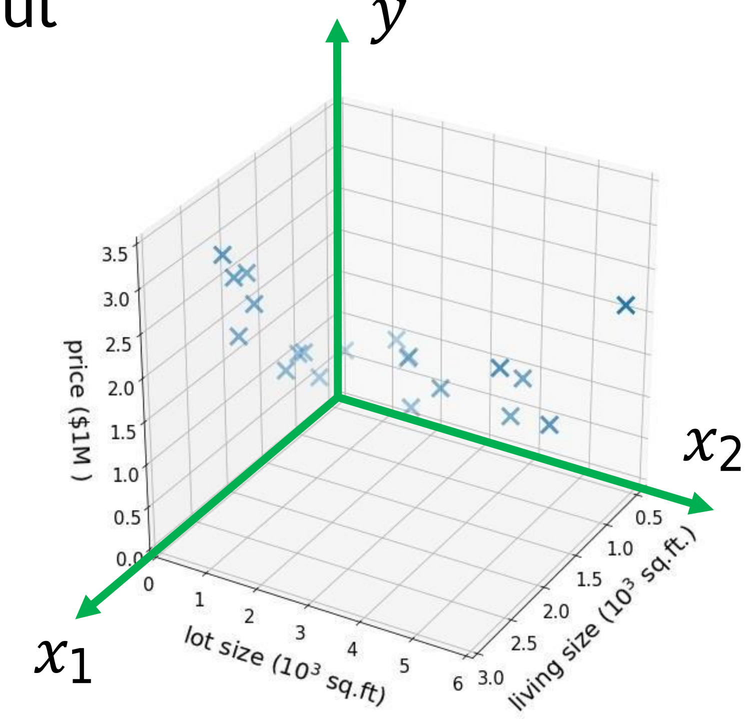


- next lectures: fitting linear/quadratic functions to the dataset

More Features

- Suppose we also know the lot size
- Task: find a function that maps

$$\underbrace{\text{(size, lot size)}}_{\substack{\text{features/input} \\ x \in \mathbb{R}^2}} \quad \rightarrow \quad \underbrace{\text{price}}_{\substack{\text{label/output} \\ y \in \mathbb{R}}}$$



- Dataset: $(x^{(1)}, y^{(1)}), \ldots, (x^{(n)}, y^{(n)})$

where $x^{(i)} = (x_1^{(i)}, x_2^{(i)})$

- "Supervision" refers to $y^{(1)}, \ldots, y^{(n)}$

High-dimensional Features

- $x \in \mathbb{R}^d$ for large d
- E.g.,

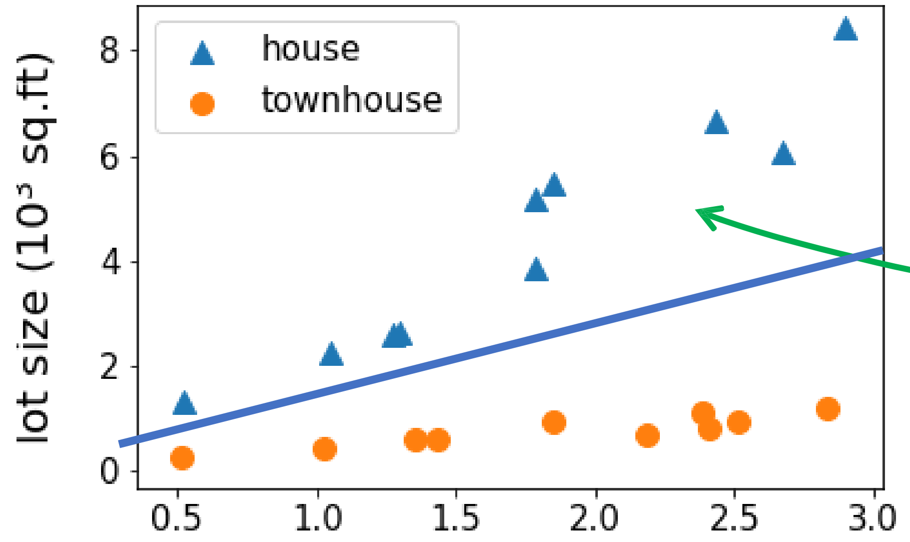
$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ x_d \end{bmatrix} \begin{array}{l} \text{--- living size} \\ \text{--- lot size} \\ \text{--- \# floors} \\ \text{--- condition} \\ \text{--- zip code} \\ \qquad\quad\vdots \end{array} \qquad \longrightarrow \qquad y \text{ --- price}$$

➢ next lectures : infinite dimensional features, select features based on the data

Regression vs Classification

- regression: if $y \in \mathbb{R}$ is a continuous variable
  - e.g., price prediction
- classification: the label is a discrete variable
  - e.g., the task of predicting the types of residence

(size, lot size) $\rightarrow$ house or townhouse?



y = house or townhouse?

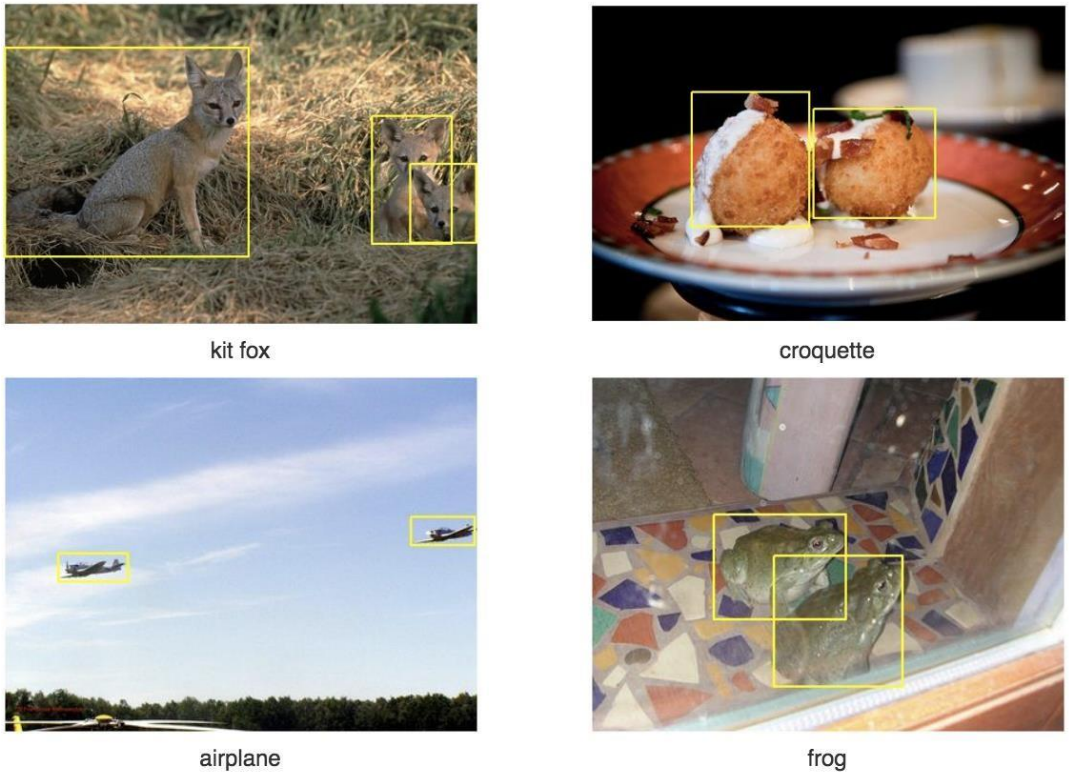
Supervised Learning in Computer Vision

- Image Classification
  - x = raw pixels of the image, y = the main object

ImageNet Large Scale Visual Recognition Challenge. Russakovsky et al.'2015

Supervised Learning in Computer Vision

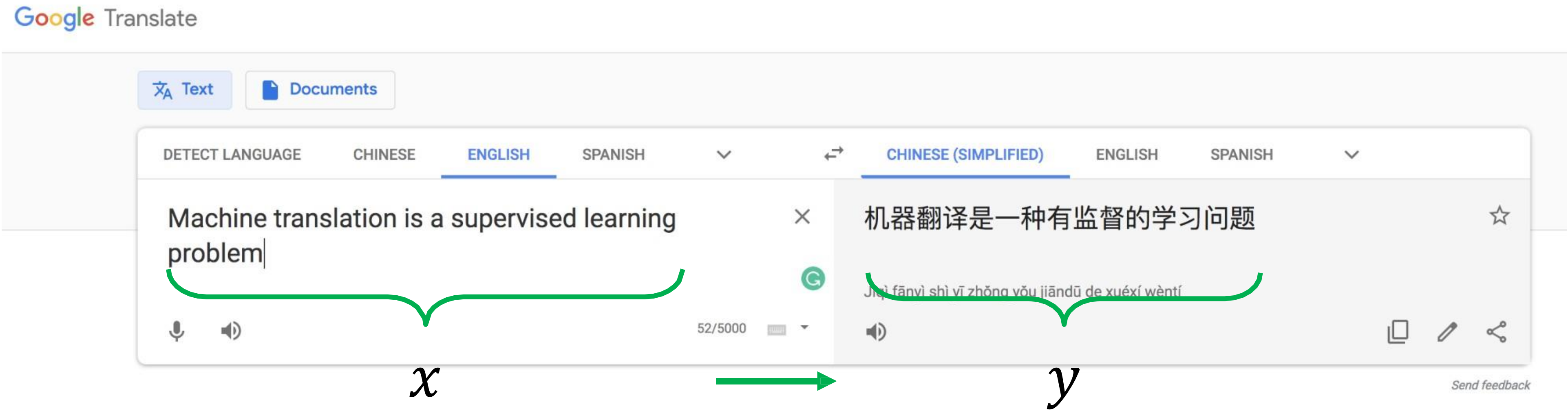
- Object localization and detection
  - $x =$ raw pixels of the image, $y =$ the bounding boxes



ImageNet Large Scale Visual Recognition Challenge. Russakovsky et al.'2015

Supervised Learning in Natural Language Processing

- Machine translation



➢Note: supervised learning algorithms that we cover may be not enough for solving hard vision or NLP problems.

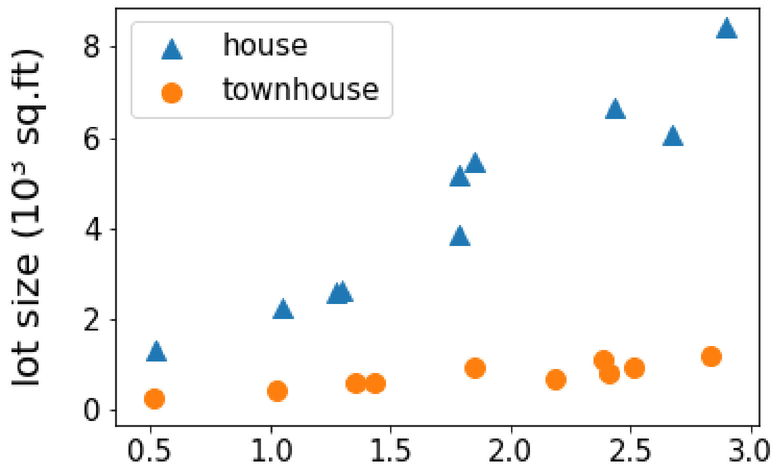
Unsupervised Learning

Let the computer learn how to do something, and use this to determine structure and patterns in data

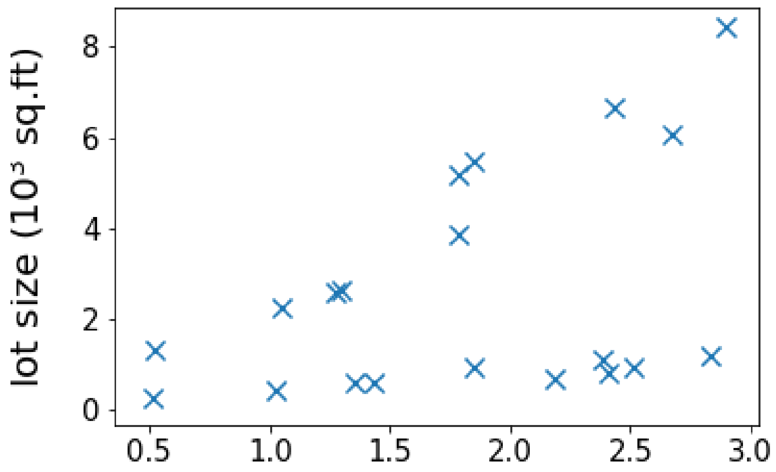
Unsupervised Learning

- Dataset contains no labels: $x^{(1)}, \ldots x^{(n)}$
- Goal (vaguely-posed): to find interesting structures in the data

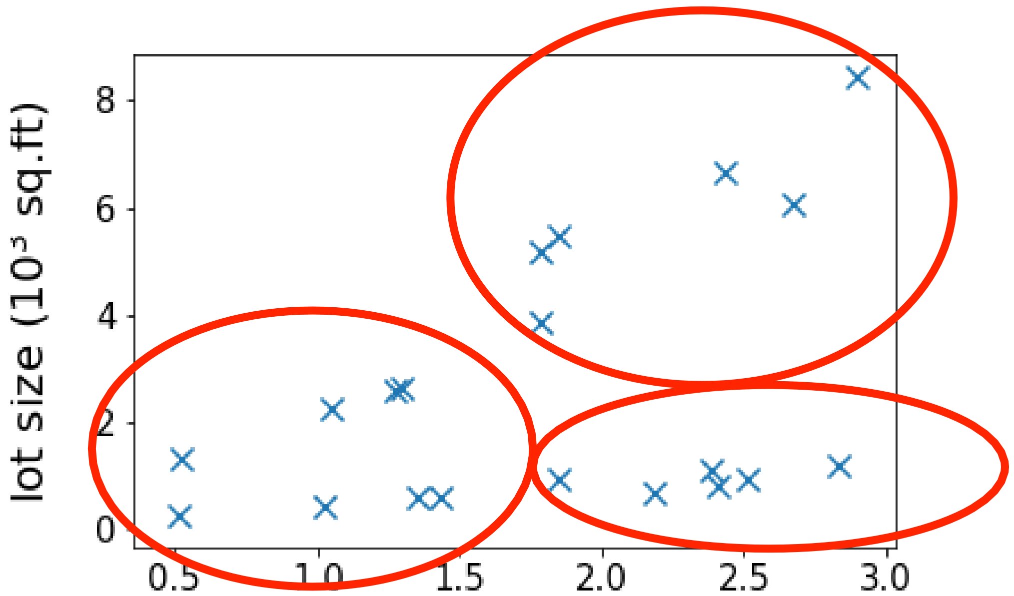
supervised



unsupervised

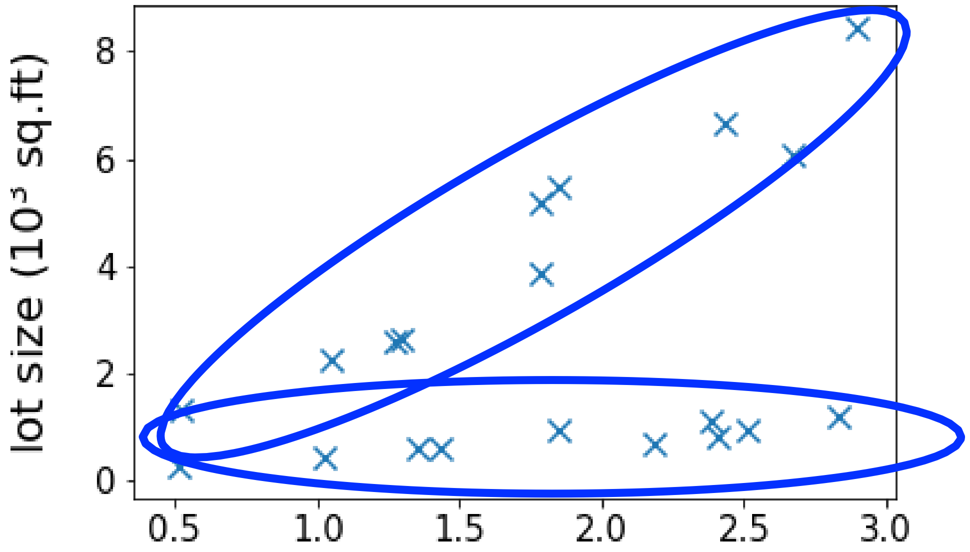


Finding patterns

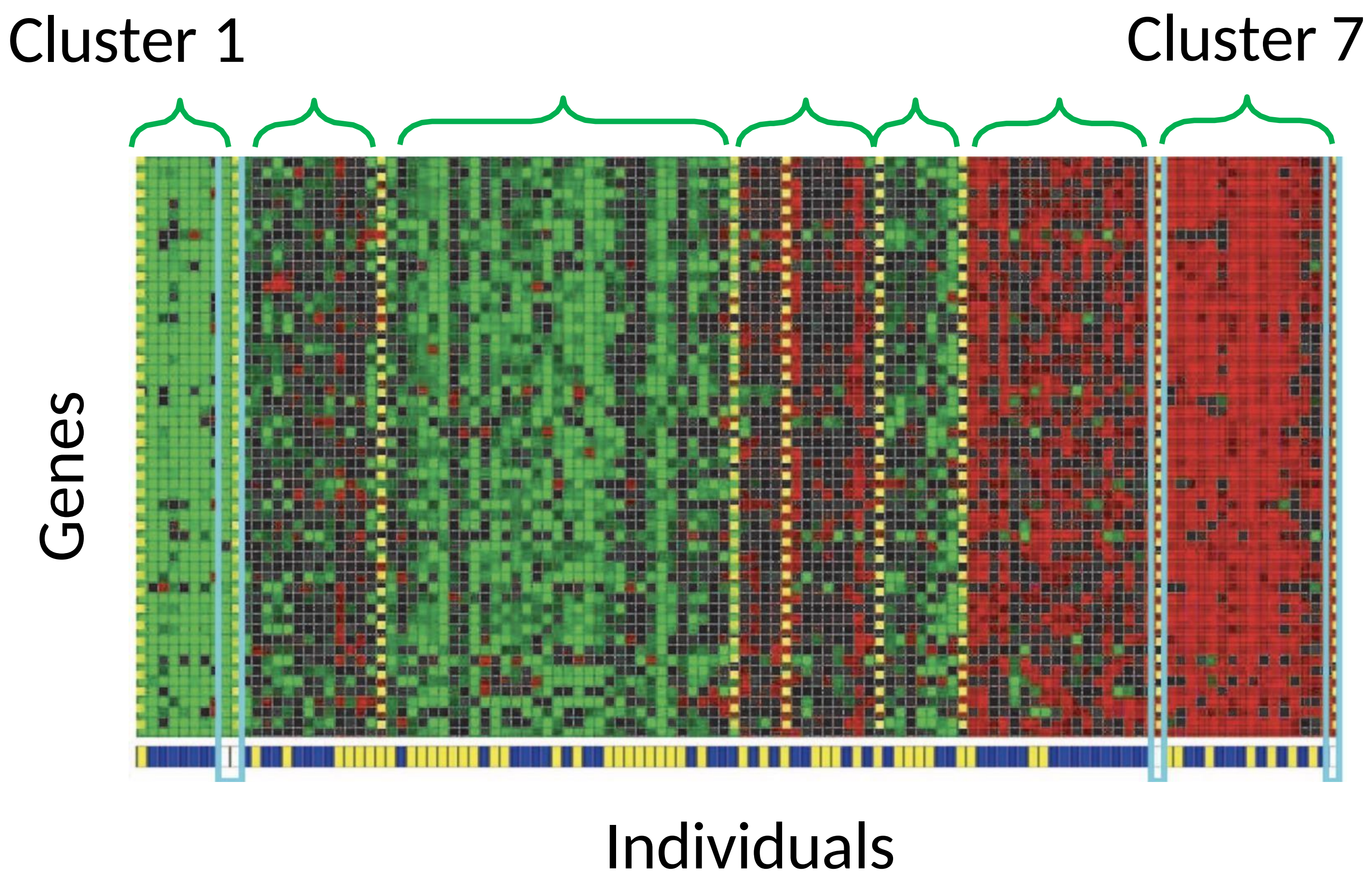


Finding patterns

- next lectures: k-mean clustering, etc



Clustering Genes



Identifying Regulatory Mechanisms using Individual Variation Reveals Key Role for Chromatin Modification. [Su-In Lee, Dana Pe'er, Aimee M. Dudley, George M. Church and Daphne Koller. ’06]

Latent Semantic Analysis (LSA)

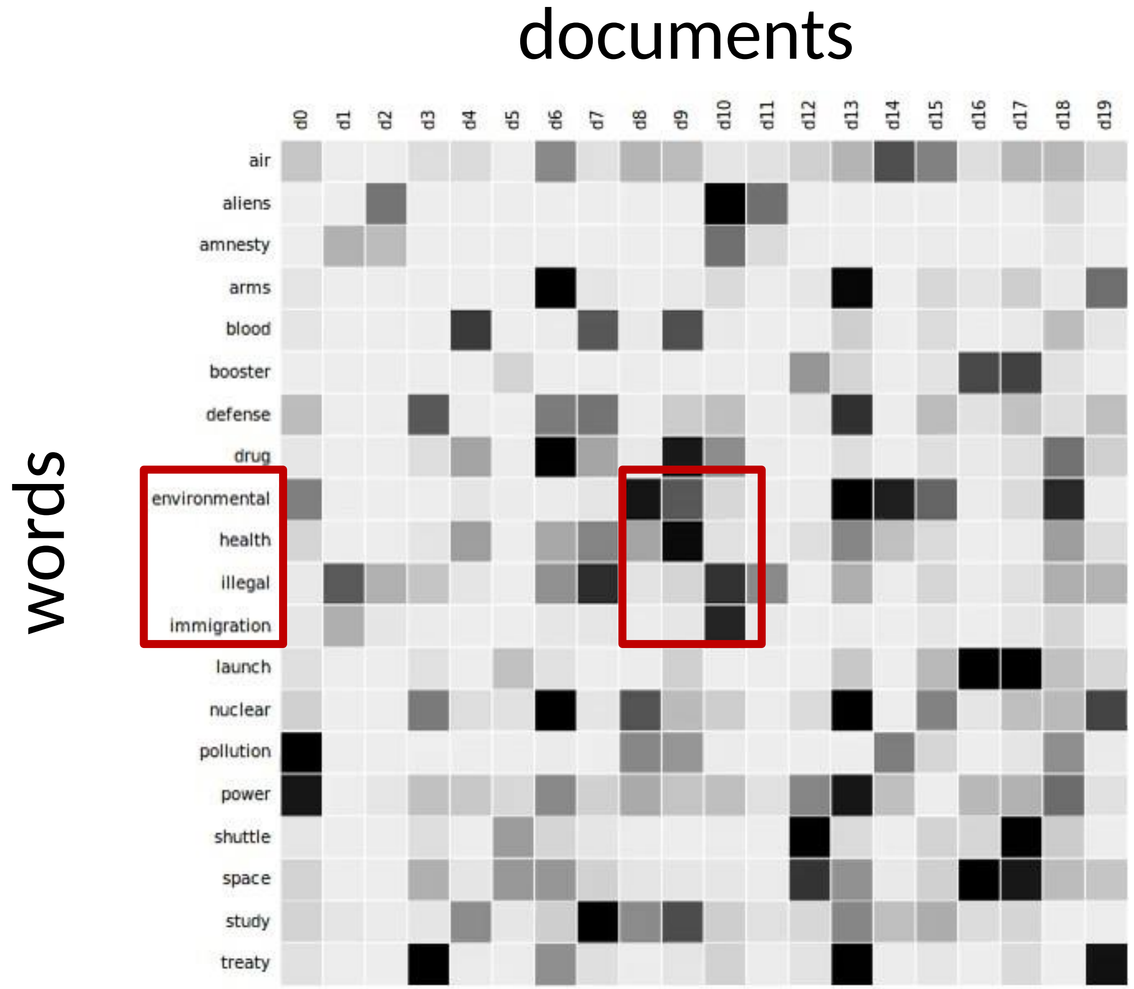
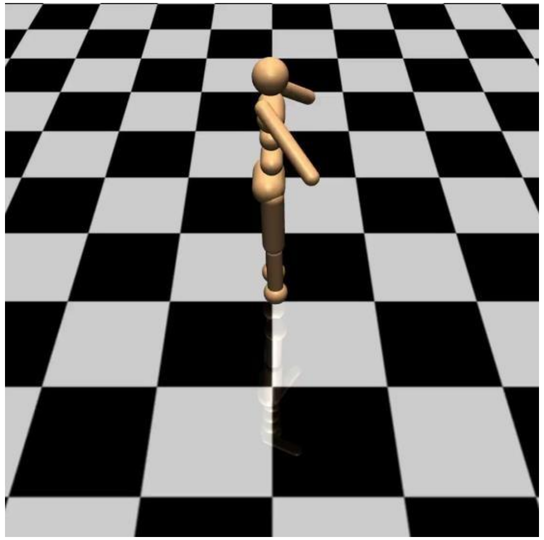


Image credit: https://commons.wikimedia.org/wiki/File:Topic_detection_in_a_document-word_matrix.gif

Reinforcement Learning

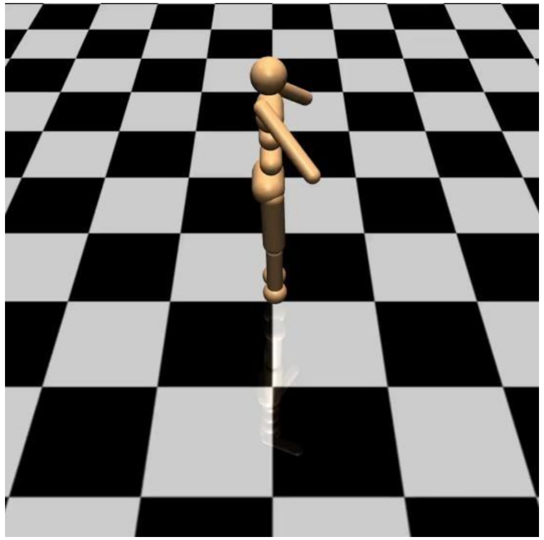
learning to walk to the right



Iteration 10

[Luo-Xu-Li-Tian-Darrell-M.'18]

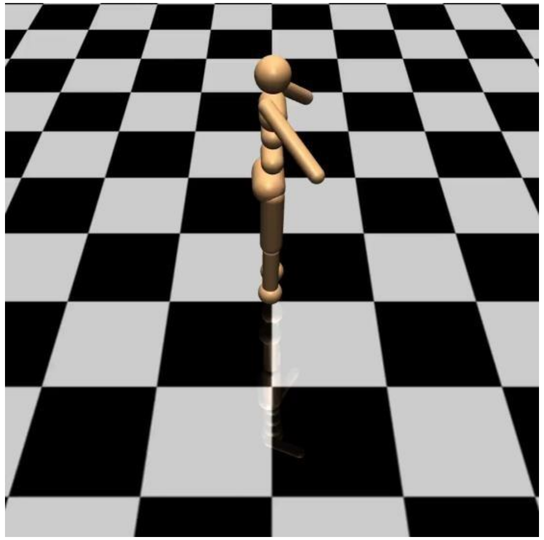
learning to walk to the right



Iteration 80

[Luo-Xu-Li-Tian-Darrell-M.'18]

learning to walk to the right



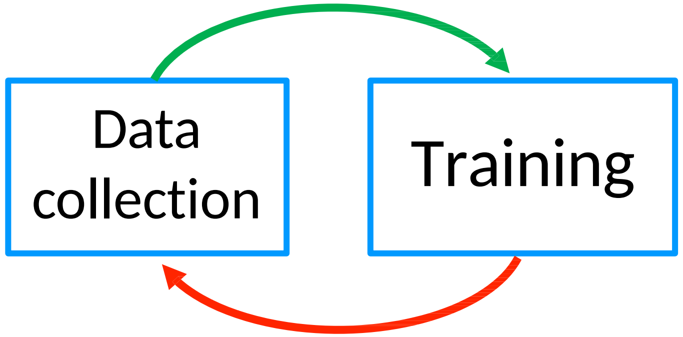
Iteration 210

[Luo-Xu-Li-Tian-Darrell-M.'18]

Reinforcement Learning

- The algorithm can collect data interactively

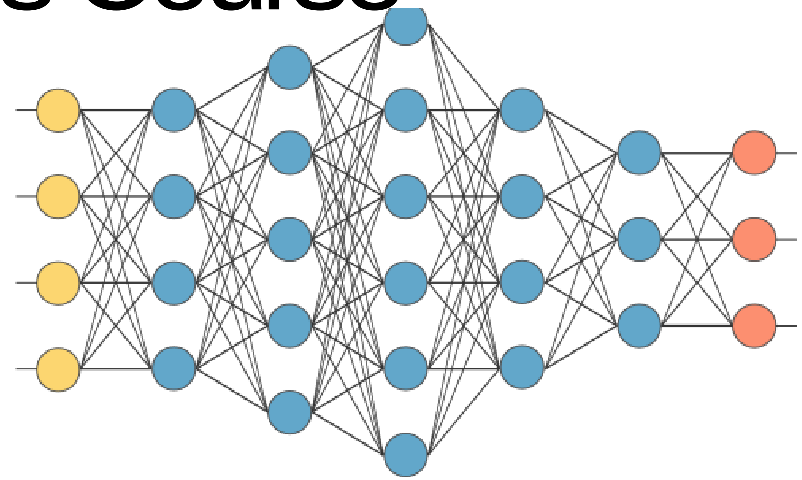
Try the strategy and collect feedbacks



Improve the strategy based on the feedbacks

Other Tools/Topics In This Course

- Deep learning basics



- Introduction to learning theory
  - Bias variance tradeoff
  - Feature selection
  - ML advice
- Broader aspects of ML
  - Robustness/fairness

Readings and next lecture

- Material on pre-reqs:
  - Chapter 2 : Linear Algebra (Goodfellow et. al)
  - Chapter 3: Probability and Information theory (Goodfellow et. al)
  - Lectures on scientific computing with Python
    - https://github.com/jrjohansson/scientific-python-lectures
- Readings:
  - Chapter 1: Tom Mitchell
  - Chapter 1 : Christopher Bishop
  - Chapter 2: Hastie et. al.
- Next lecture : Supervised learning