University of
South Australia

COMP 2019

Week 6

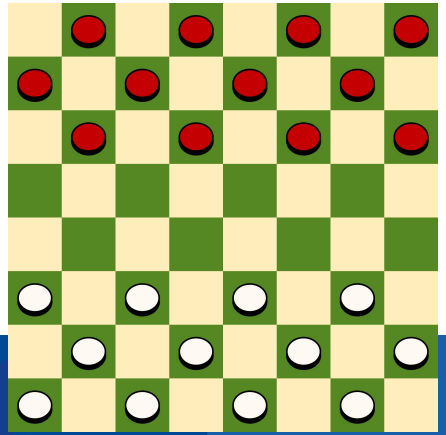
Introduction to Machine Learning

Learning Objectives

- Explain supervised, unsupervised, and reinforcement learning (CO1)

Definition of Machine Learning (ML)

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.

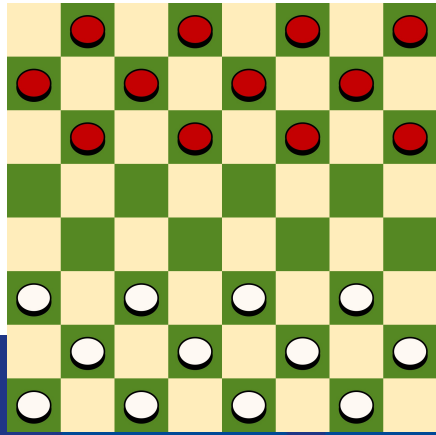


Photos from Wikipedia

Definition of Machine Learning (ML)

- Tom Mitchell (1998) Well-posed Learning Problem: A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

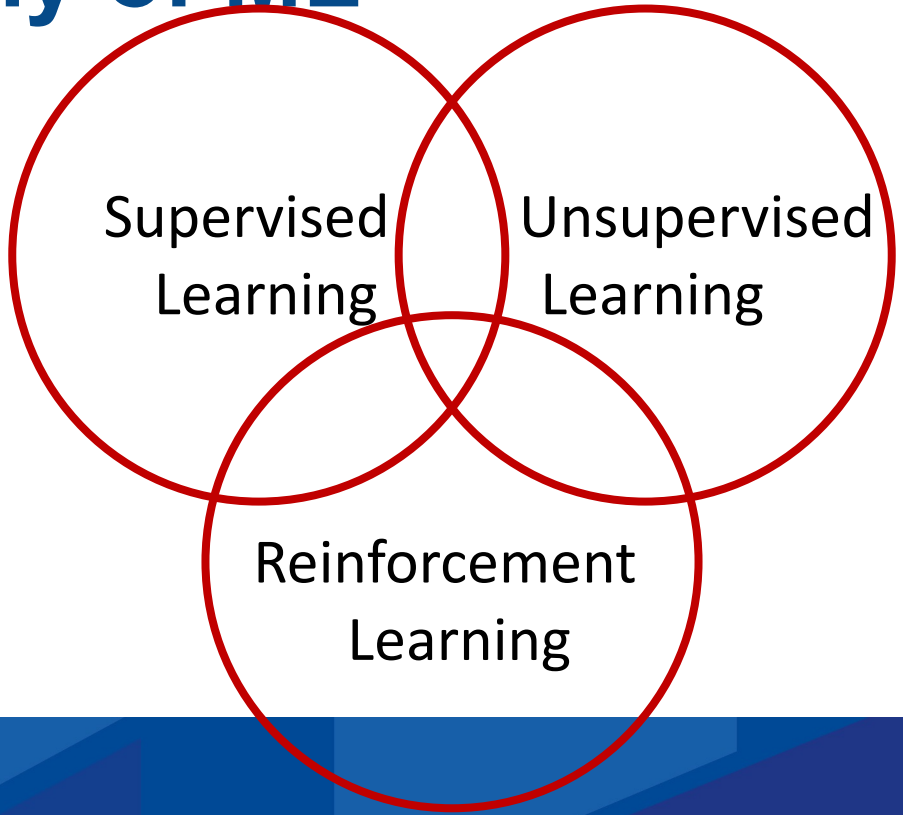
- Experience (data):
  games played by the program (with itself)
- Performance measure: winning rate



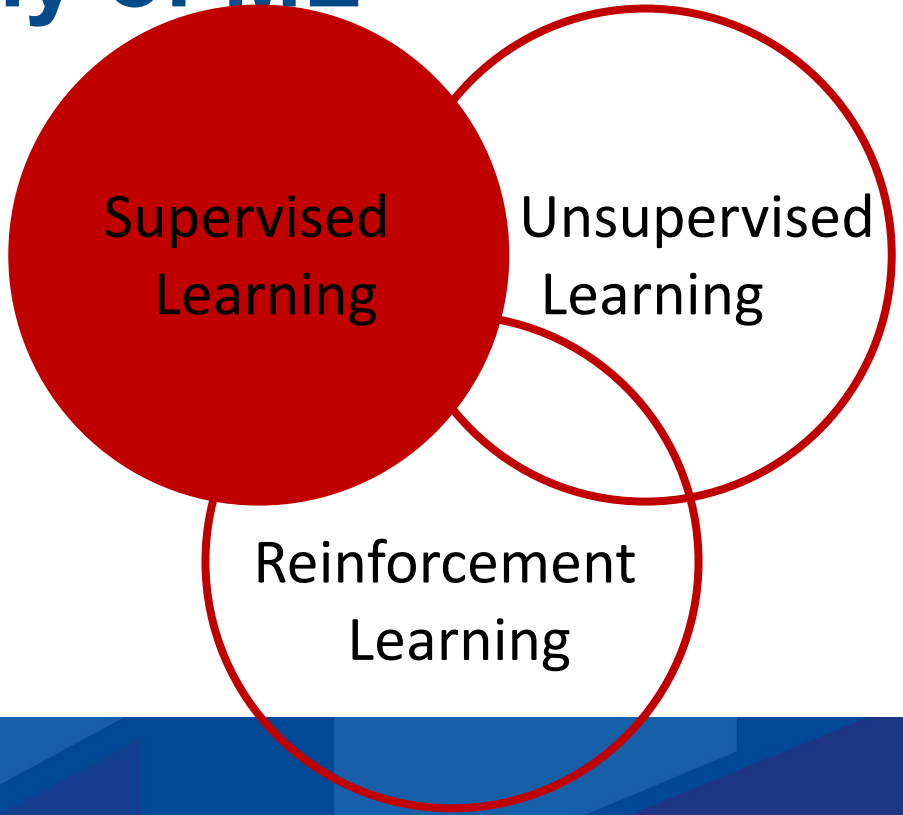
Machine Learning Examples

- Database mining, “Big Data Analytics”
  - Large datasets from growth of automation/web
  - Web click data, medical records, biology, engineering
- Applications we can’t program by hand
  - Autonomous helicopter pilot, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision, Super-human Games
- Self-customizing programs
  - Amazon, Netflix product recommendations
- Understanding human learning (brain, real AI)

Taxonomy of ML



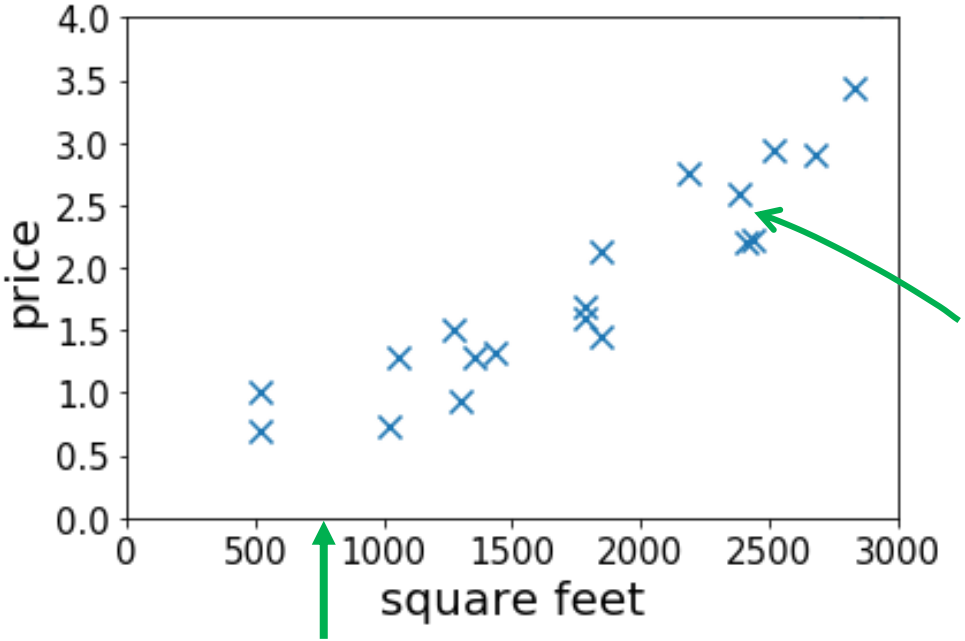
Taxonomy of ML



House Price Prediction

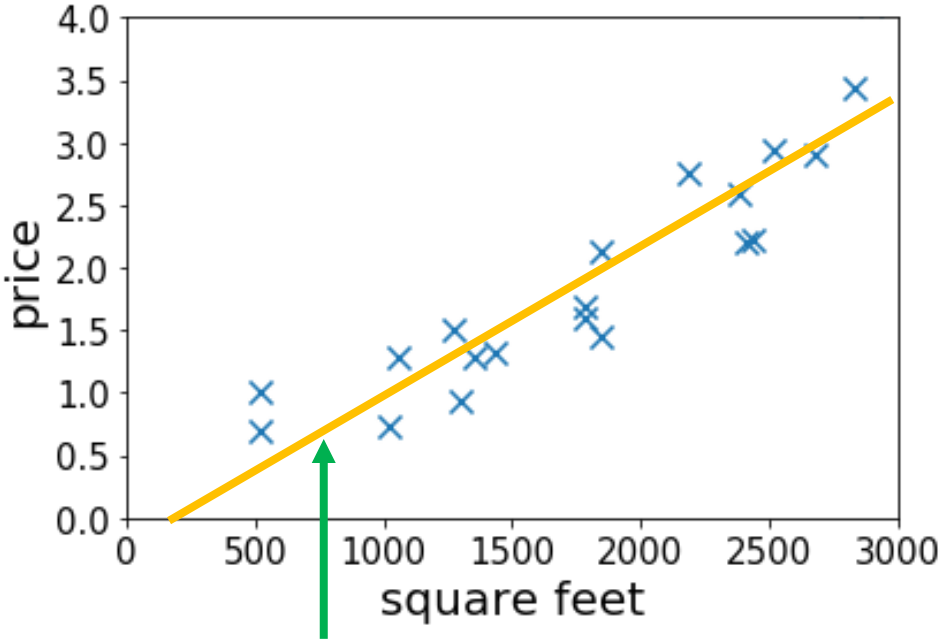
- Given: a dataset that contains n samples
$$\left(x^{(1)}, y^{(1)}\right), \ldots \left(x^{(n)}, y^{(n)}\right)$$

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



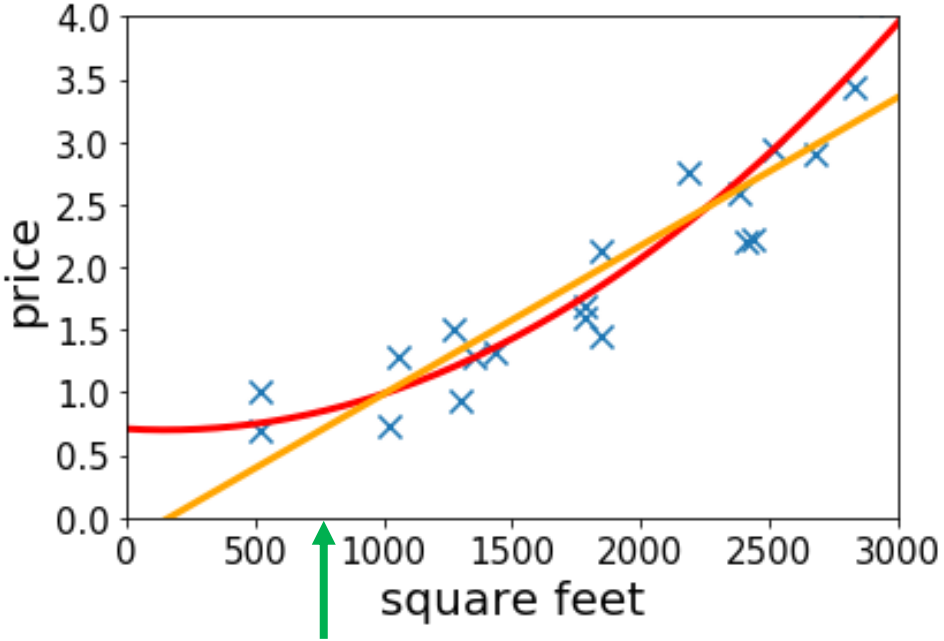
15th sample
$(x^{(15)}, y^{(15)})$

$x = 800$
$y = ?$


4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
price
0
500
1000
1500
2000
2500
3000
square feet

$x = 800$

$y = ?$


4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
price
0
500
1000
1500
2000
2500
3000
square feet

$x = 800$

$y = ?$

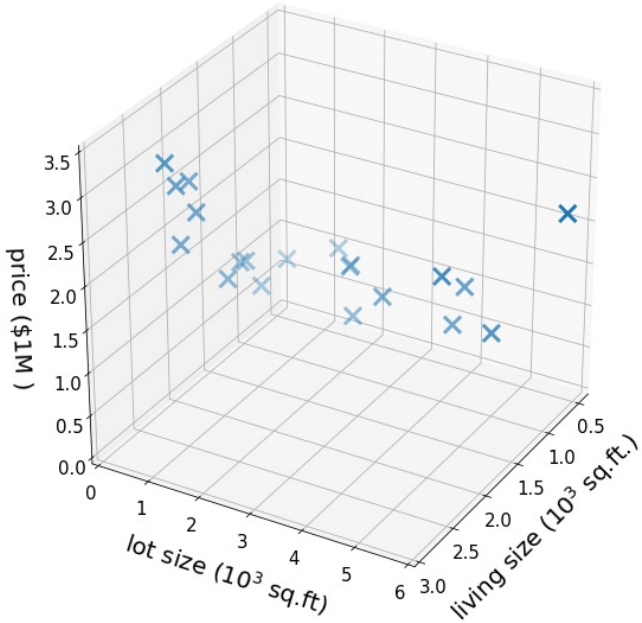
More Features

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

(size, lot size) → price

features/input label/output

- Dataset: $\left(x^{(1)}, y^{(1)}\right), \ldots, \left(x^{(n)}, y^{(n)}\right)$ where $x^{(i)} = (x_1^{(i)}, x_2^{(i)})$
- *Supervision* refers to $y^{(1)}, \ldots, y^{(n)}$



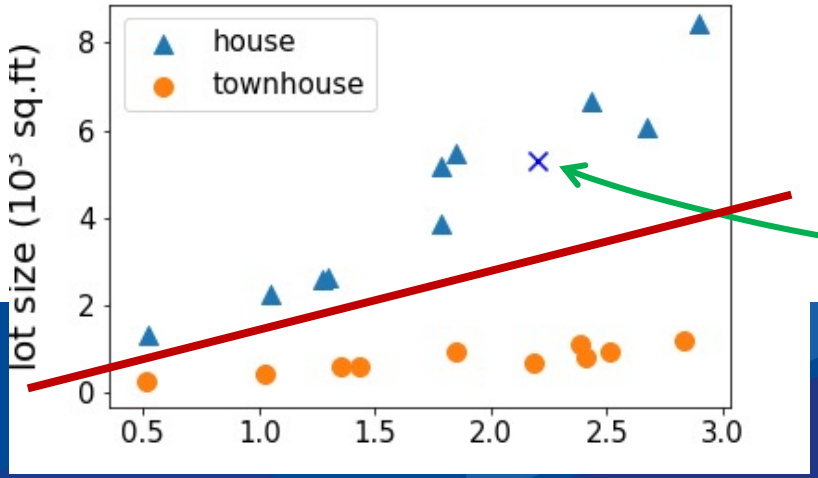
Even More Features

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	PRICE
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	18.7	396.90	5.33	36.2

x y

Regression vs Classification

- Regression: if y 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



y = house or townhouse?

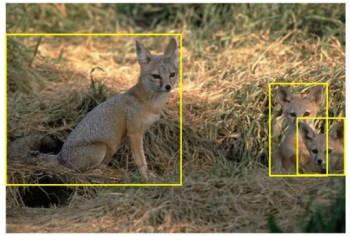
Supervised Learning in Computer Vision

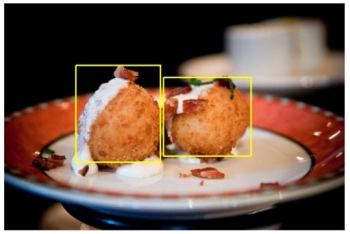
- Image Classification
  - x = raw pixels of the image
  - y = the main object

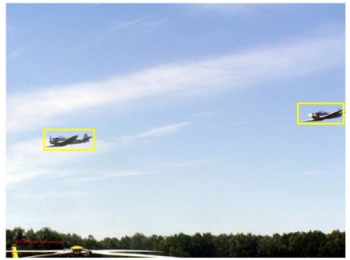
ImageNet Large Scale Visual Recognition Challenge. Russakovsky et al.'2015

Supervised Learning in Computer Vision

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


kit fox


croquette

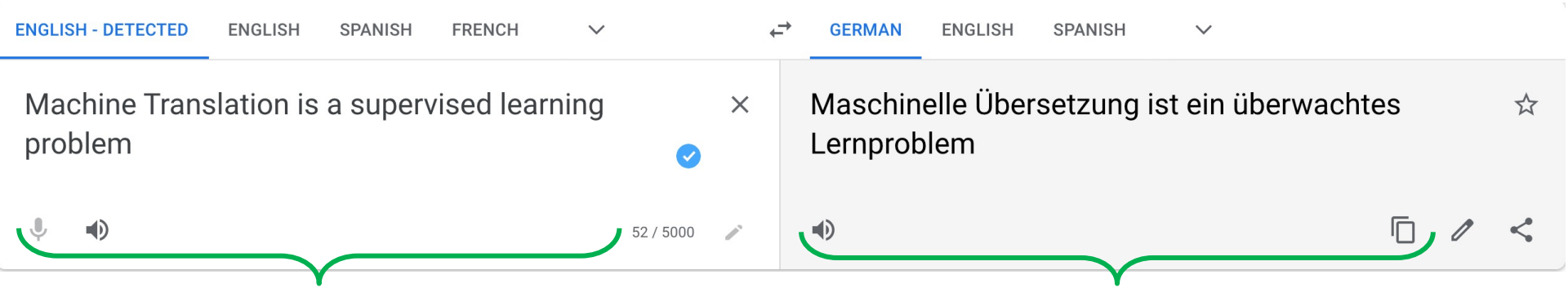

airplane


frog

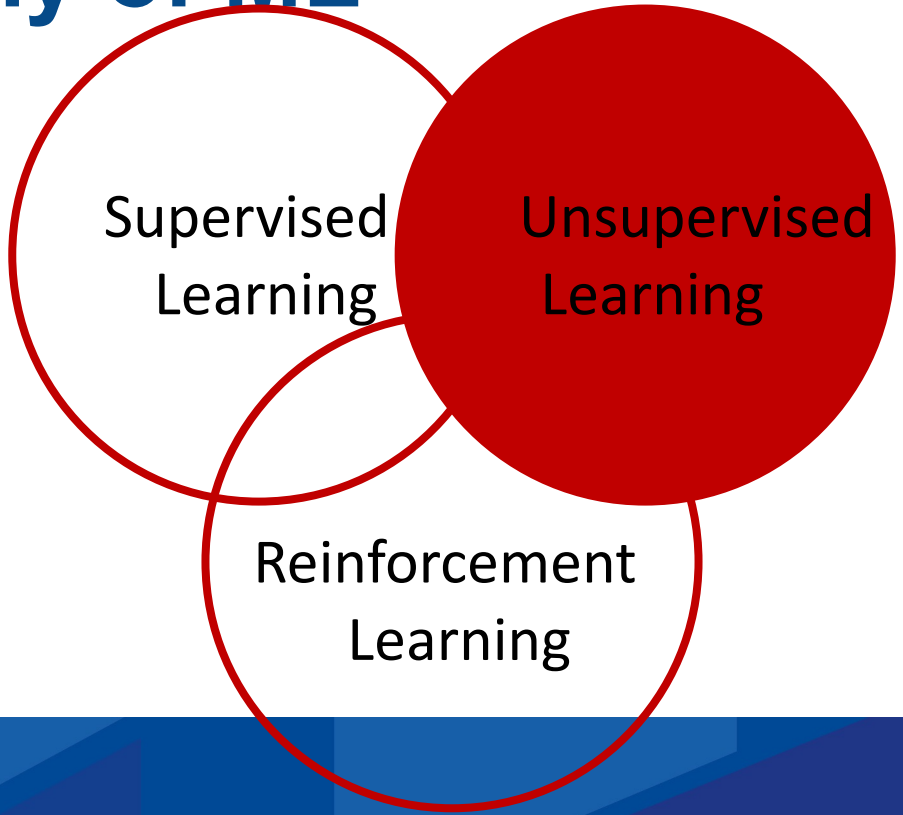
ImageNet Large Scale Visual Recognition Challenge. Russakovsky et al.'2015

Supervised Learning in NLP

- Machine translation



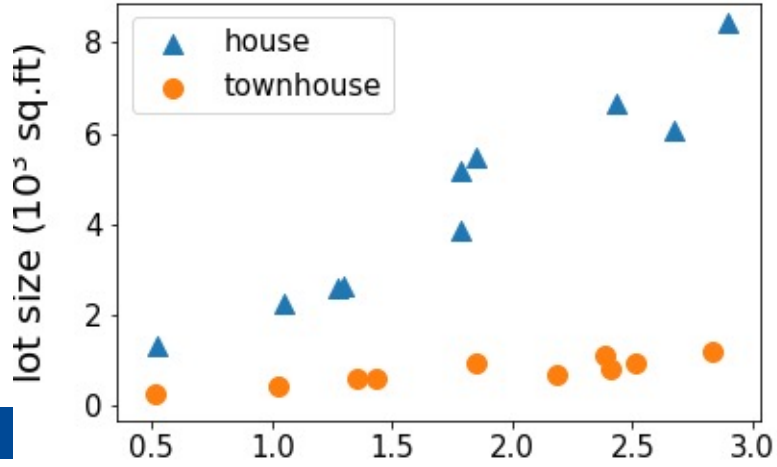
Taxonomy of ML



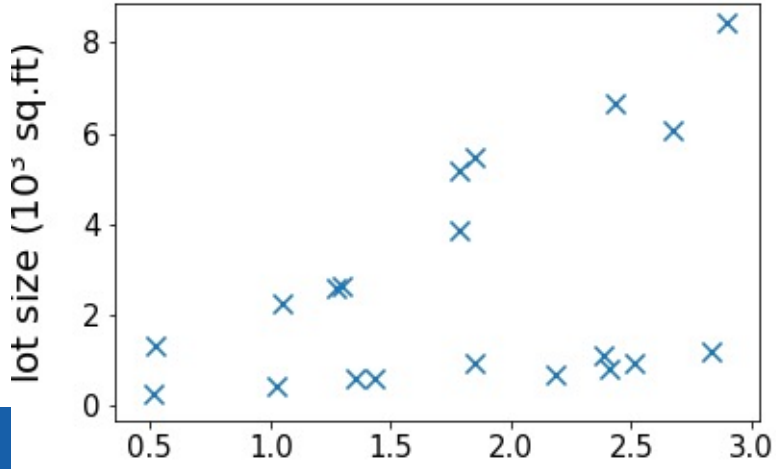
Unsupervised Learning

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

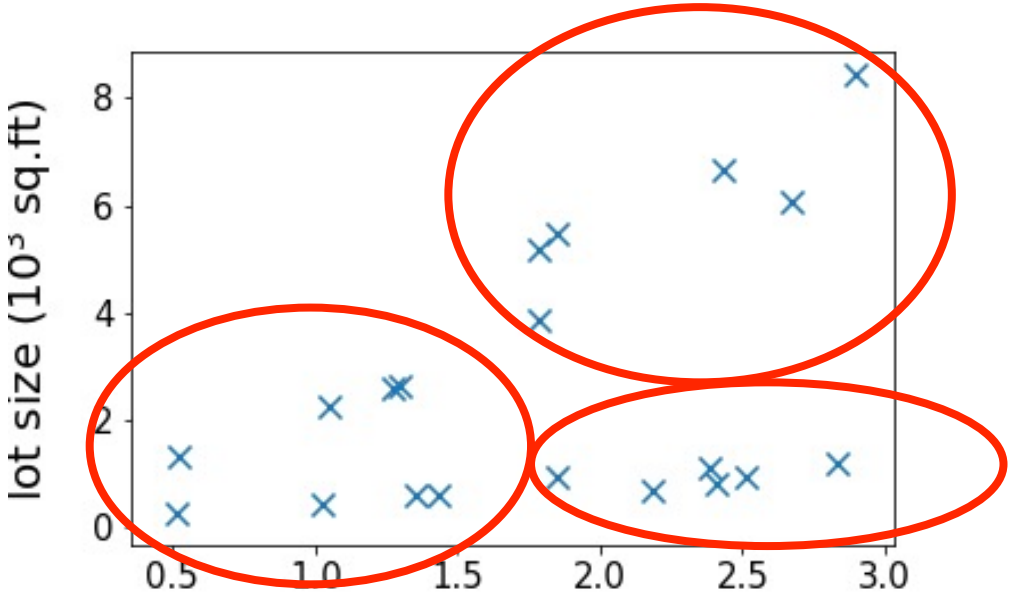
supervised



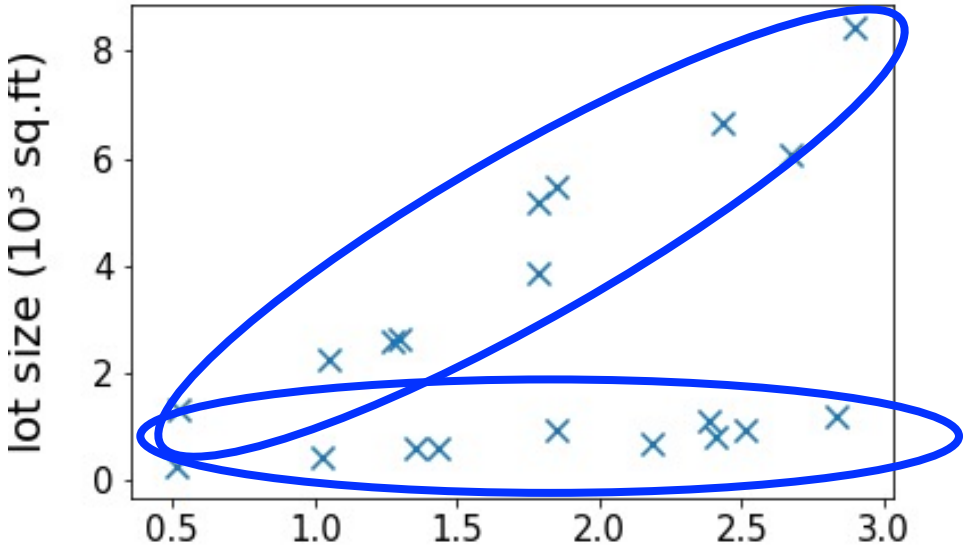
unsupervised



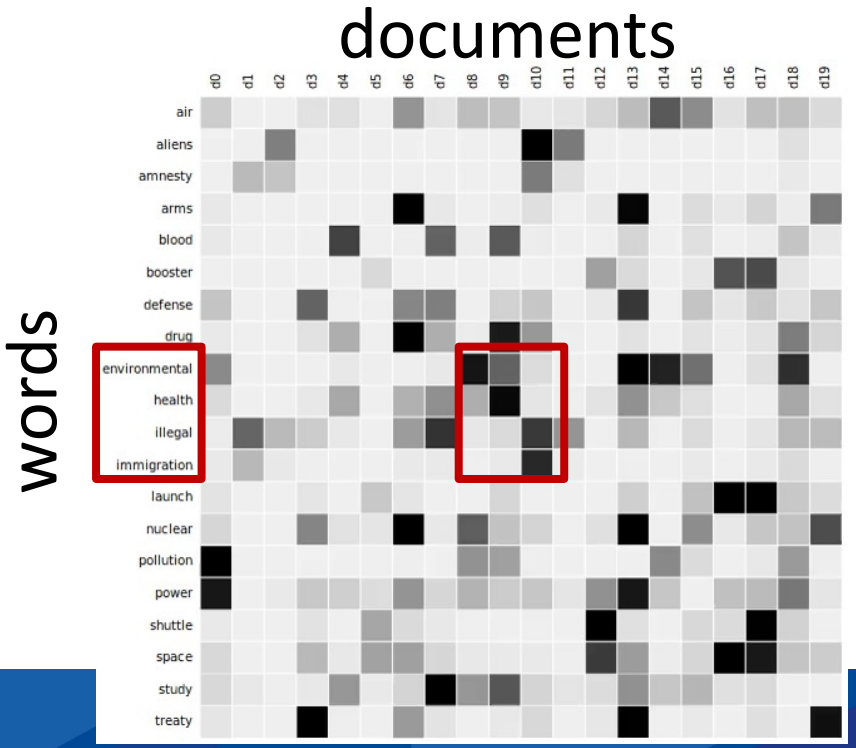
Clustering



Clustering



Latent Semantic Analysis



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

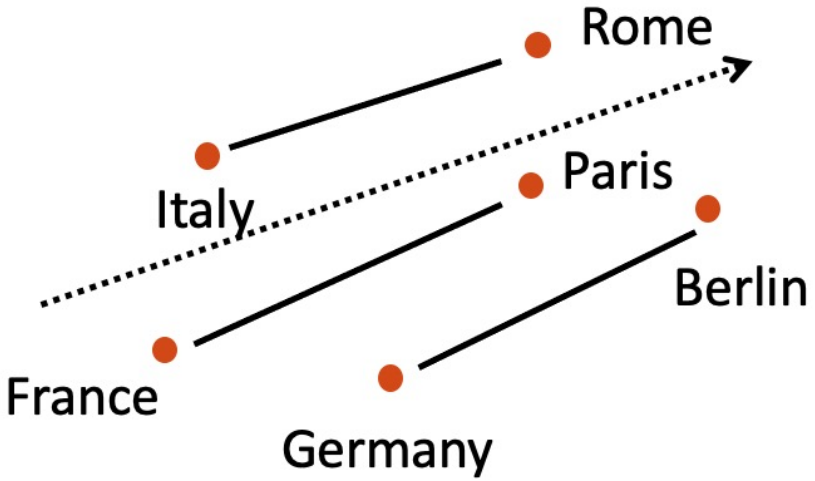
Word Embeddings

Represent words by vectors

word $\xrightarrow{\text{encode}}$ vector

relation $\xrightarrow{\text{encode}}$ direction

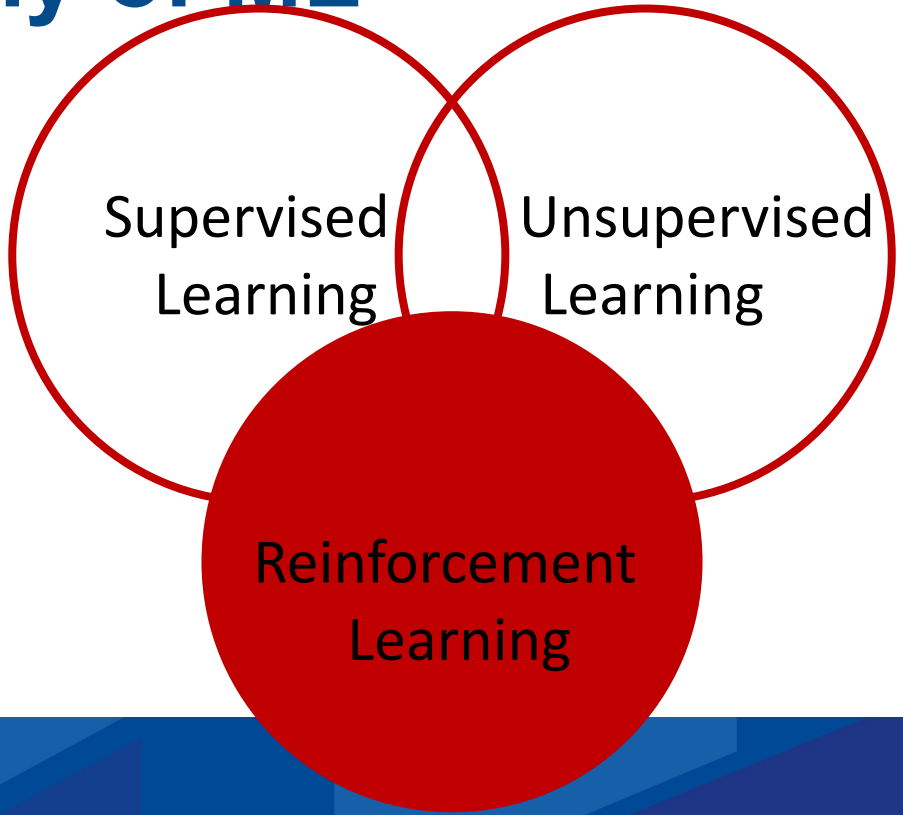
Unlabeled dataset



Word2vec [Mikolov et al'13]

GloVe [Pennington et al'14]

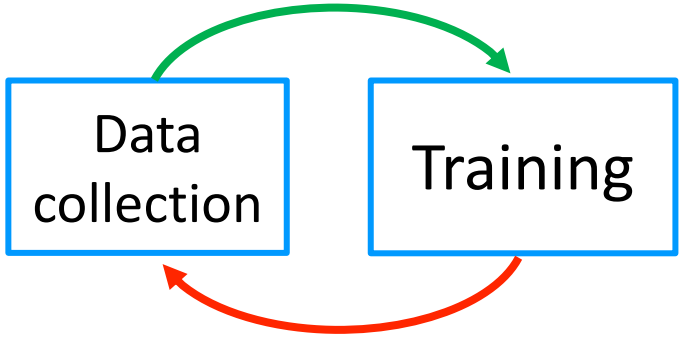
Taxonomy of ML



Learning from Feedbacks

The algorithm can collect data interactively through trial and error in an environment

Try the strategy and collect feedbacks



Improve the strategy based on the feedbacks

https://www.youtube.com/watch?v=V1eYniJ0Rnk

RL Successes

This Course

Mainly supervised learning

- Preprocessing
- Training
- Validation

Applications

- Computer Vision
- NLP

Summary

- Machine Learning is a hot topic in AI and most other fields at the moment
- There are three main types of ML
  - Supervised, unsupervised, reinforcement learning
  - Can work well if data is abundant
- ML has achieved super-human performance on some tasks

Acknowledgment

- Some of the slides were adapted from CS229 Machine Learning course at Stanford

University of
South Australia

Questions?