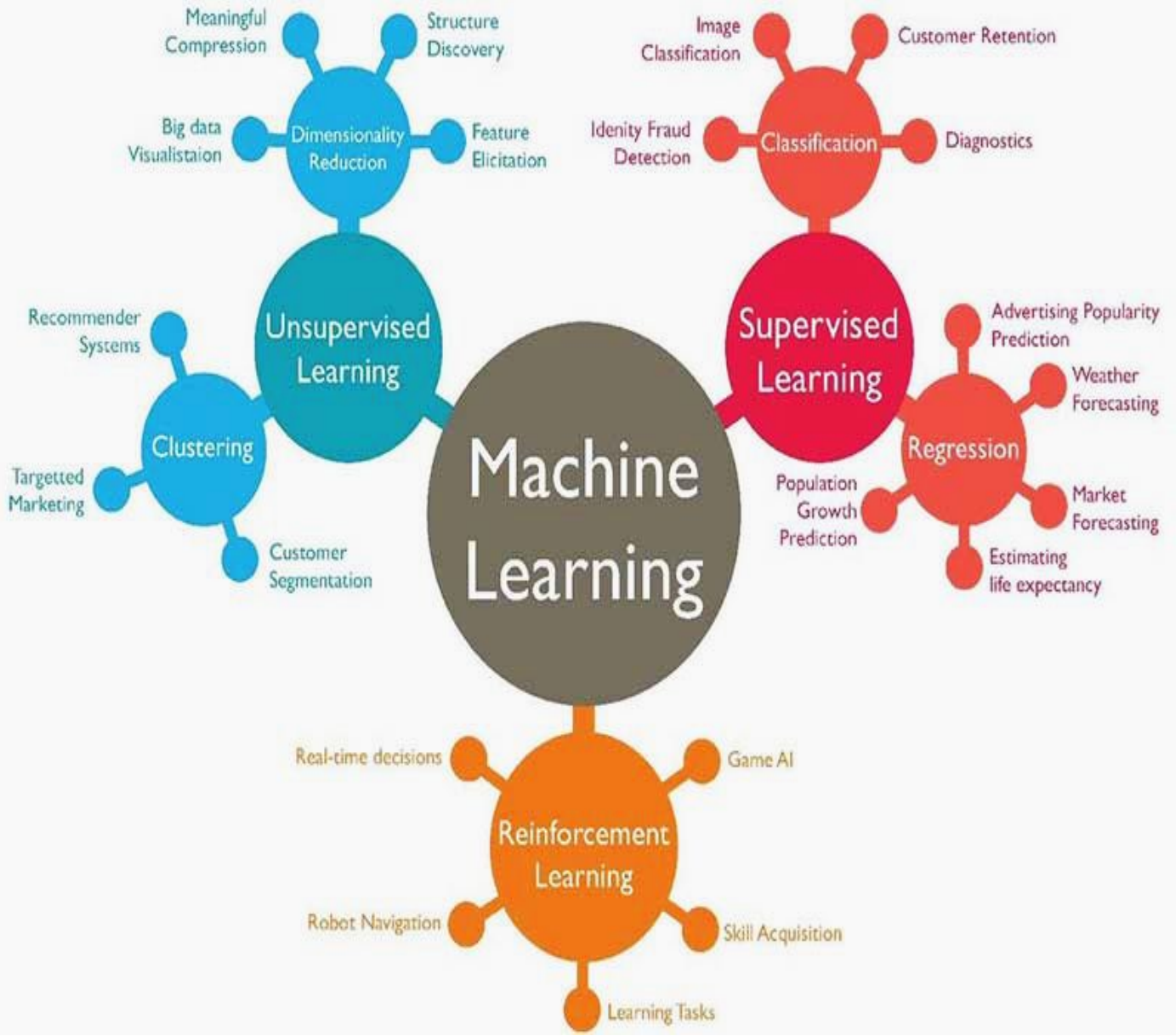
Supervised and Unsupervised Learning

Different Machine Learning Paradigms



Supervised Learning

- To learn an unknown *target function* f
- Input: a *training set* of *labeled examples* (x_j, y_j) where $y_j = f(x_j)$
  - E.g., x_j is an image, $f(x_j)$ is the label "giraffe"
- Output: *hypothesis* h that is "close" to f, i.e., predicts well on unseen examples ("*test set*")
- Many possible hypothesis families for h
  - Linear models, logistic regression, neural networks, decision trees, examples (nearest-neighbor) etc.

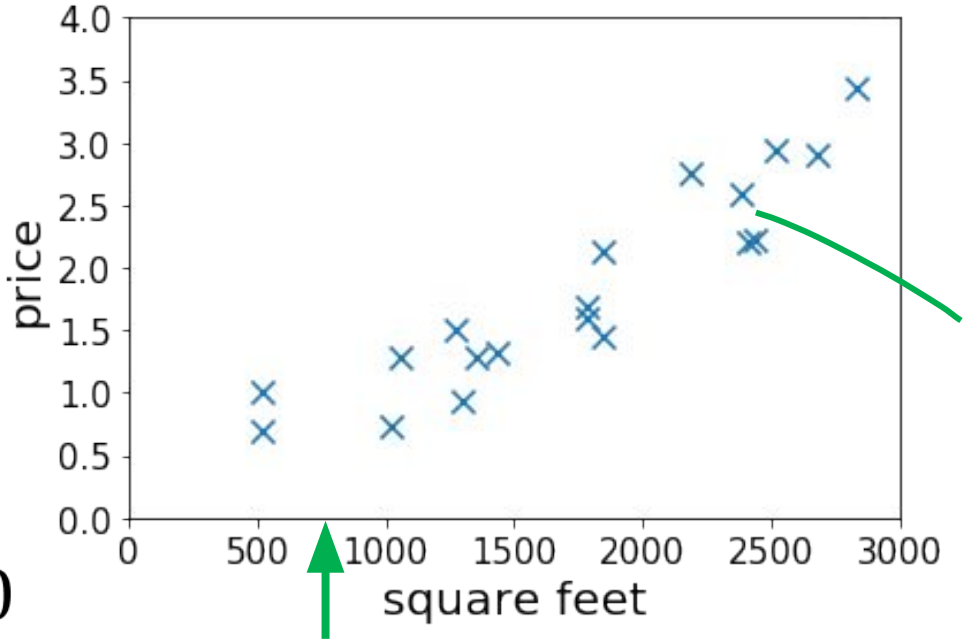
Slide Adapted from **Stuart Russell**

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?



15th sample $(x^{(15)}, y^{(15)})$

$x = 800$

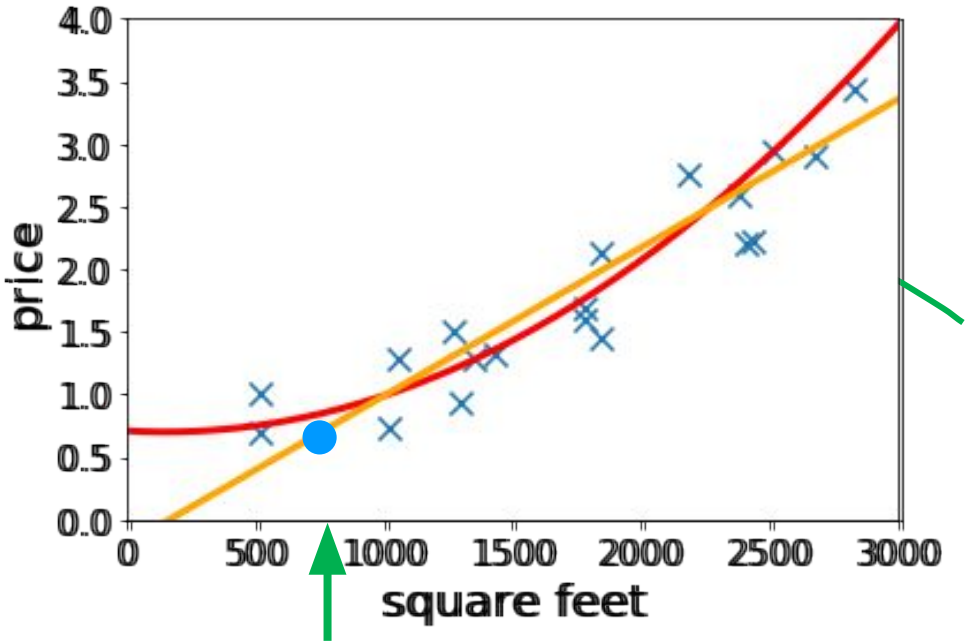
$y = ?$

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?



15th sample $(x^{(15)}, y^{(15)})$

$x = 800$

$y = ?$

- Solution: fitting linear/quadratic functions to the dataset.

High-dimensional Features

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

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ \vdots \\ x_d \end{bmatrix} \begin{matrix} \text{--- living size} \\ \text{--- lot size} \\ \text{--- \# floors} \\ \text{--- condition} \\ \text{--- zip code} \\ \vdots \end{matrix} \rightarrow y \text{ --- price}$$

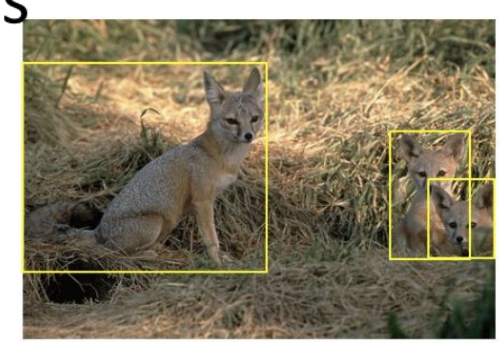
Supervised Learning in CV

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

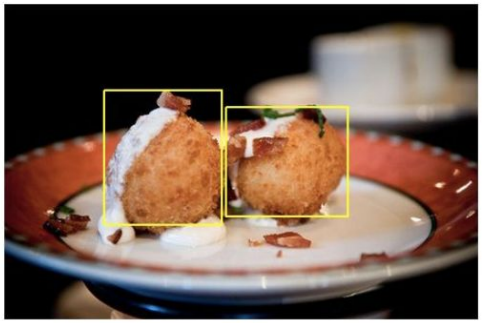
ImageNet Large Scale Visual Recognition Challenge. Russakovsky et al.'2015

Supervised Learning in CV

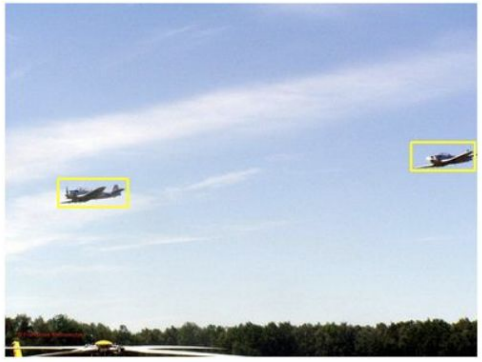
- Object localization and detection
  - x = raw pixels of the image, y = the bounding boxes



kit fox



croquette



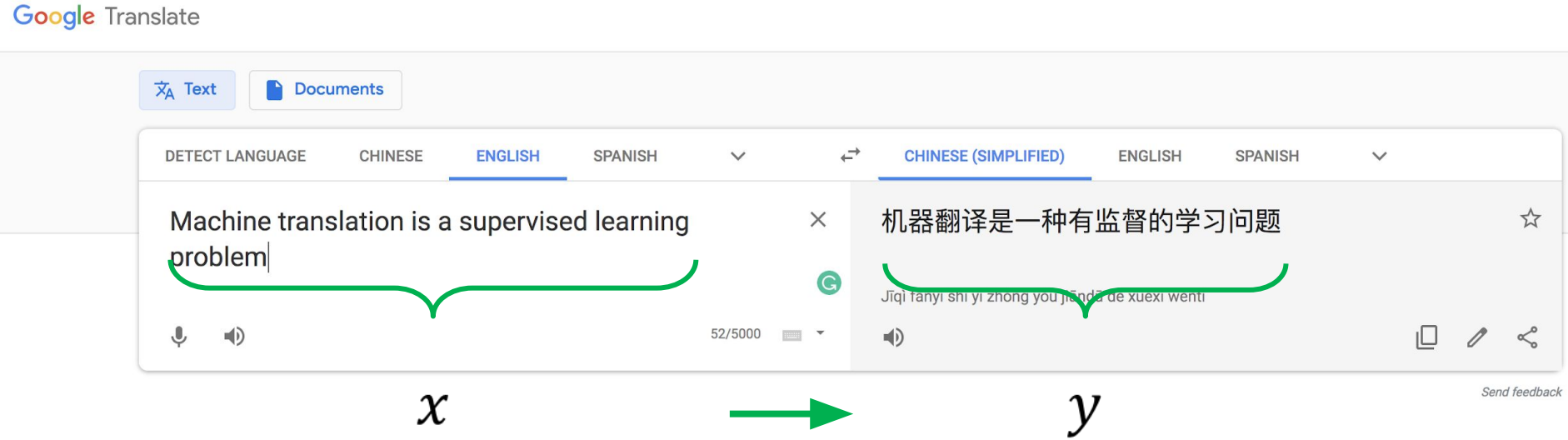
airplane

frog

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.)

Supervised Learning

Advantage

- You have full control over what the machine is learning.

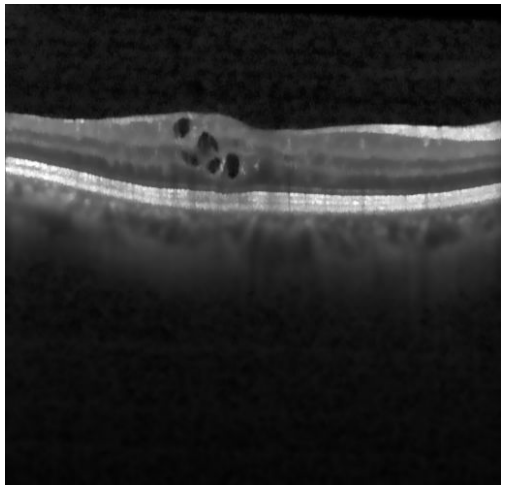
- You can easily test and debug your learning machine.
  - Since the labelled data is available you can easily inspect its output and find out what errors it's making on what type of input data.

Supervised Learning

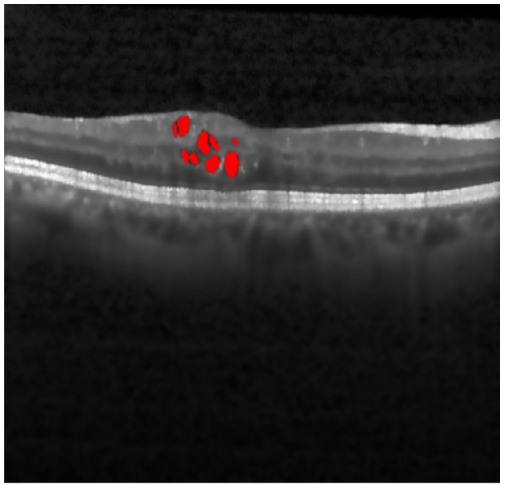
Disadvantage

- Collecting and labelling data is expensive and time-consuming.

  Example: Speech Recognition, Medical Image Analysis, etc.



Original Retinal Scan



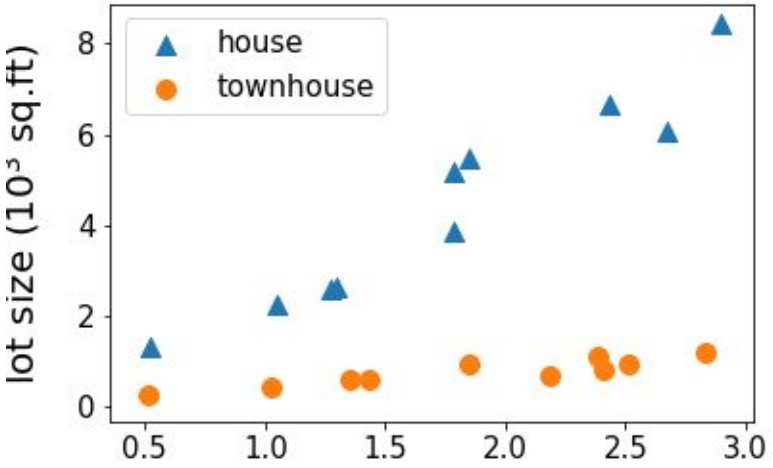
Intra-retinal cysts annotated scan

- Errors in your training data might confuse your algorithm and lower its accuracy. Garbage-in -> Garbage-out

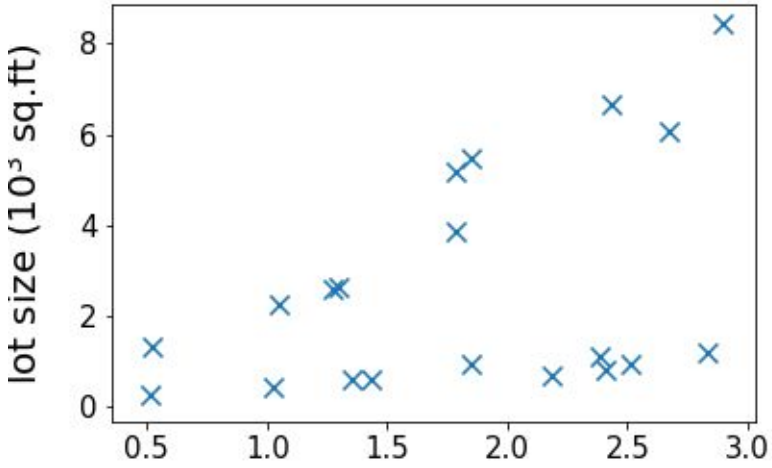
Unsupervised Learning

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

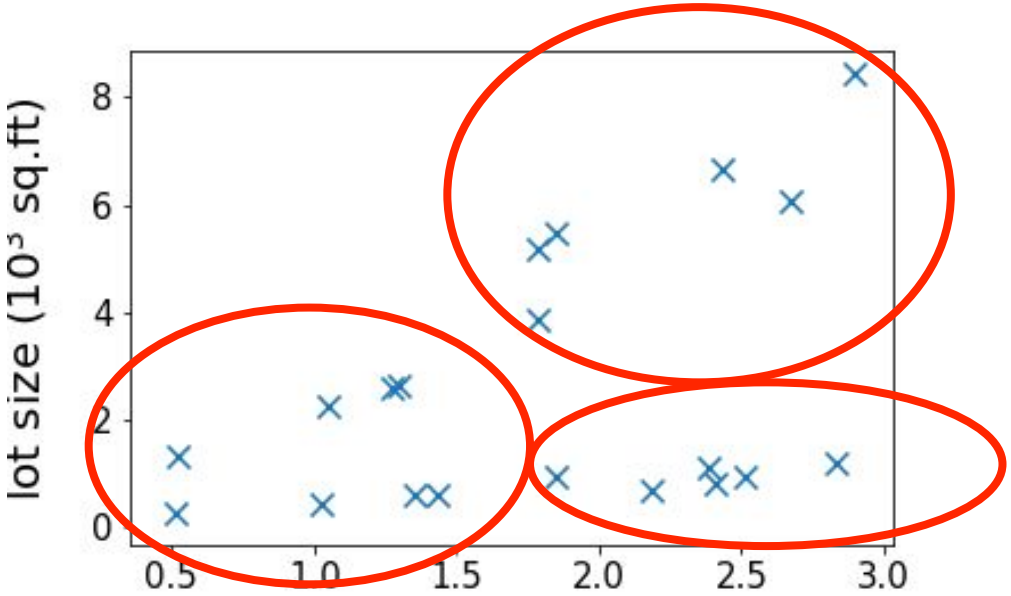
supervised



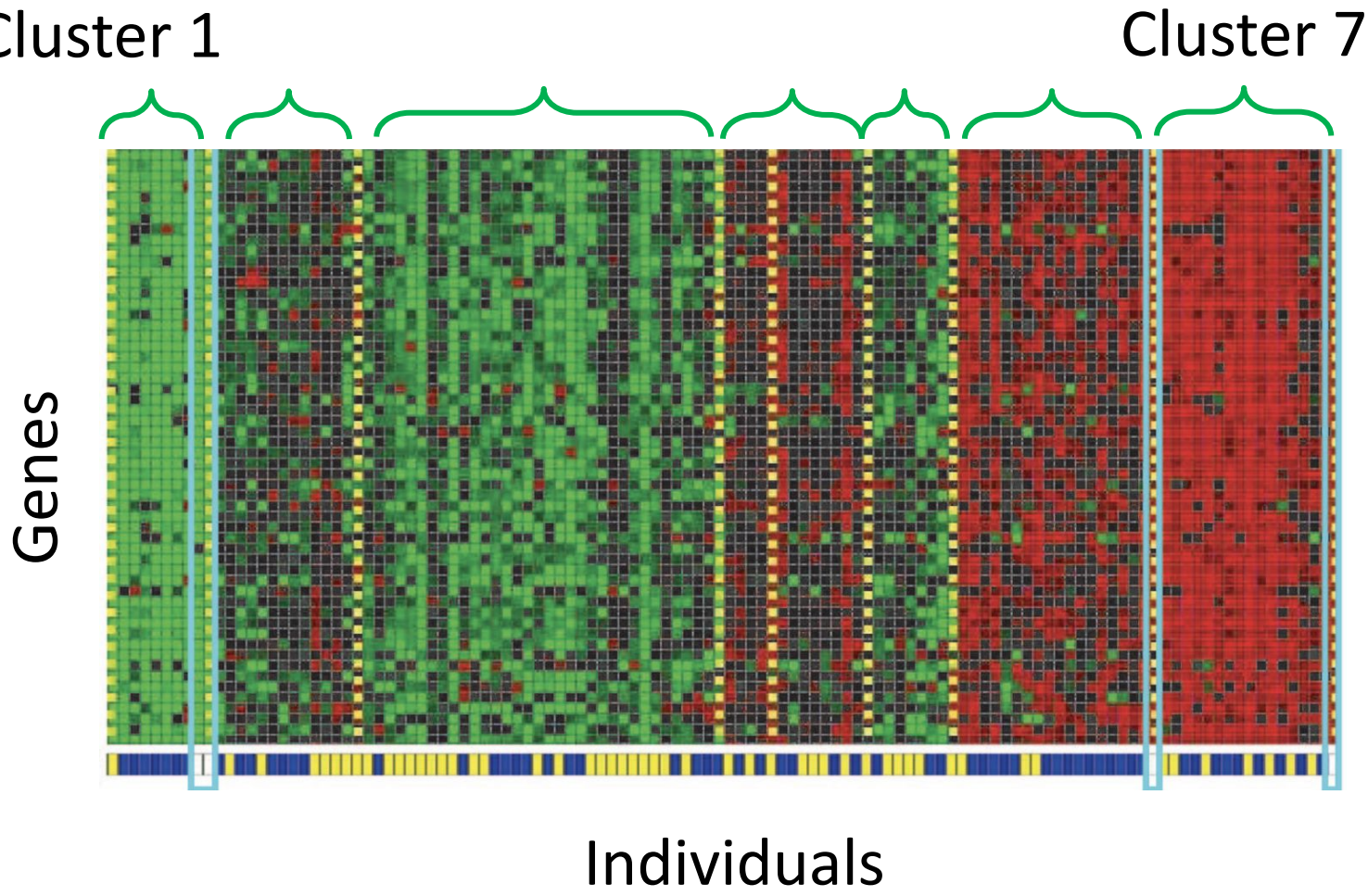
unsupervised



Clustering



Clustering 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]

Need for Unsupervised Learning

- Annotating large datasets is very costly and time consuming. Example: Speech Recognition, Medical Image Analysis, etc.

- There may be cases where we don’t know how many/what classes is the data divided into.

  Example: Data Mining, Sentimental Analysis.

- We may want to use clustering to gain some insight into the structure of the data before designing a classifier.

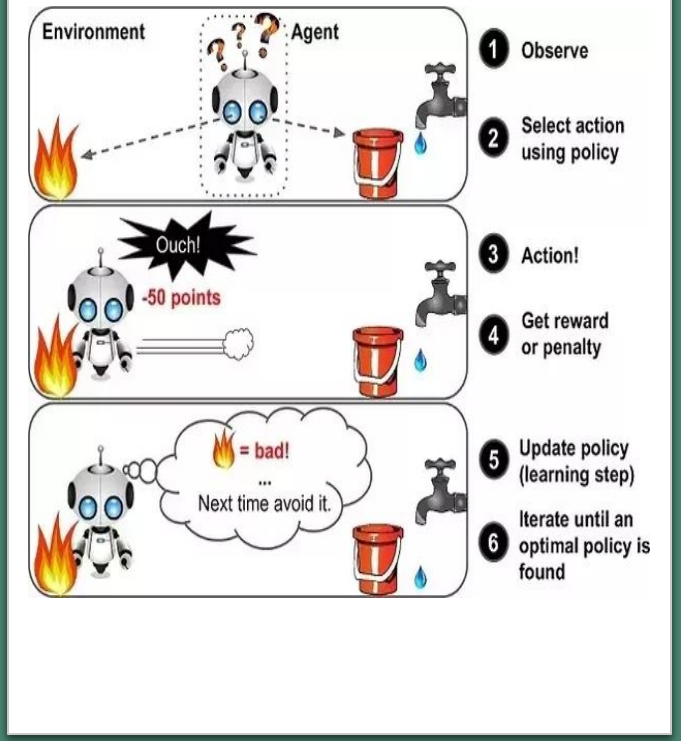
Disadvantages of Unsupervised Learning

- Unsupervised Learning is harder as compared to Supervised Learning. Since, making the inference is difficult due to unavailable labels.

- How do we know if results are meaningful since it has unlabelled data?
  - External evaluation- Expert analysis.
  - Internal evaluation- Objective function.

Reinforcement Learning

A reinforcement learning algorithm, or agent, learns by interacting with its environment. The agent receives rewards by performing correctly and penalties for performing incorrectly.



Reinforcement learning example: Fire fighting agent

Need for Reinforcement Learning

- Reinforcement learning can be used to solve very complex problems that cannot be solved by conventional techniques.
- In the absence of a training dataset, it is bound to learn from its experience.
- Reinforcement learning models can outperform humans in many tasks and learning process is similar to human learning.
- DeepMind's AlphaGo program, a reinforcement learning model, beat the world champion *Lee Sedol* at the game of *Go* in March 2016.

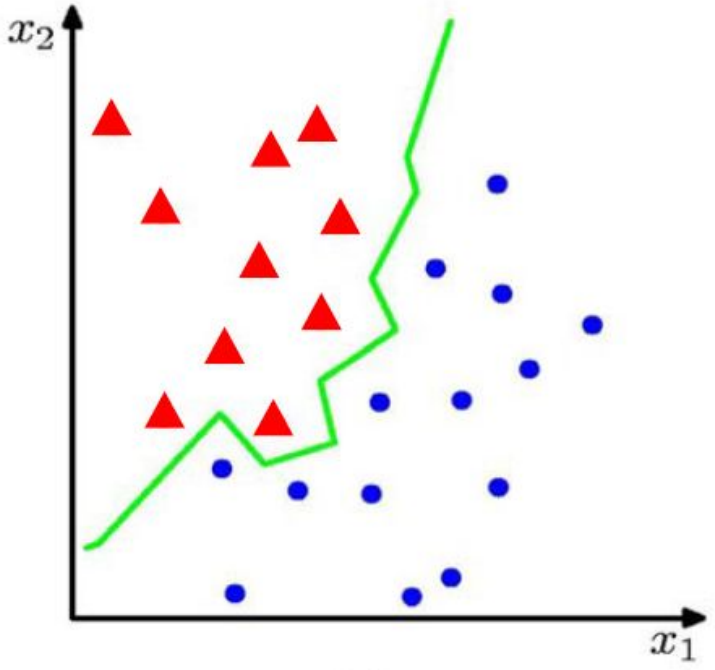
Disadvantages of Reinforcement Learning

- Reinforcement learning needs a lot of data and a lot of computation. It is data-hungry.
  - So for solving video games and puzzles it performs well.
- Reinforcement learning assumes the world is Markovian, which it is not.
  - The Markovian model describes a sequence of possible events in which the probability of each event depends only on the state attained in the previous event.

What is classification problem?

- Let there are two classes of objects.
  - Class 1: Set of dog pictures
  - Class 2: Set of cat pictures
- Problem is
  - Given a picture, you should say whether it is cat or dog.
  - For a human being it is easy…, but for a machine it is a non-trivial problem.

What is classification problem?



- Suppose we are given a training set of N observations

$$(x_1, \ldots, x_N) \text{ and } (y_1, \ldots, y_N), x_i \in \mathbb{R}^d, y_i \in \{-1, 1\}$$

- Classification problem is to estimate f(x) from this data such that

$$f(x_i) = y_i$$

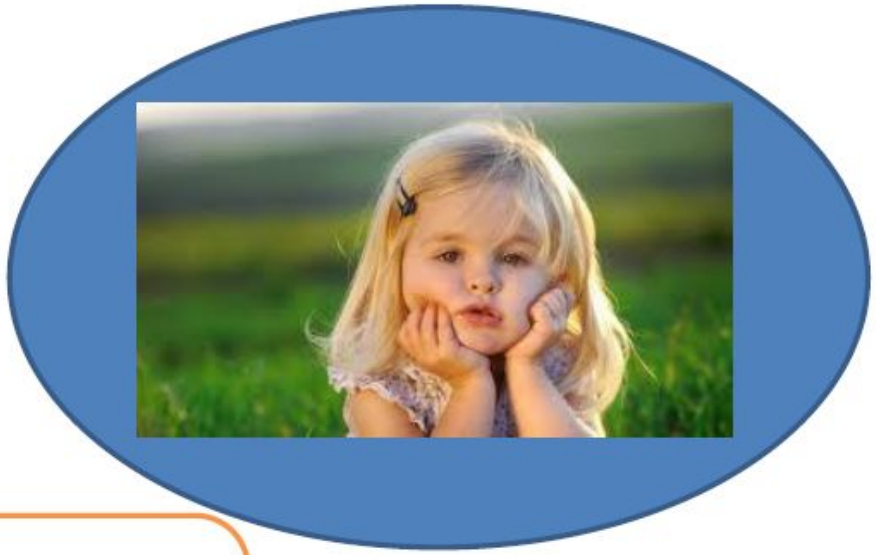
Classification: Supervised Learning

Training Phase

We have shown a set of dog pictures and a set of cat pictures to a child.

Classification: Supervised Learning

Testing Phase



DOG

This picture as it is may not be in the training set

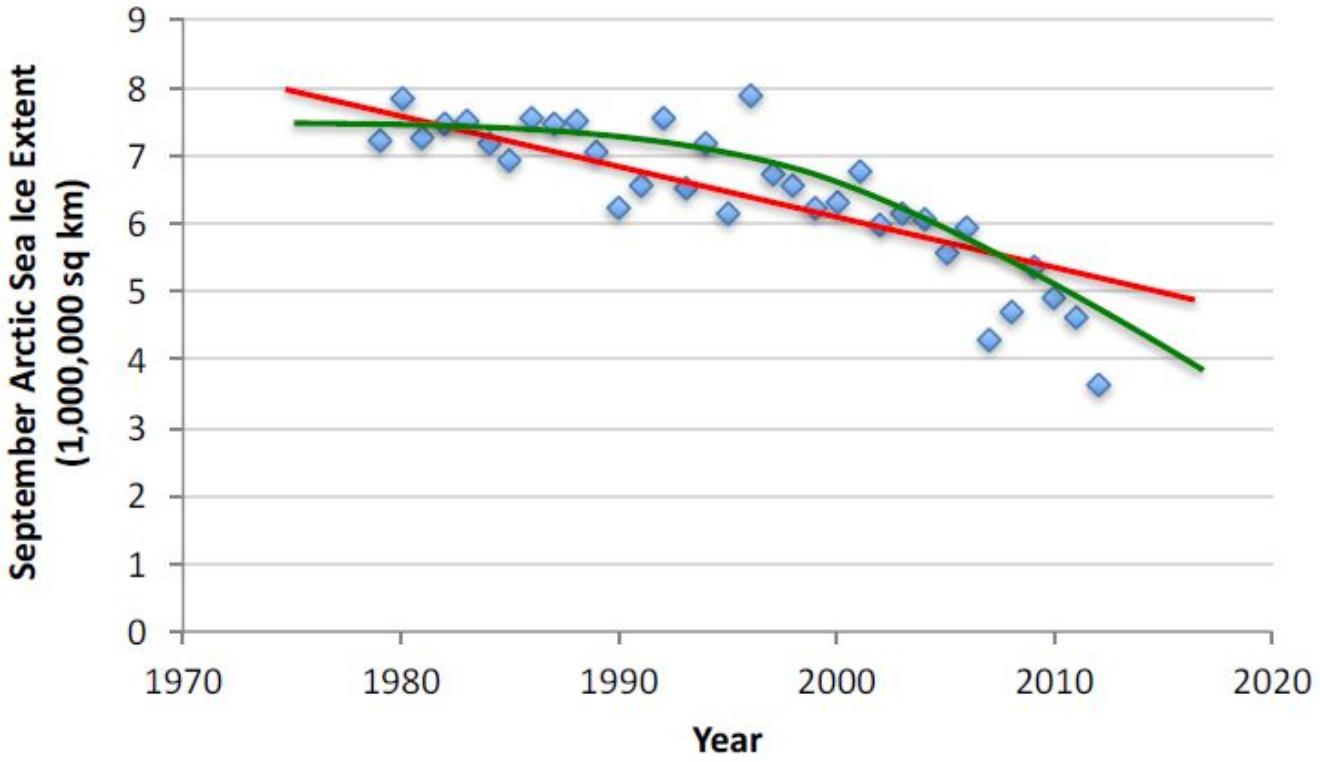
Child has done more than just remembering

What is Learning?

- Child has learnt what is it that is common among dogs … and, what is it that is common among cats… also, what are the distinguishing features/attributes.
- Child has learnt the pattern (regularity) behind all dogs and the pattern behind all cats.
- Child then recognized a test image as having a particular pattern that is unique to dogs.

What is Regression Problem?

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression



Popular ML algorithms

Classification

- Linear Classifiers
- Support Vector Machines
- Decision Trees
- K-Nearest Neighbor
- Random Forest

Regression

- Linear Regression
- Logistic Regression
- Polynomial Regression

Resources: Journals

- Journal of Machine Learning Research www.jmlr.org
- Machine Learning
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association

Resources: Conferences

- International Conference on Machine learning (ICML)
- European Conference on Machine Learning (ECML)
- Neural Information Processing Systems (NIPS)
- Computational Learning
- International Joint Conference on Artificial Intelligence (IJCAI)
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)
- IEEE Int. Conf. on Data Mining (ICDM)

Thank You:
Question?