Machine Learning

Amin Golzari Oskouei

a.golzari@azaruniv.ac.ir

a.golzari@tabrizu.ac.ir

https://github.com/Amin-Golzari-Oskouei

Azarbaijan Shahid Madani University
2023

Machine Learning : Introduction

References (1)

- Introduction To Machine Learning

[Ethem Alpaydin, 2014]

- Machine Learning: A Probabilistic Perspective

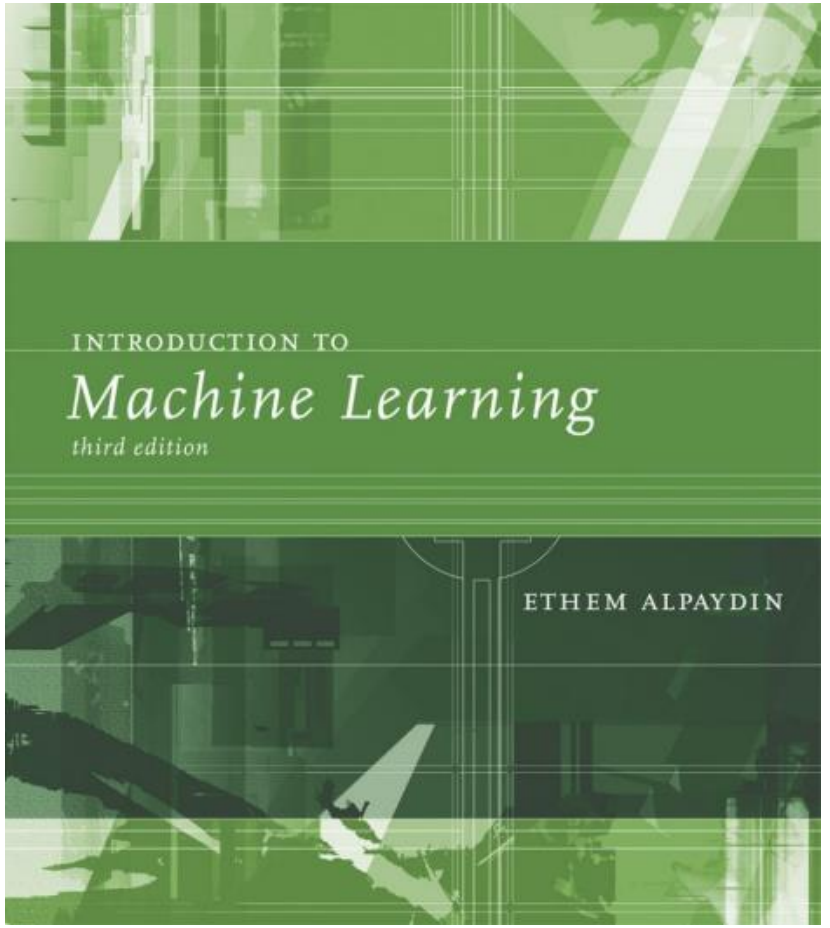
[Kevin P. Murphy, 2012]

- Pattern Recognition and Machine Learning

[Christopher M. Bishop, 2006]

- Machine Learning

[Razavi S. Naser, 2018]



References (2)

- Introduction To Machine Learning

[Ethem Alpaydin, 2014]

- Machine Learning: A Probabilistic Perspective

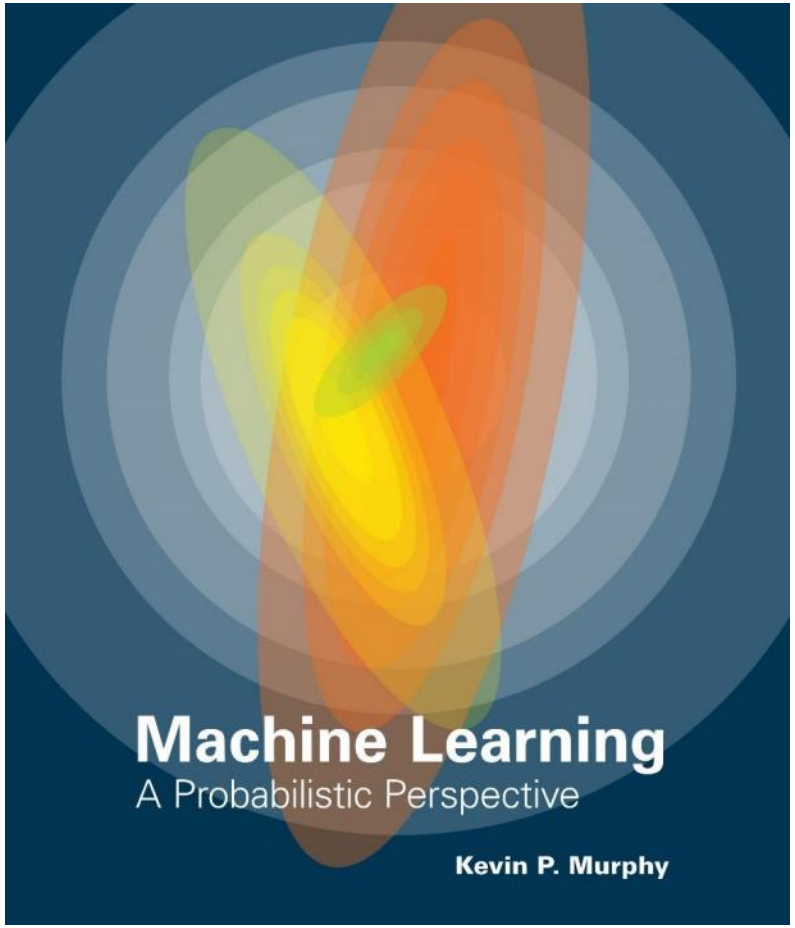
[Kevin P. Murphy, 2012]

- Pattern Recognition and Machine Learning

[Christopher M. Bishop, 2006]

- Machine Learning

[Razavi S. Naser, 2018]



References (4)

- Introduction To Machine Learning

[Ethem Alpaydin, 2014]

- Machine Learning: A Probabilistic Perspective

[Kevin P. Murphy, 2012]

- Pattern Recognition and Machine Learning

[Christopher M. Bishop, 2006]

- Machine Learning

[Razavi S. Naser, 2018]

References (5)

□ These slides have been translated and edited based on Dr. Seyed Naser Razavi's machine learning course.

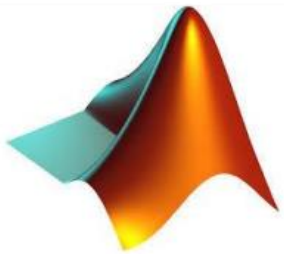
Requirements

- Algorithm analysis and design methods
  - Analysis of computational complexity of learning algorithms
- Linear algebra
  - Matrices, vectors, matrix operations and linear equations
  - Inverse matrix, eigenvectors, order of matrix, decomposition of singular values
- Multivariate calculus
  - Derivative, integral, tangent planes
- Probabilities
  - Random variables, expected value, variance and …

Evaluation

- Tasks [45%]
  - Theoretical topics
  - Programming
- Final Exam [50%]
- Effective attendance in class [5%]

OCTAVE



MATLAB

Table of Contents

Some Quotes

"If you invent a breakthrough in artificial intelligence, so machines can learn, that is worth 10 Microsofts."

-Bill Gates (Former CEO of Microsoft)

"Machine Learning is the next internet."

-Anthony Tether (Director of DARPA)

"Machine learning is going to result in a real revolution."

-Greg Papadopoulos (Creator of Redshift)

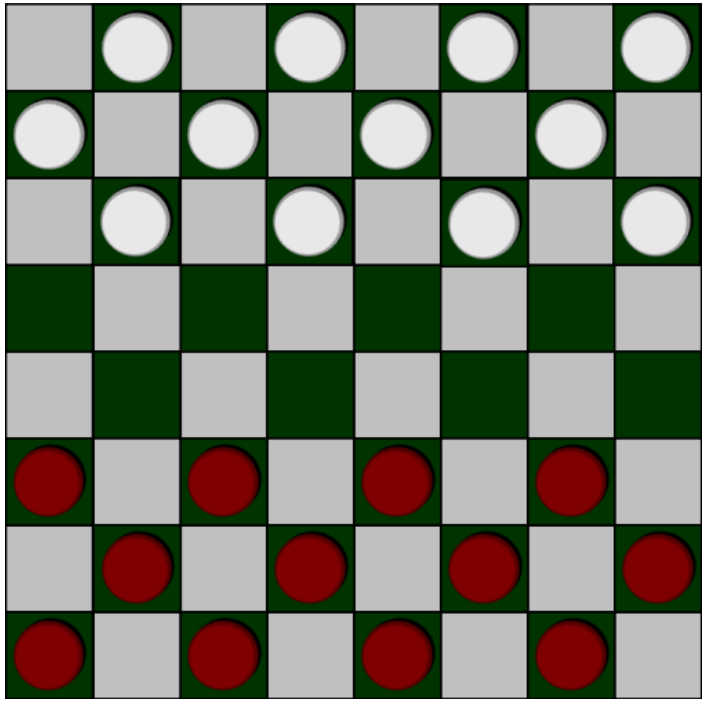
What is Machine Learning?

Machine Learning: Definitions

- Arthur Samuel [1959]

  “Field of study that gives computers the ability to learn without being explicitly programmed.”

- Samuel’s Checkers Player [Samuel, 1959, 1967]

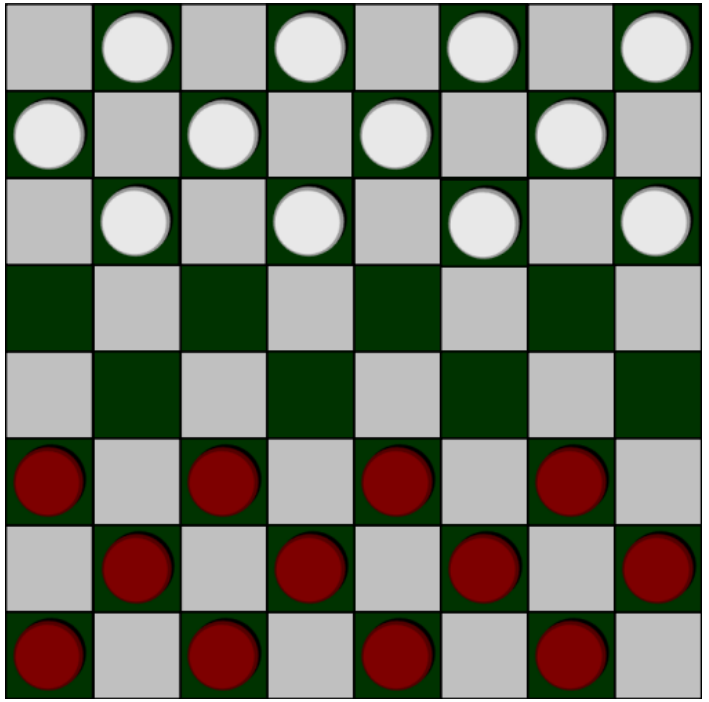


Machine Learning: Definitions

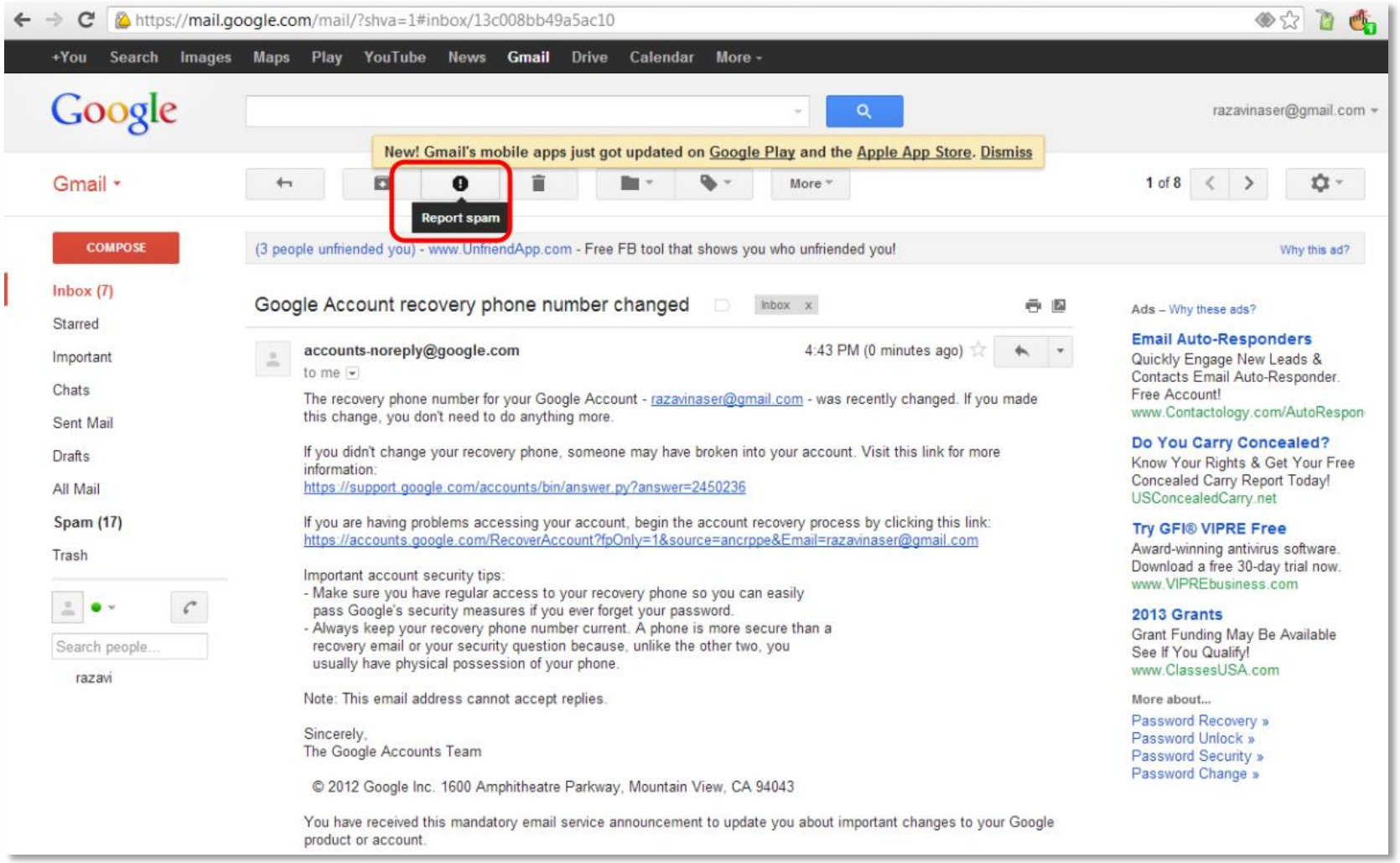
- Tom Mitchell [1998]

  "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."

- Example: Checkers Game
  - Task: Playing checkers
  - Experience: Playing thousands of time against yourself
  - Efficiency Measure: Number of wins against new opponents



Example: Detecting Spam Mail



Example: Detecting Spam Mail

- Example: Detecting Spam Mail

  Let's say your email program allows you to mark incoming emails as spam and learns how to detect spam mails accordingly.

- Task: Categorizing emails as spam or email
- Experience: Monitoring which emails you mark as spam
- Efficiency Measure: The number of emails that are correctly categorized

Types of Machine Learning Methods

- Machine learning Improving the machine's performance in performing a task with experience.

- Q. How can a machine know that its performance has improved?

  - We can give the machine the correct answer for some limited examples of input in the hope that it can generalize it to other examples - supervised learning

  - We can tell the machine to what extent its answer was correct (for example by giving a point) and the machine itself is responsible for finding the correct answers - reinforcement learning

  - We may not give the machine any information about the correct answer and only ask the machine to find inputs that have common features - unsupervised learning

Supervised Learning

Supervised Learning

- Input: A training set where the correct answer is given for each input.

$$\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$$

Training Set

- Goal: Find a suitable approximation for the following mapping:

$$f: X \rightarrow Y$$

- Example:
  - Spam Detection: Mapping emails to the {Spam, Non-Spam} collection
  - Digit Recognition: Mapping a set of pixels to the set {0, 1, 2, ..., 9}
  - Cancer Diagnosis: Mapping medical data to the set {Malignant, Benign}

Example: Detecting Spam Mail

- Input: Email
- Output: Spam, Non-Spam

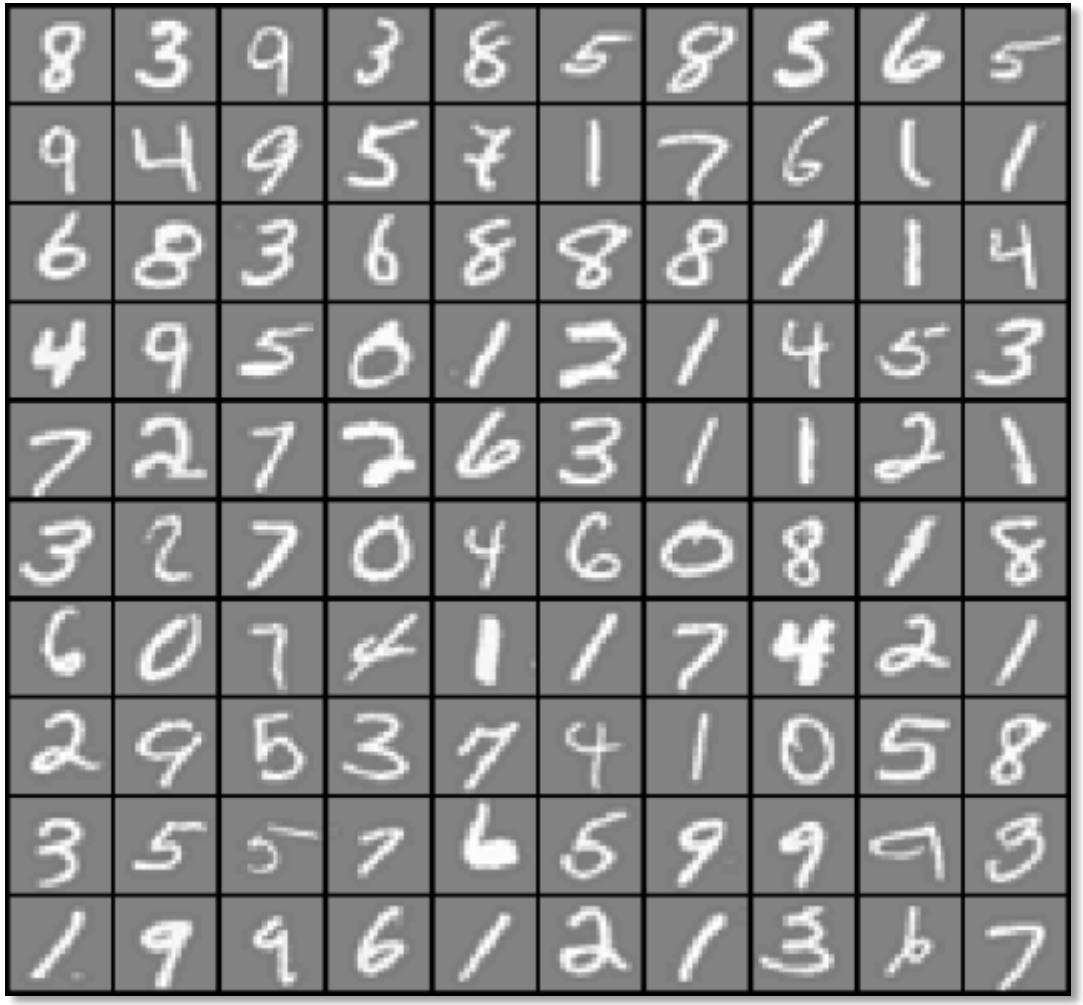
Dear Sir. First, I must solicit your confidence in this transaction, this is by virture of its nature as being utterly confidencial and top secret. …

To be removed from future mailings, simply reply to this message and put "remove" in the subject. 99 million email addresses for only $99

Ok, I know this is blatantly OT but I'm beginning to go insane. Had an old Dell Dimension XPS sitting in the corner and decided to put it to use, I know it was working pre being stuck in the corner, but when I plugged it in, hit the power nothing happened.

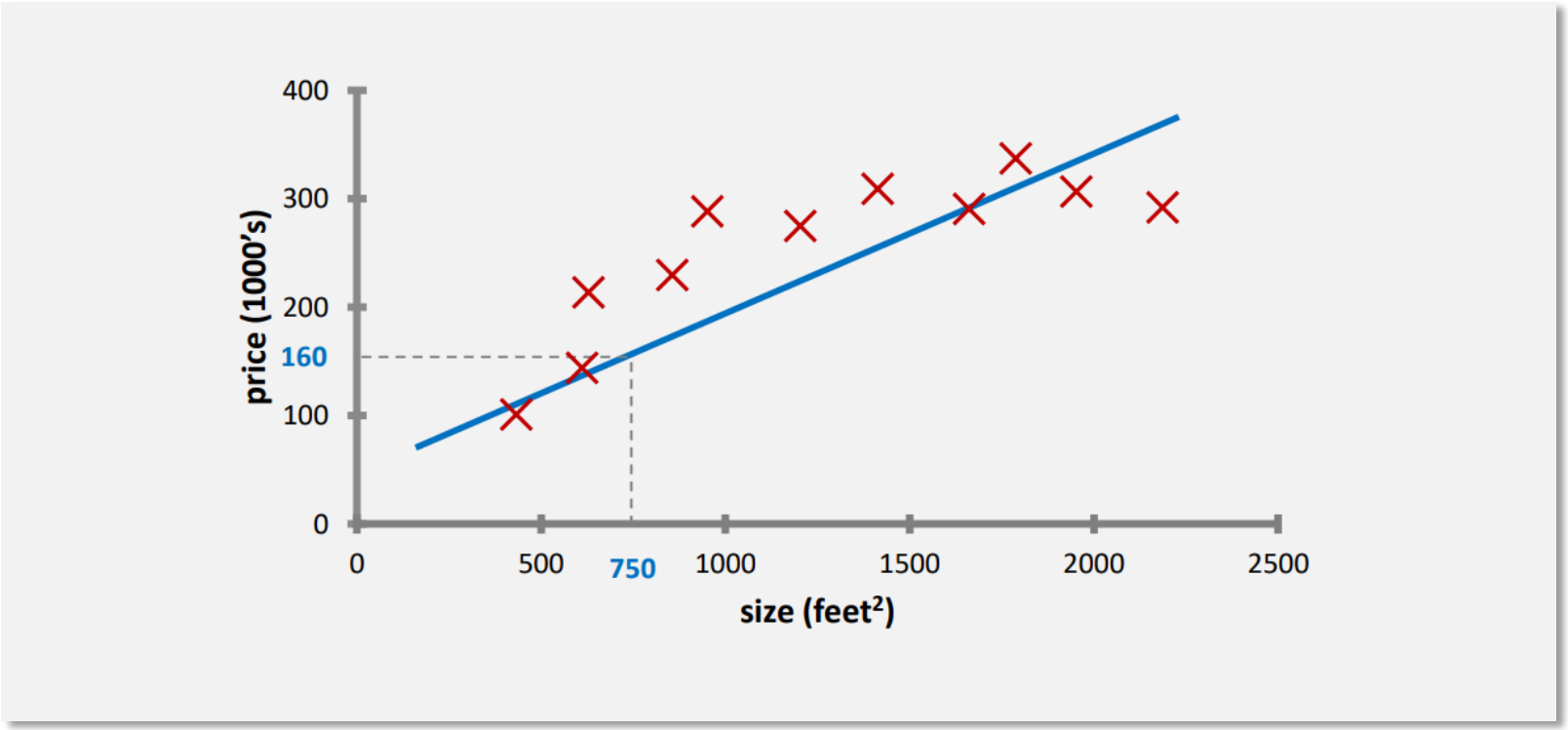
Example: Recognizing Handwritten Numbers

- Input: Image of A Number
- Output: A Number



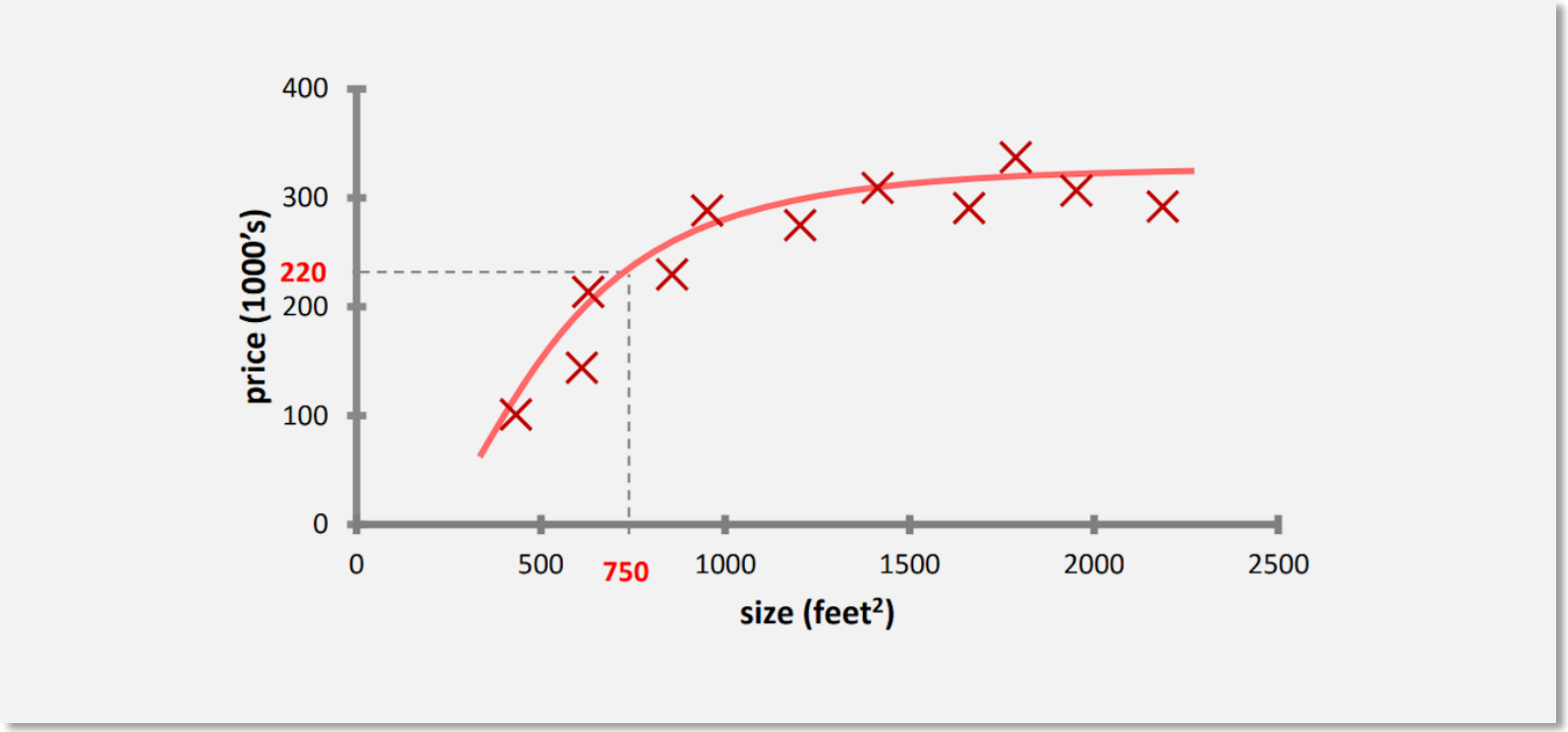
Example: Pricing A House

- Input: The Size of The House [In Feet^2]
- Output: Estimated Price

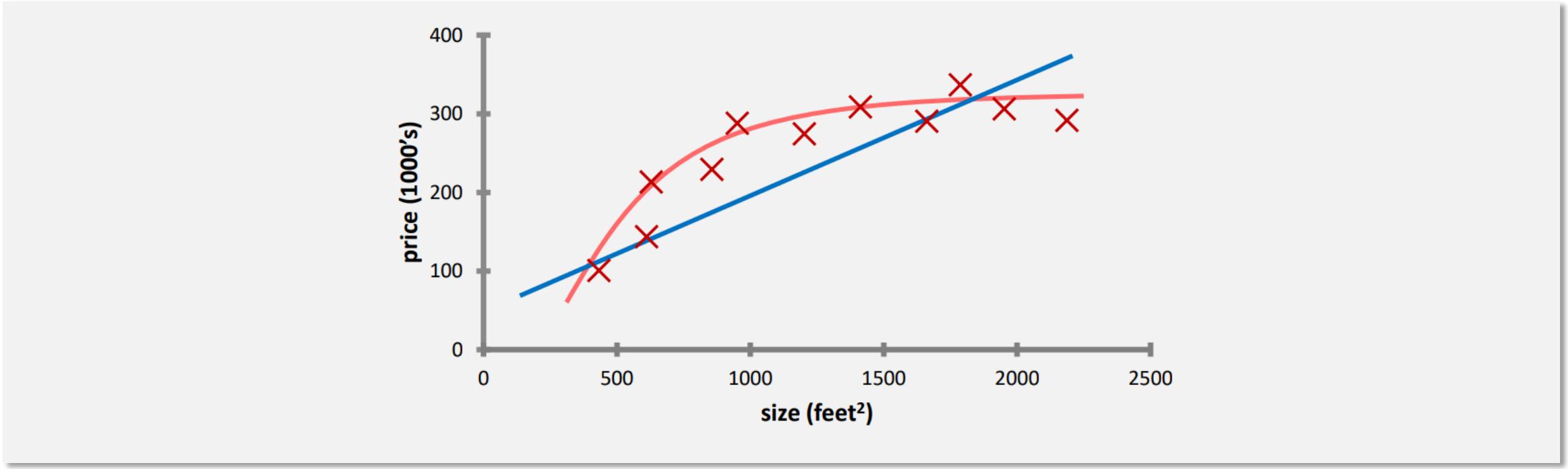


Example: Pricing A House

- Q. Which one is better? A linear function or a quadratic function?



Example: Pricing A House



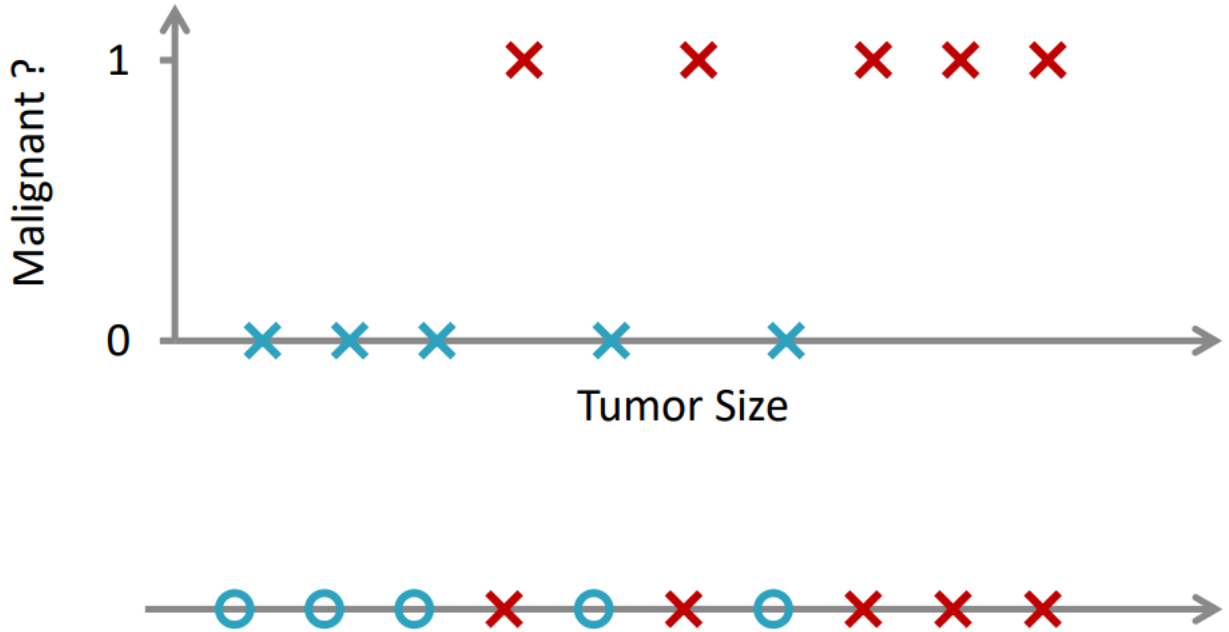
☐ Supervised Learning:

For each training sample, a "correct answer" is given.

☐ Regression:

Predicting quantities with continuous values (such as the price of a house)

Example: Cancer Diagnosis (Malignant, Benign)



☐ Supervised Learning:

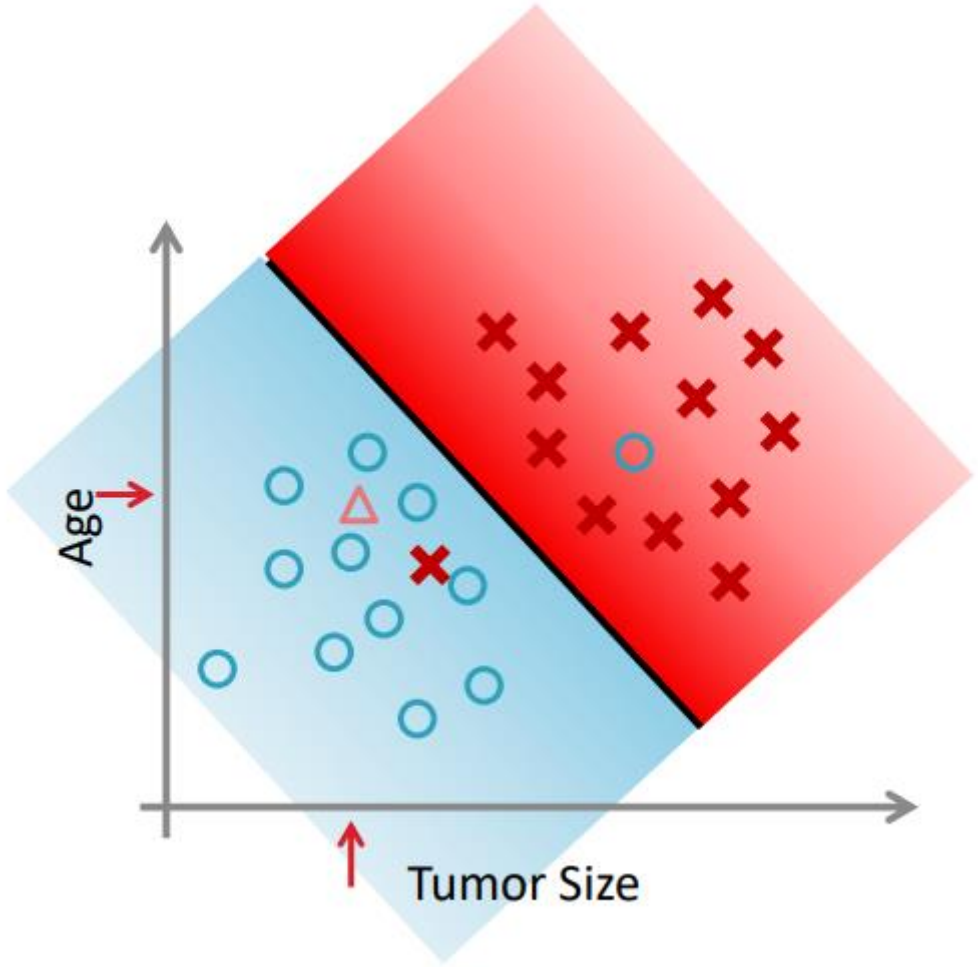
For each educational example, a "correct answer" is given.

☐ Classification:

Predicting quantities with discrete values (such as zero and one).

Example: Cancer Diagnosis (Malignant, Benign)

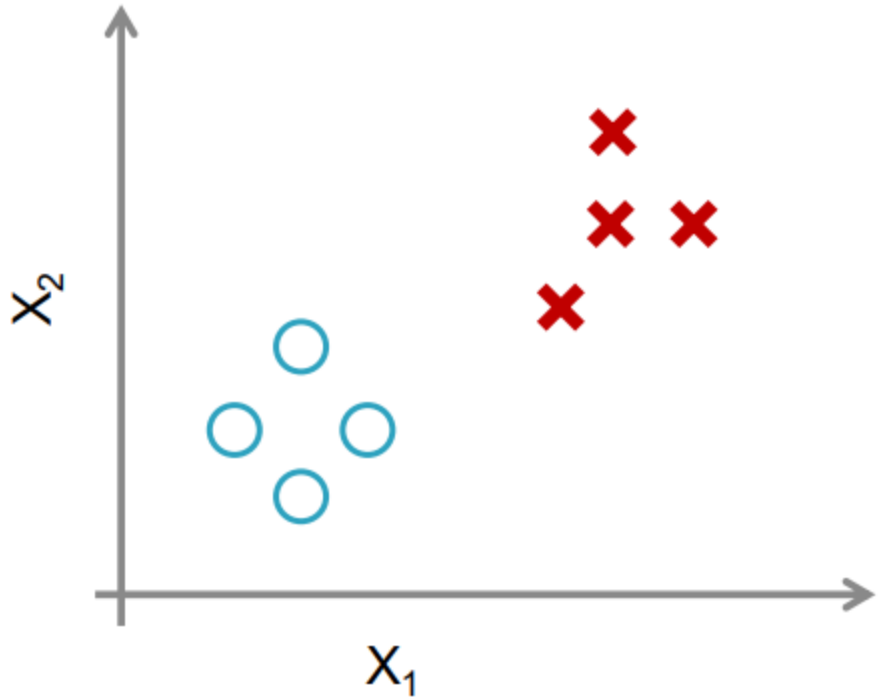
- Other Features:
  - Uniformity of cell shape
  - Uniformity of cell size
  - …



Unsupervised Learning

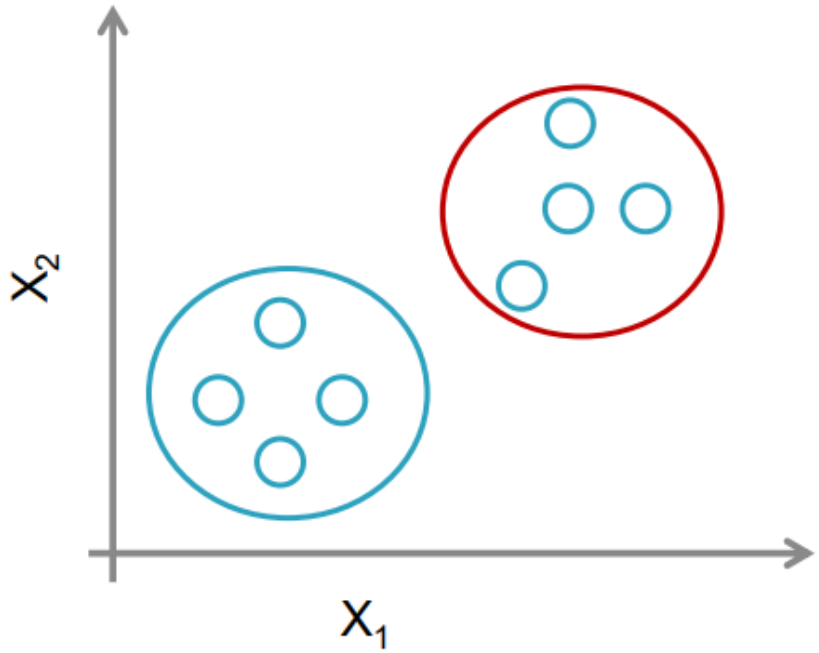
Supervised Learning

- Supervised Learning: For each training example, a "correct answer" is given.



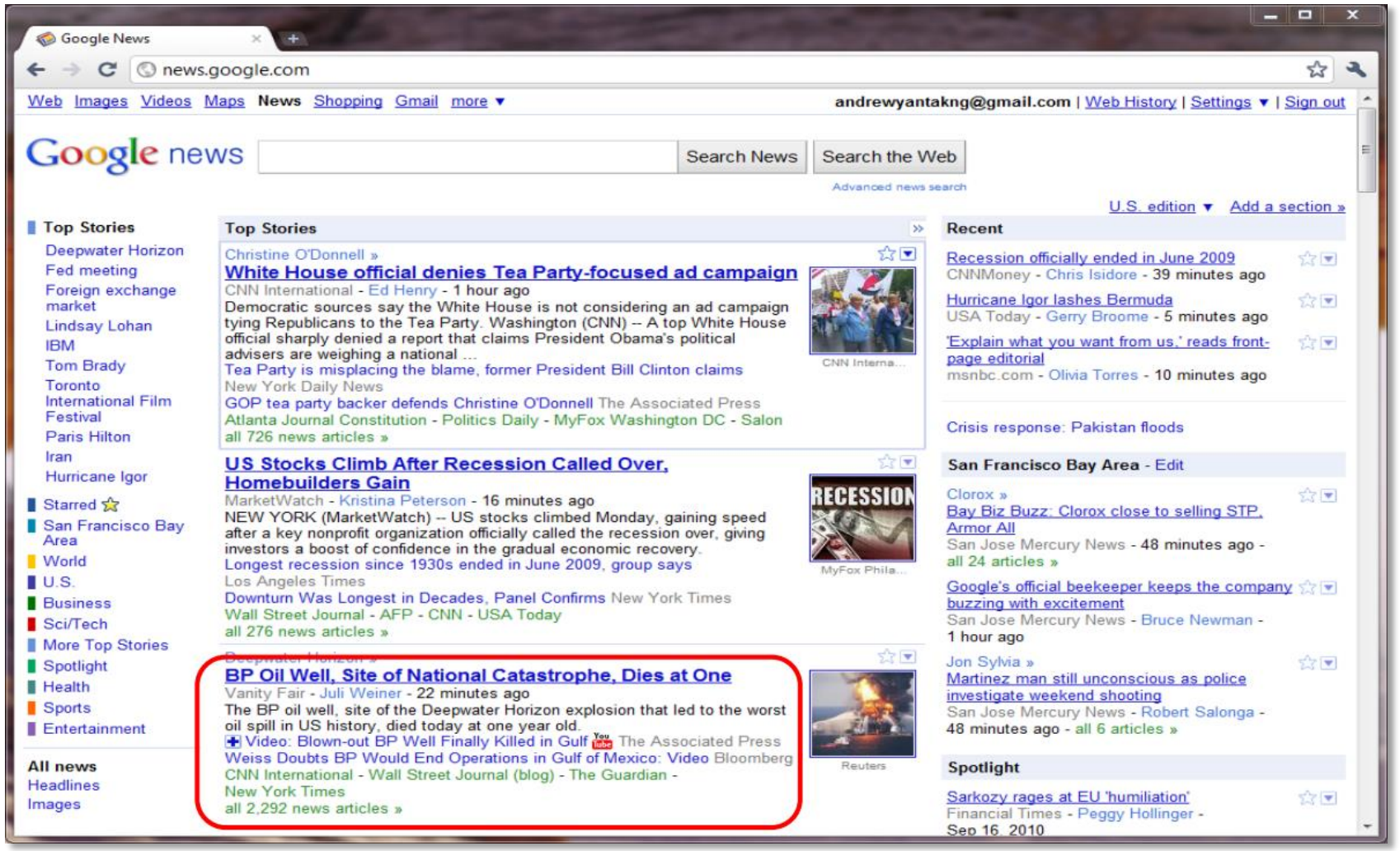
Unsupervised Learning

- Unsupervised Learning: No information about correct answers is given!

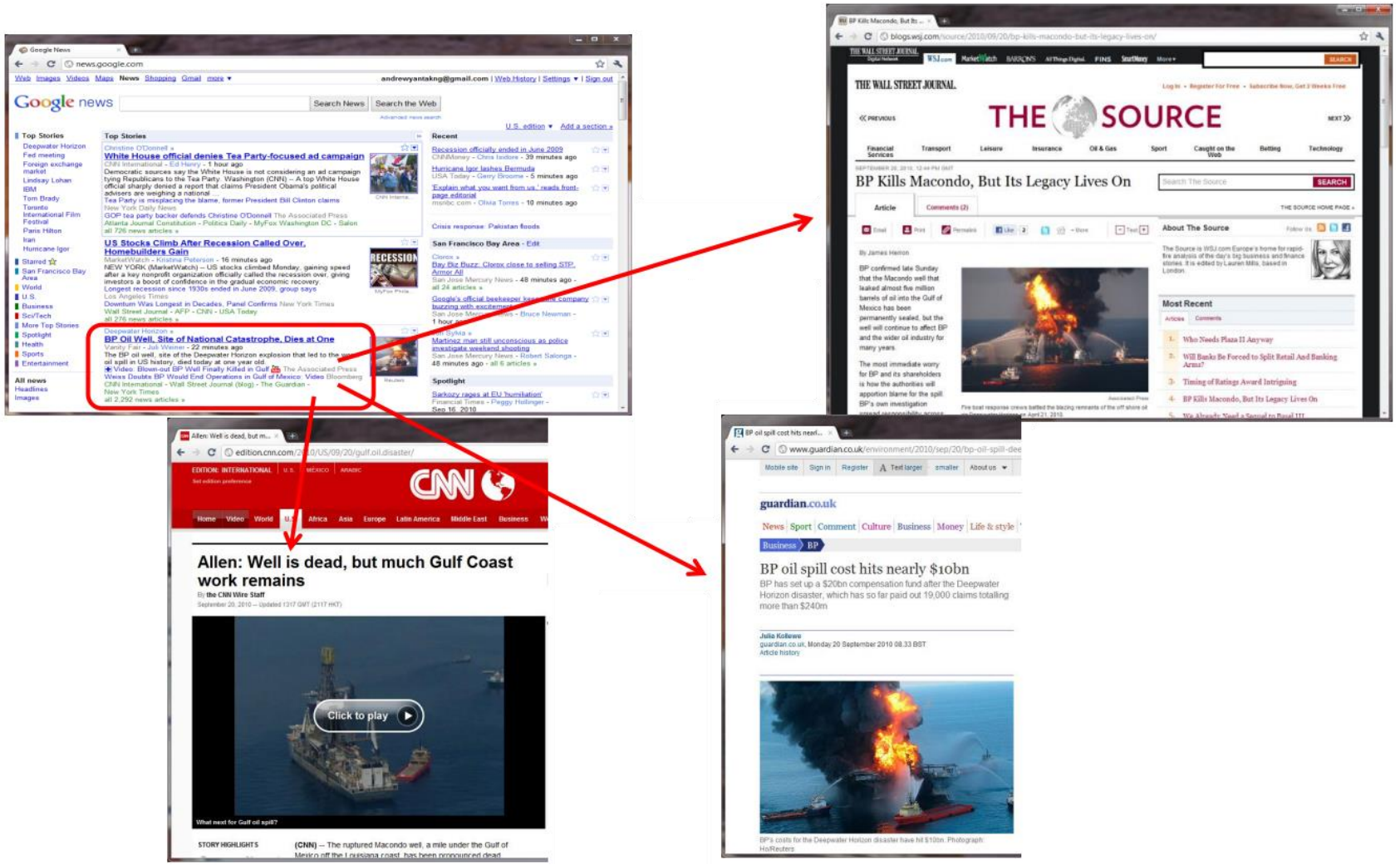


- Goal: Detecting structure in input data (grouping similar data).

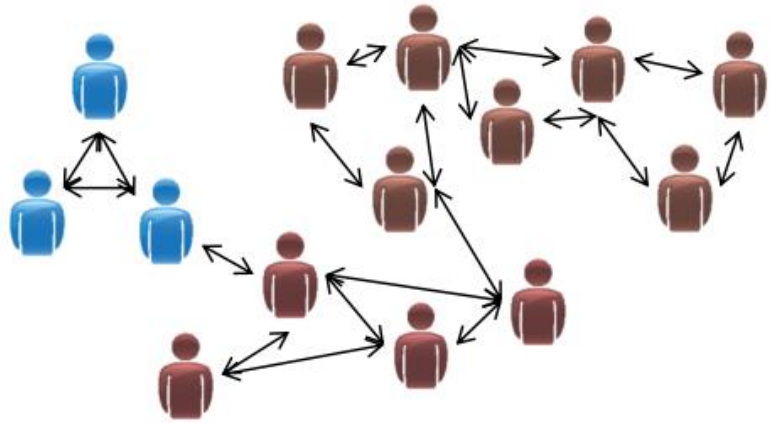
Application of Clustering: Grouping Related News



Application of Clustering: Grouping Related News



Some Other Applications of Unsupervised Learning



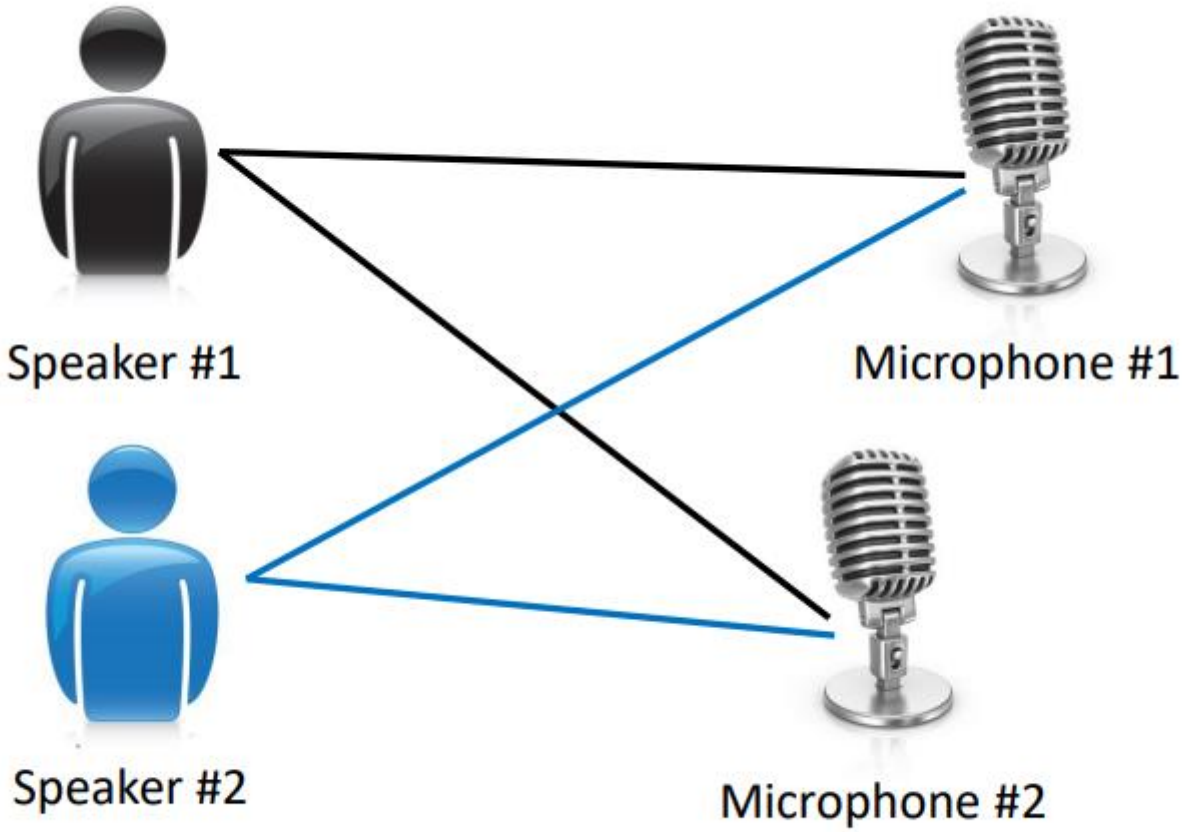
Analysis of social networks

Organization of computing clusters (data centers)

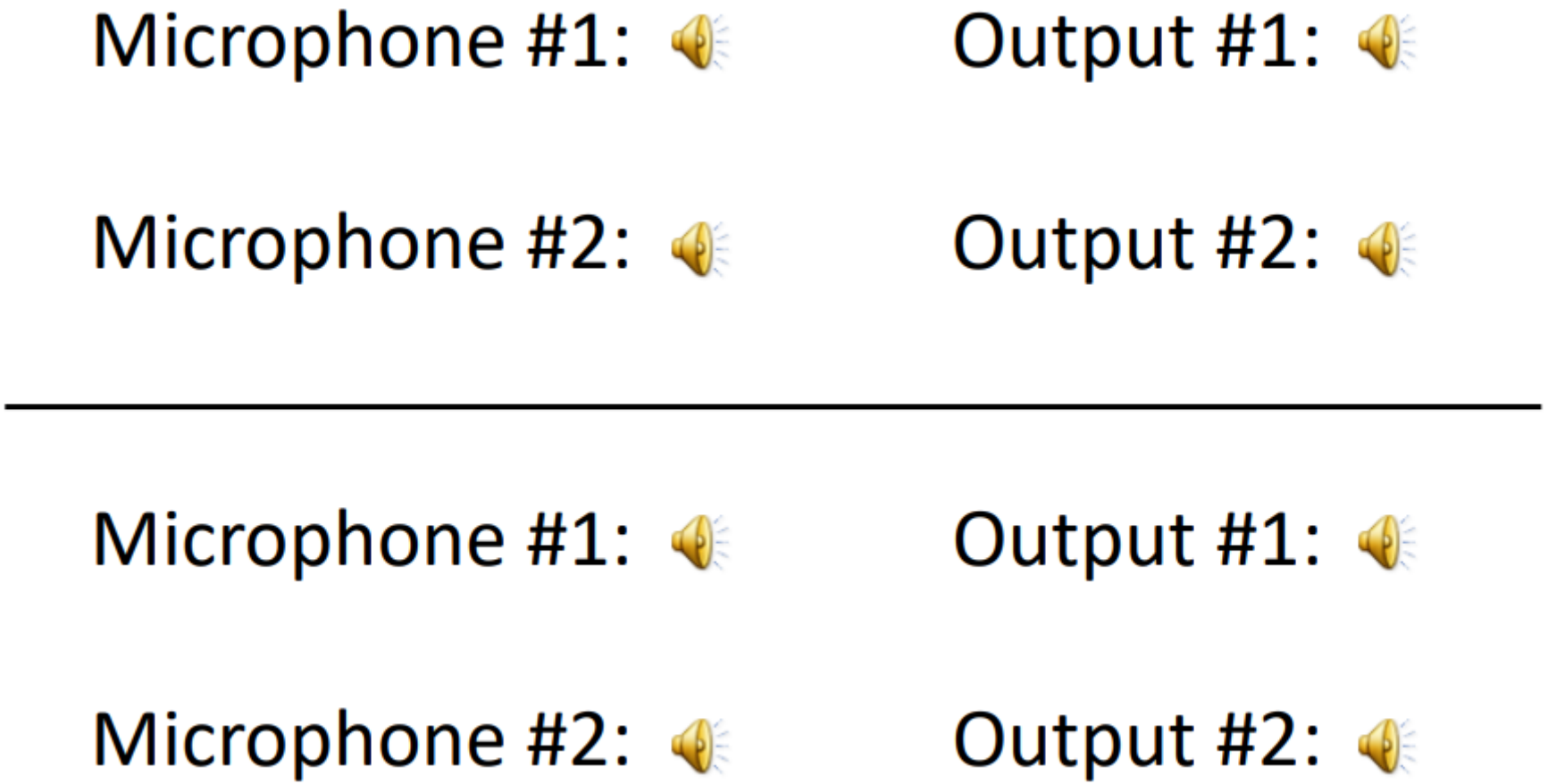
Analysis of astronomical data (how galaxies form)

Market segmentation

Cocktail Party Problem



Cocktail Party Problem



Cocktail Party Problem Algorithm

- Octave Code:

```
[W, s, v] = svd((repmat(sum(x .* x, 1), size(x, 1), 1) .* x) * x');
```

A Question For Class

- For which one of the following situations should an unsupervised learning algorithm be used?
  - Developing a program to filter spam mails by having some normal emails and some spam emails
  - Group a collection of newly found articles on the web by topic
  - Diagnosing diabetes in new patients by having data on a number of healthy and diabetic people
  - Grouping a set of customers into several different market segments by having a database about customers