Machine Learning Introduction

AI Summer School

University of Tehran

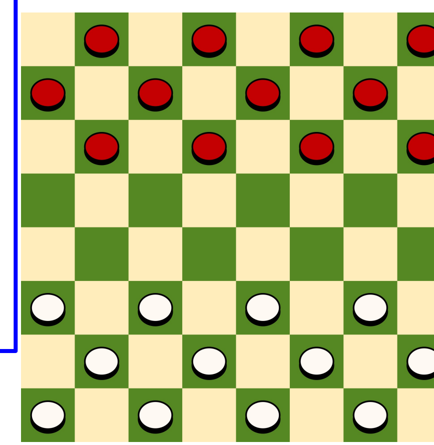
Slides adopted from Stanford CS229

Definition of Machine Learning

Arthur Samuel (1959): Machine Learning is the field of study that gives the computer the ability to learn without being explicitly programmed.

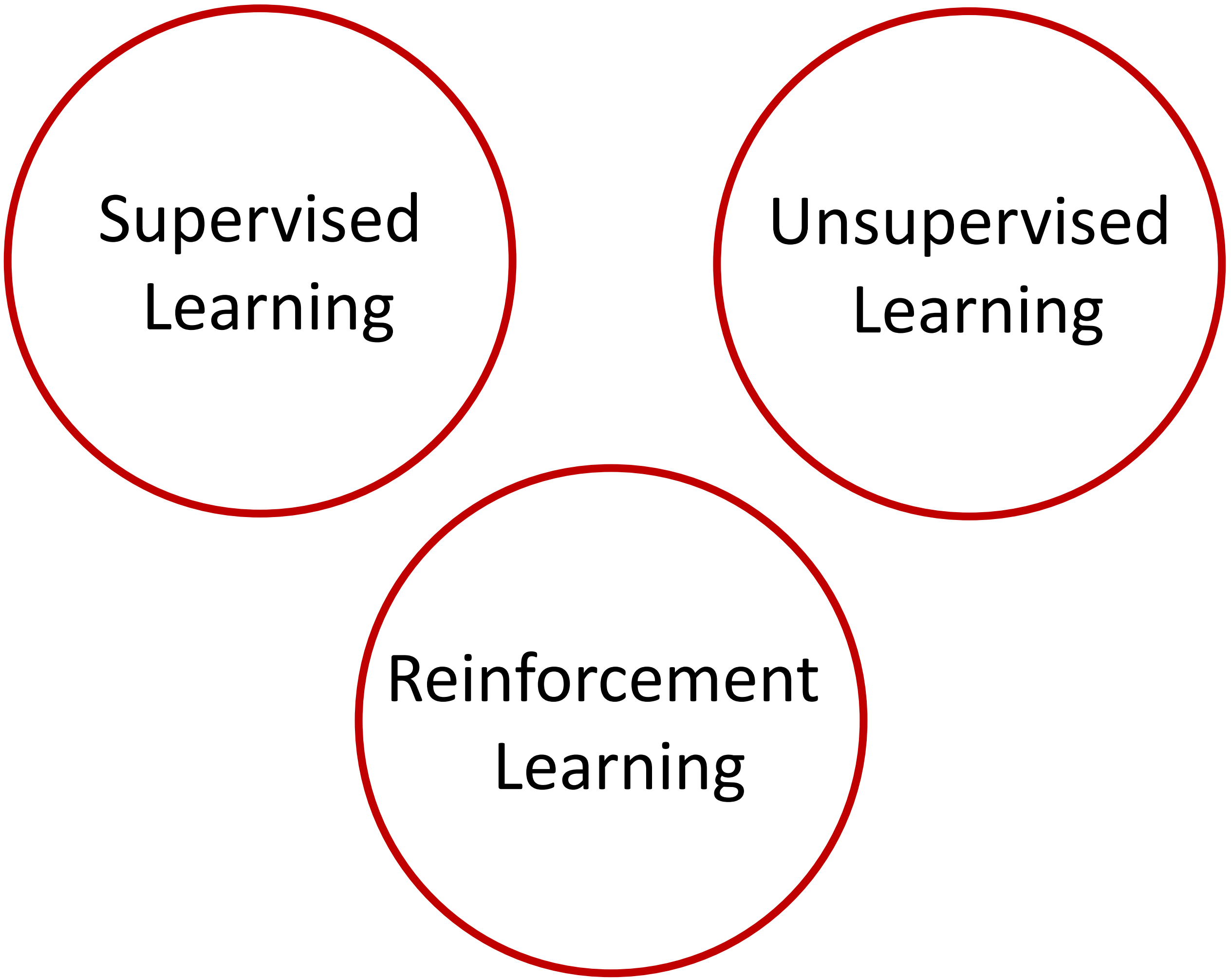
A. L. Samuel*

Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress



Photos from Wikipedia

Taxonomy of Machine Learning
(A Simplistic View Based on Tasks)



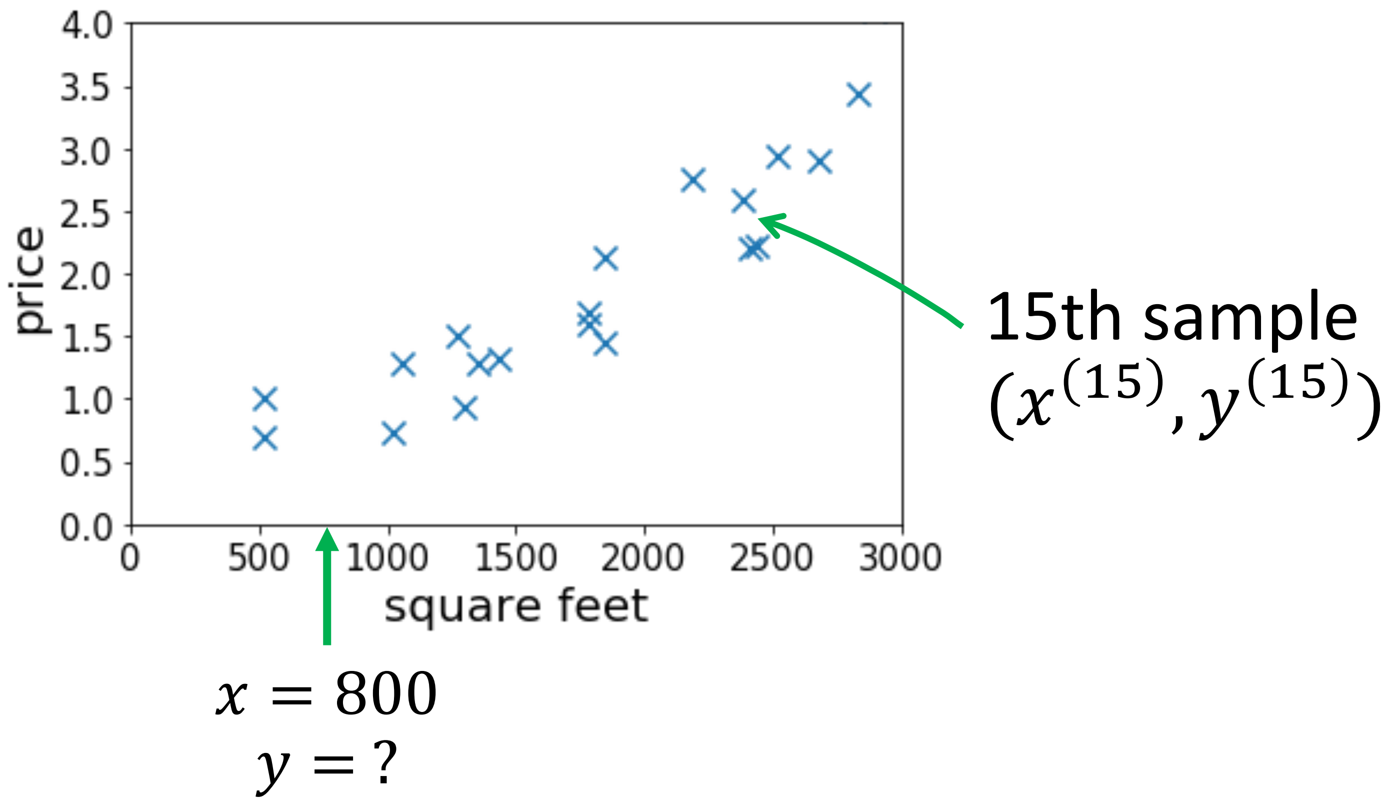
Supervised Learning

Housing Price Prediction

- Given: a dataset that contains n samples

$$\left(x^{(1)}, y^{(1)}\right), \dots \left(x^{(n)}, y^{(n)}\right)$$

- Task: if a residence has x square feet, predict its price?

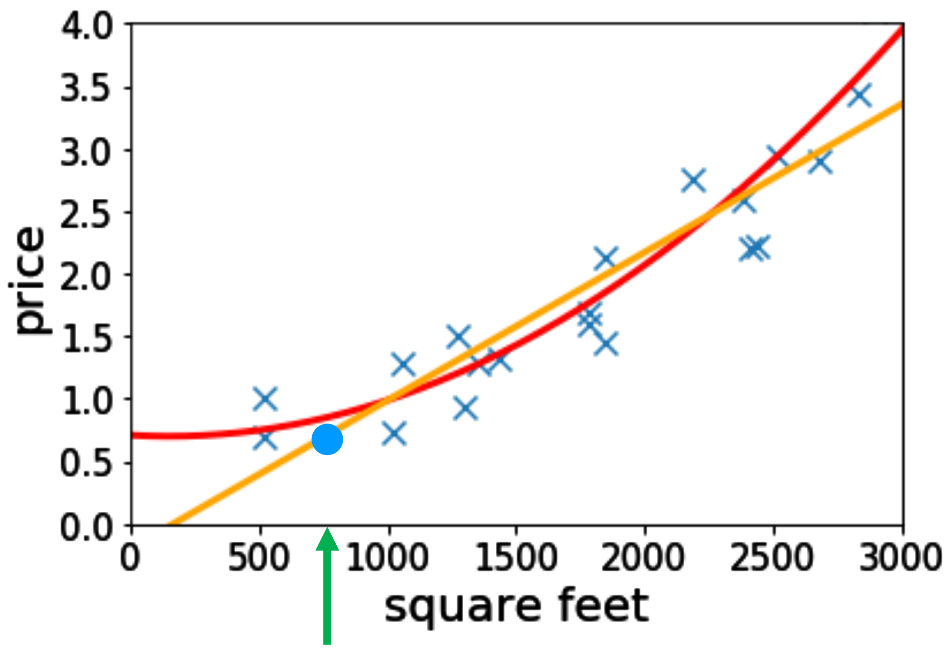


Housing Price Prediction

- Given: a dataset that contains n samples

$$\left(x^{(1)}, y^{(1)}\right), \ldots \left(x^{(n)}, y^{(n)}\right)$$

- Task: if a residence has x square feet, predict its price?



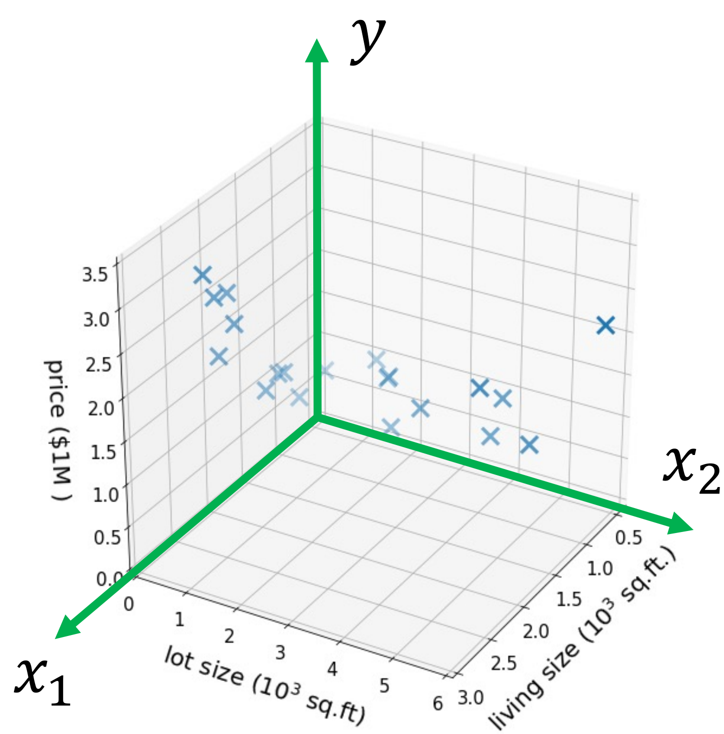
$$x = 800$$
$$y = ?$$

- Lecture 2&3: fitting linear/quadratic functions to the dataset

More Features

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

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



➢ Dataset: $\left(x^{(1)}, y^{(1)}\right), \ldots, \left(x^{(n)}, y^{(n)}\right)$

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

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

High-dimensional Features

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

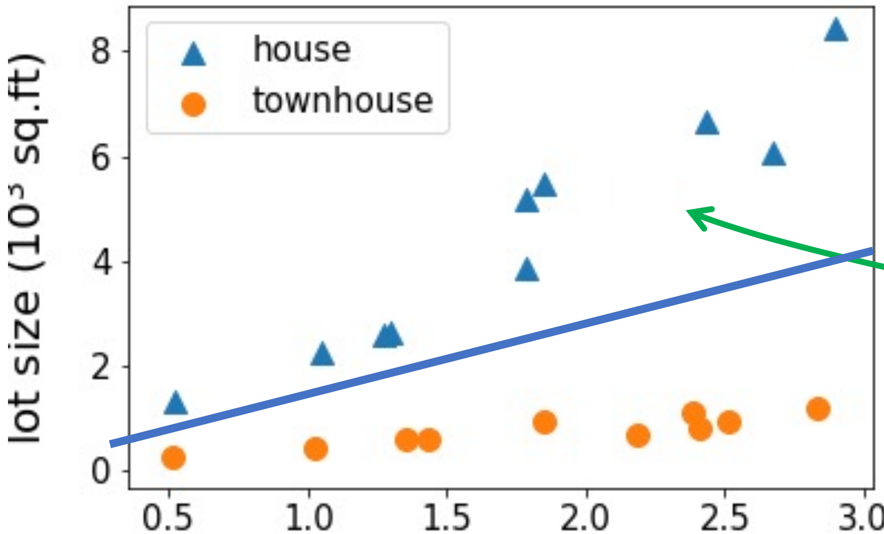
$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ \vdots \\ x_d \end{bmatrix} \begin{array}{l} \text{--- living size} \\ \text{--- lot size} \\ \text{--- \# floors} \\ \text{--- condition} \\ \text{--- zip code} \\ \qquad\vdots \\ \\ \end{array} \longrightarrow \quad y \text{ --- price}$$

- Lec. 6-7: infinite dimensional features (kernels)
- Lec. 10-11: select features based on data (deep learning)

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?

Lecture 3&4: classification

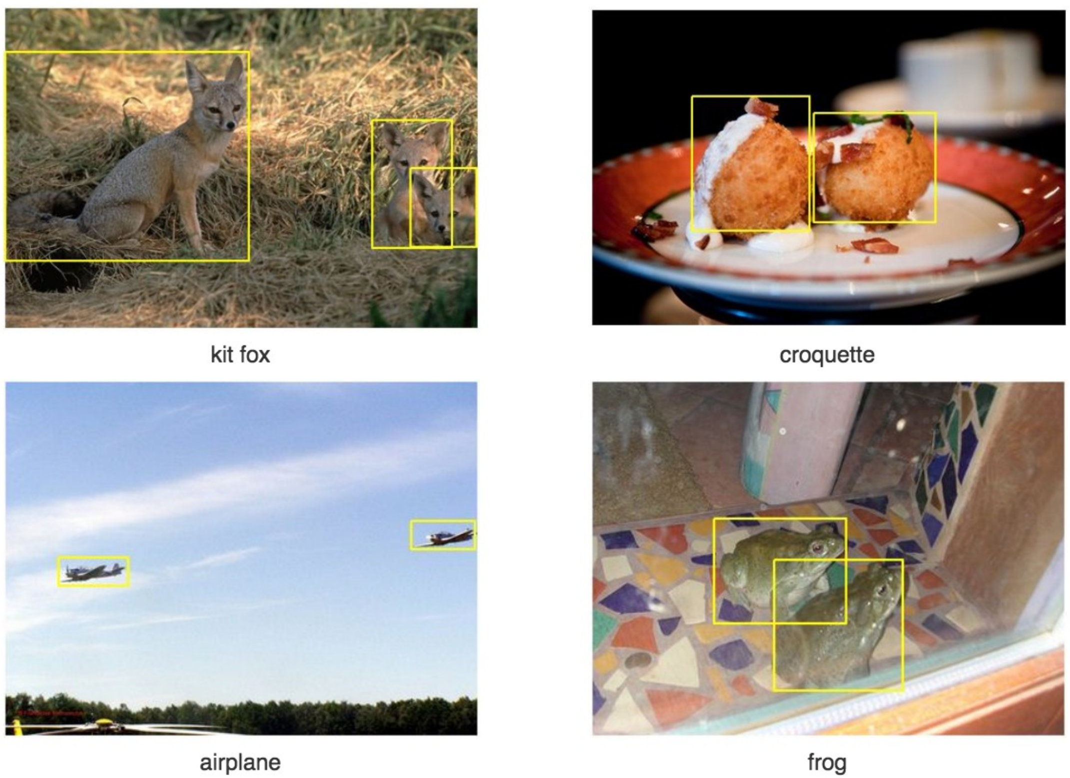
Supervised Learning in Computer Vision

- 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

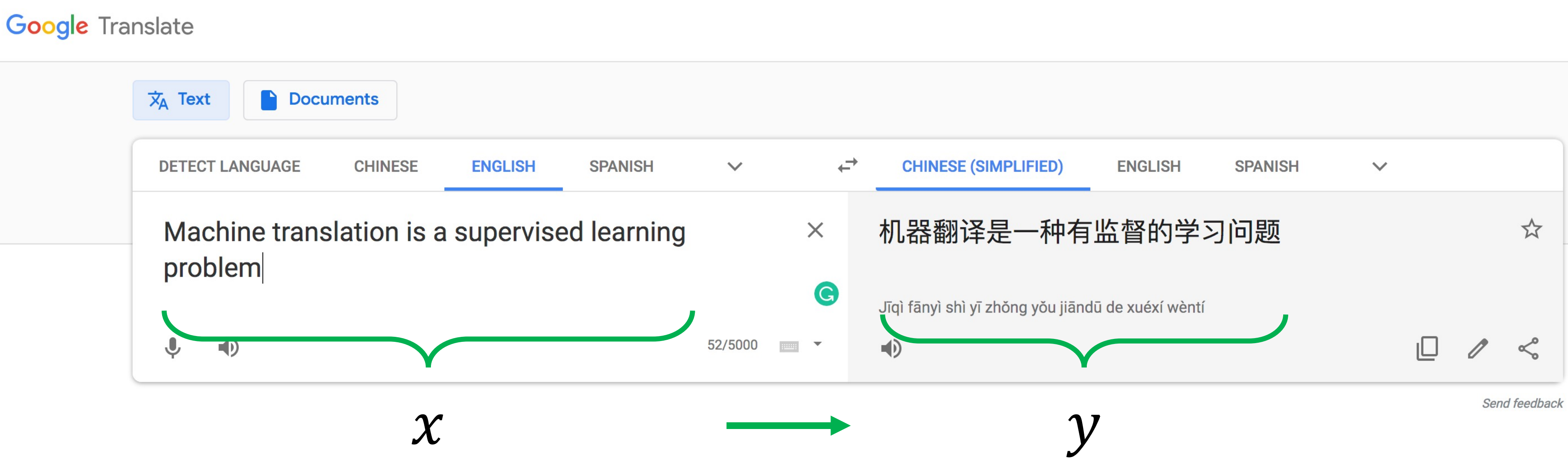
- Object localization and detection
  - $x =$ raw pixels of the image, $y =$ the bounding boxes



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

Supervised Learning in Natural Language Processing

- Machine translation



➢ Note: this course only covers the basic and fundamental techniques of supervised learning (which are not enough for solving hard vision or NLP problems.)

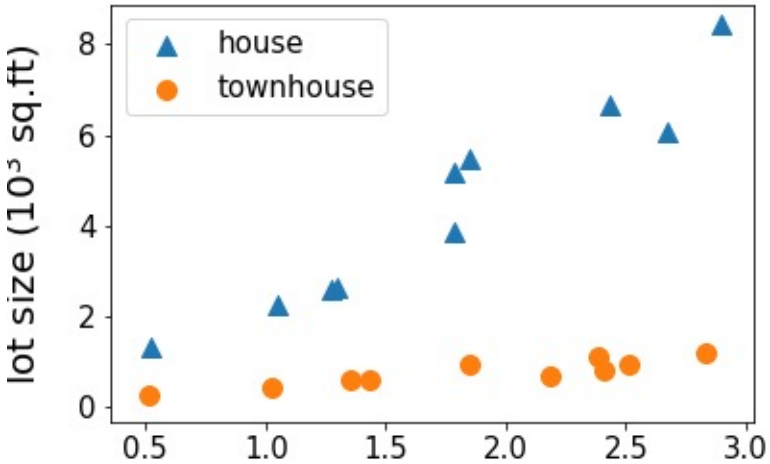
➢ CS224N and CS231N, if you are interested in the particular applications.

Unsupervised Learning

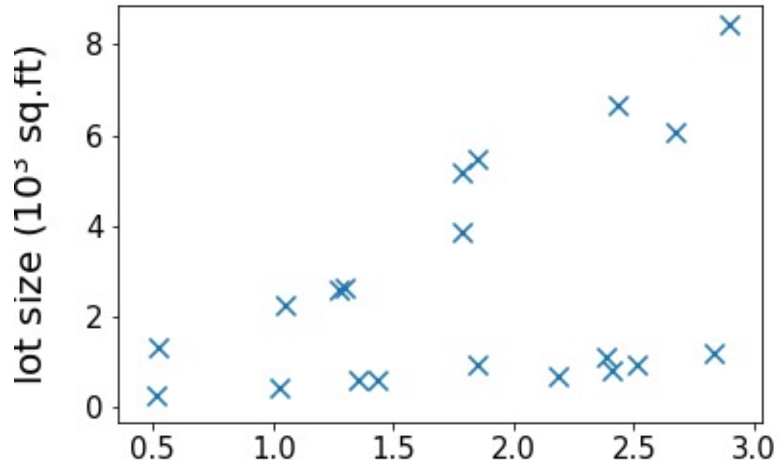
Unsupervised Learning

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

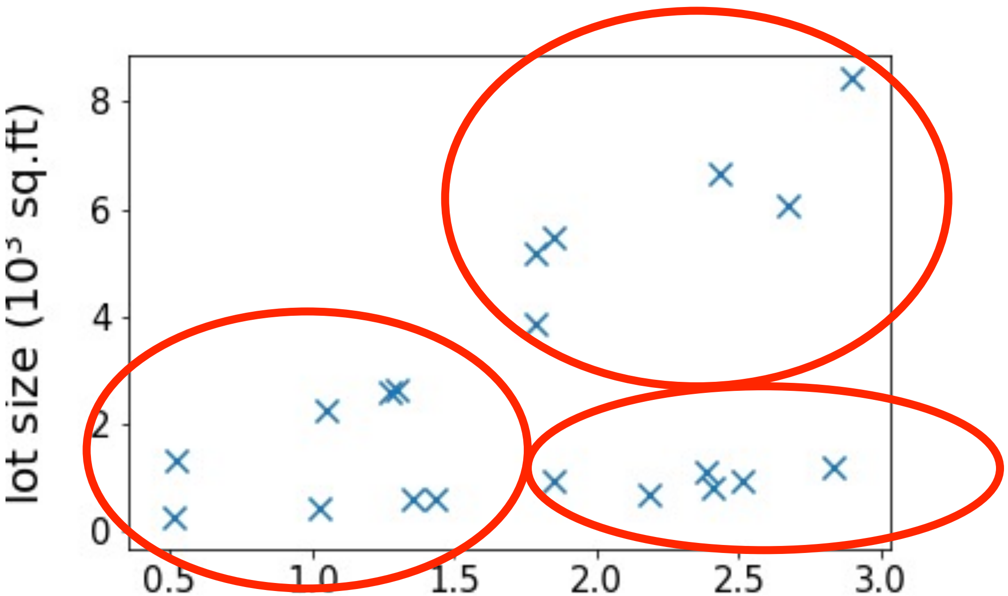
supervised



unsupervised

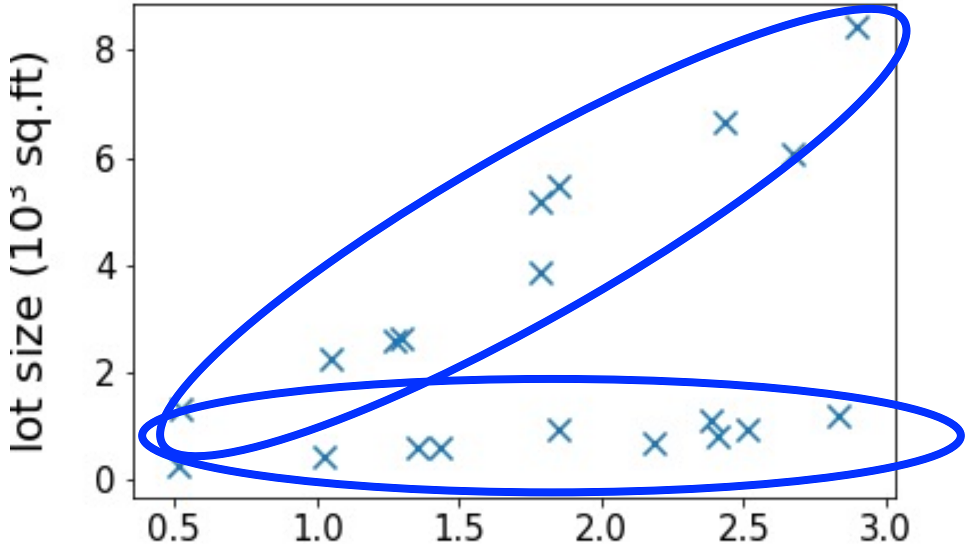


Clustering

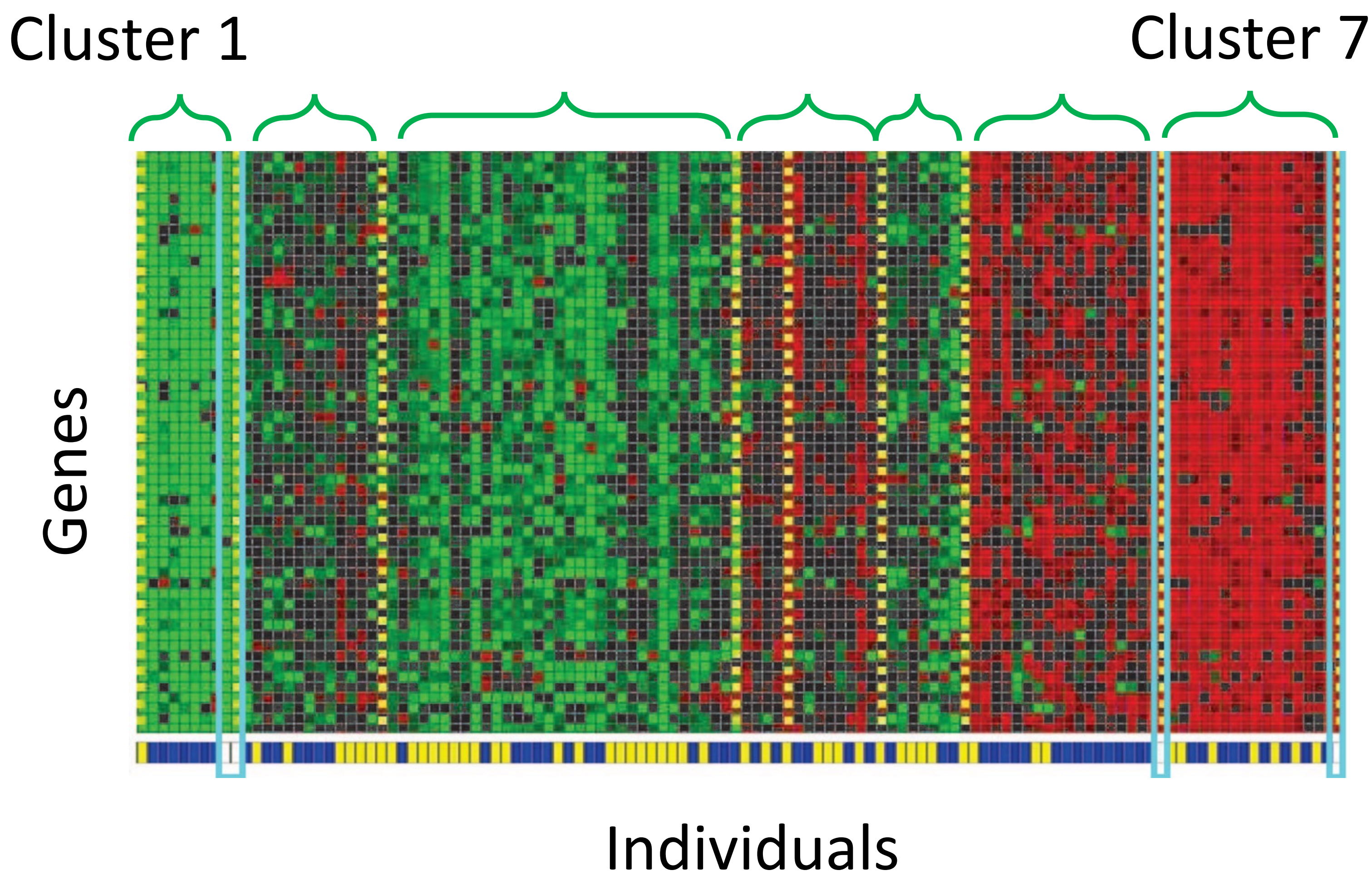


Clustering

- Lecture 12&13: k-mean clustering, mixture of Gaussians



Clustering Genes



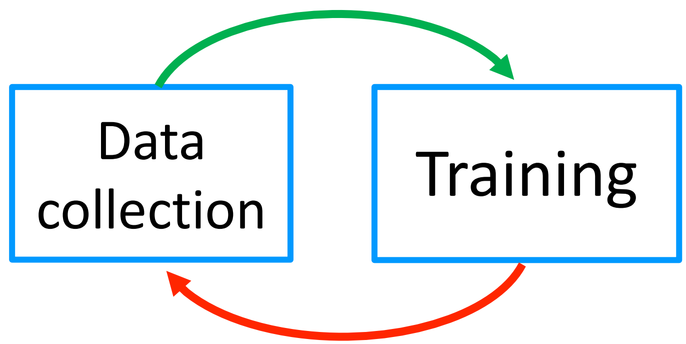
Identifying Regulatory Mechanisms using Individual Variation Reveals Key Role for Chromatin Modification. [Su-In Lee, Dana Pe'er, Aimee M. Dudley, George M. Church and Daphne Koller. '06]

REINFORCEMENT LEARNING

Reinforcement Learning

- The algorithm can collect data interactively

Try the strategy and collect feedback



Improve the strategy based on the feedback