

Machine Learning

# Introduction

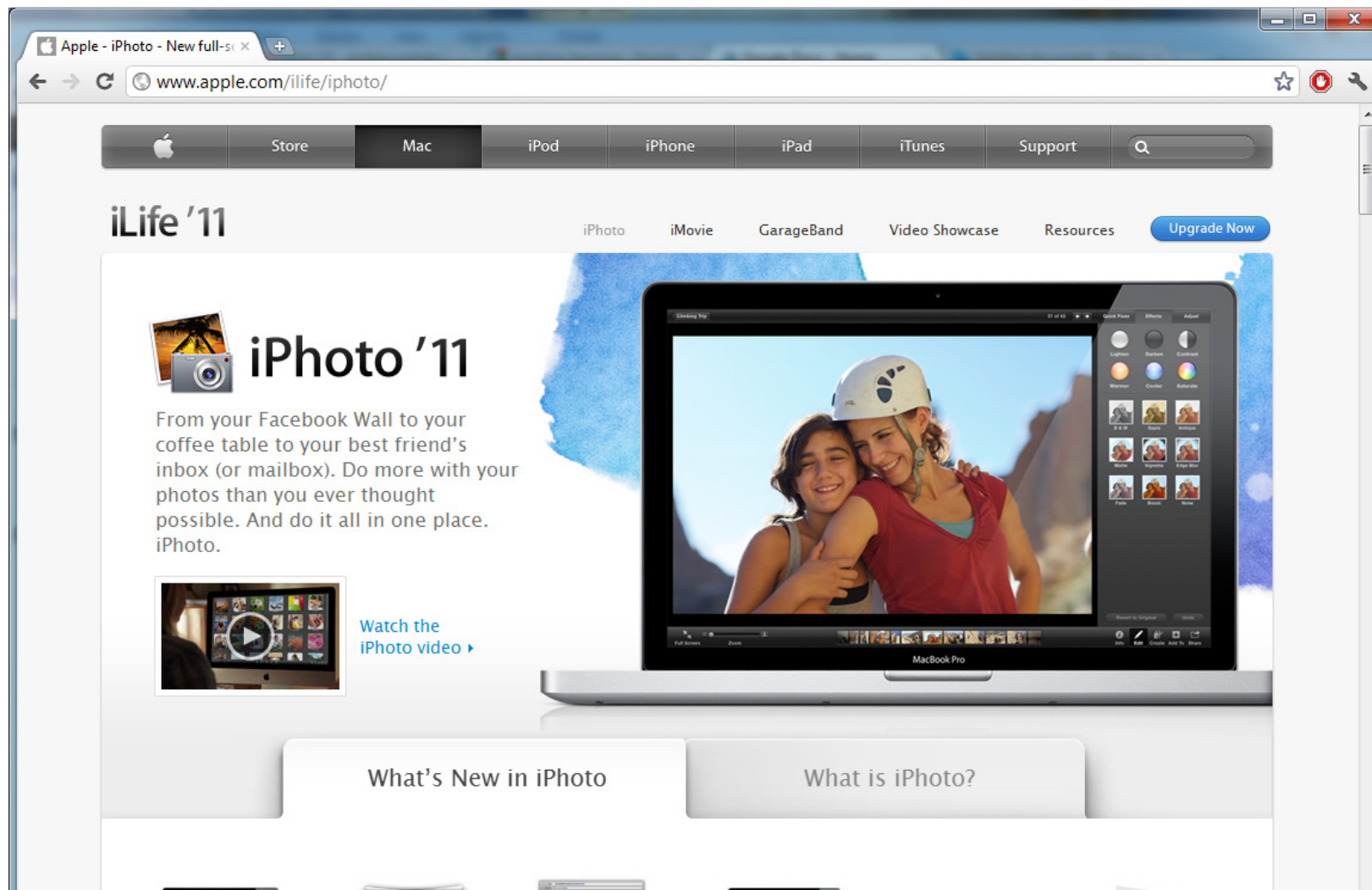
---

# Welcome



**Taught by:**

**Andrew Ng**, CEO/Founder Landing AI; Co-founder, Coursera; Adjunct Professor, Stanford University; formerly Chief Scientist, Baidu and founding lead of Google Brain





# Machine Learning

- Grew out of work in AI
- New capability for computers

## Examples:

- Database mining

Large datasets from growth of automation/web.

E.g., Web click data, medical records, biology, engineering

- Applications can't program by hand.

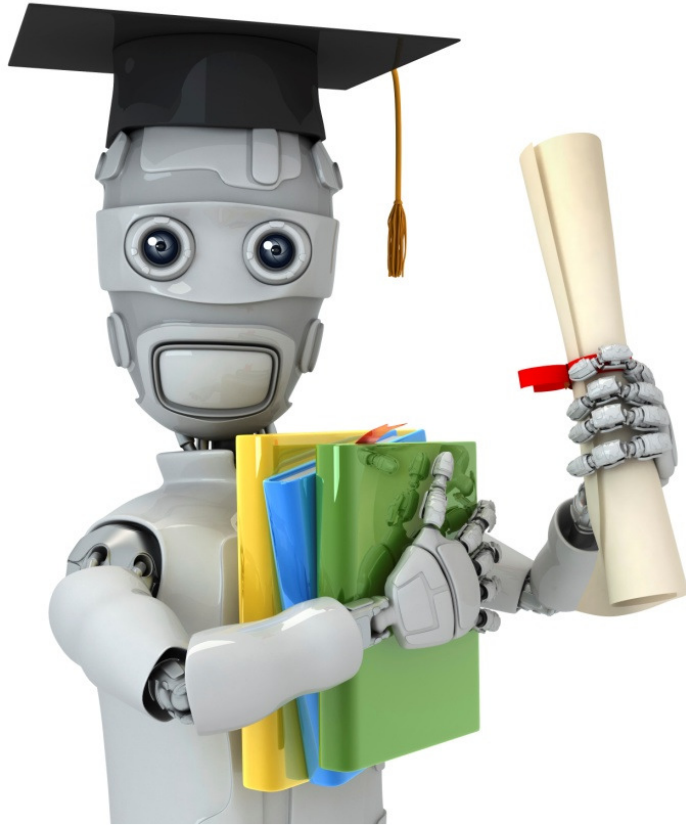
E.g., Autonomous helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.

# Machine Learning

- Grew out of work in AI
- New capability for computers

## Examples:

- Database mining
  - Large datasets from growth of automation/web.
  - E.g., Web click data, medical records, biology, engineering
- Applications can't program by hand.
  - E.g., Autonomous helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.
- Self-customizing programs
  - E.g., Amazon, Netflix product recommendations
- Understanding human learning (brain, real AI).



Machine Learning

# Introduction

---

## What is machine learning

# Machine Learning definition

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

“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  $E$ .”

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task  $T$  in this setting?

- ☐ Classifying emails as spam or not spam.
- ☐ Watching you label emails as spam or not spam.
- ☐ The number (or fraction) of emails correctly classified as spam/not spam.
- ☐ None of the above—this is not a machine learning problem.

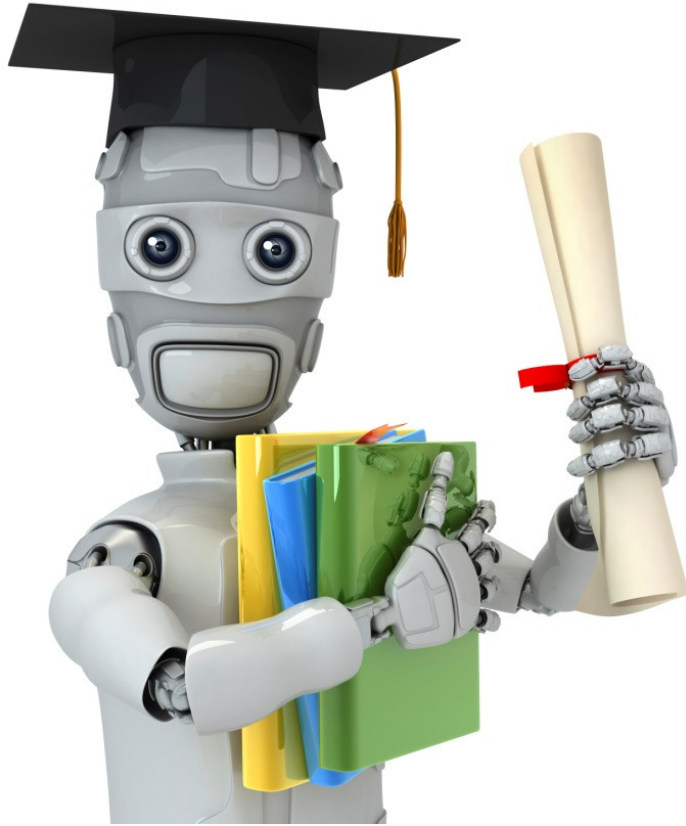


Machine learning algorithms:

- Supervised learning
- Unsupervised learning

Others: Reinforcement learning, recommender systems.

Also talk about: Practical advice for applying learning algorithms.



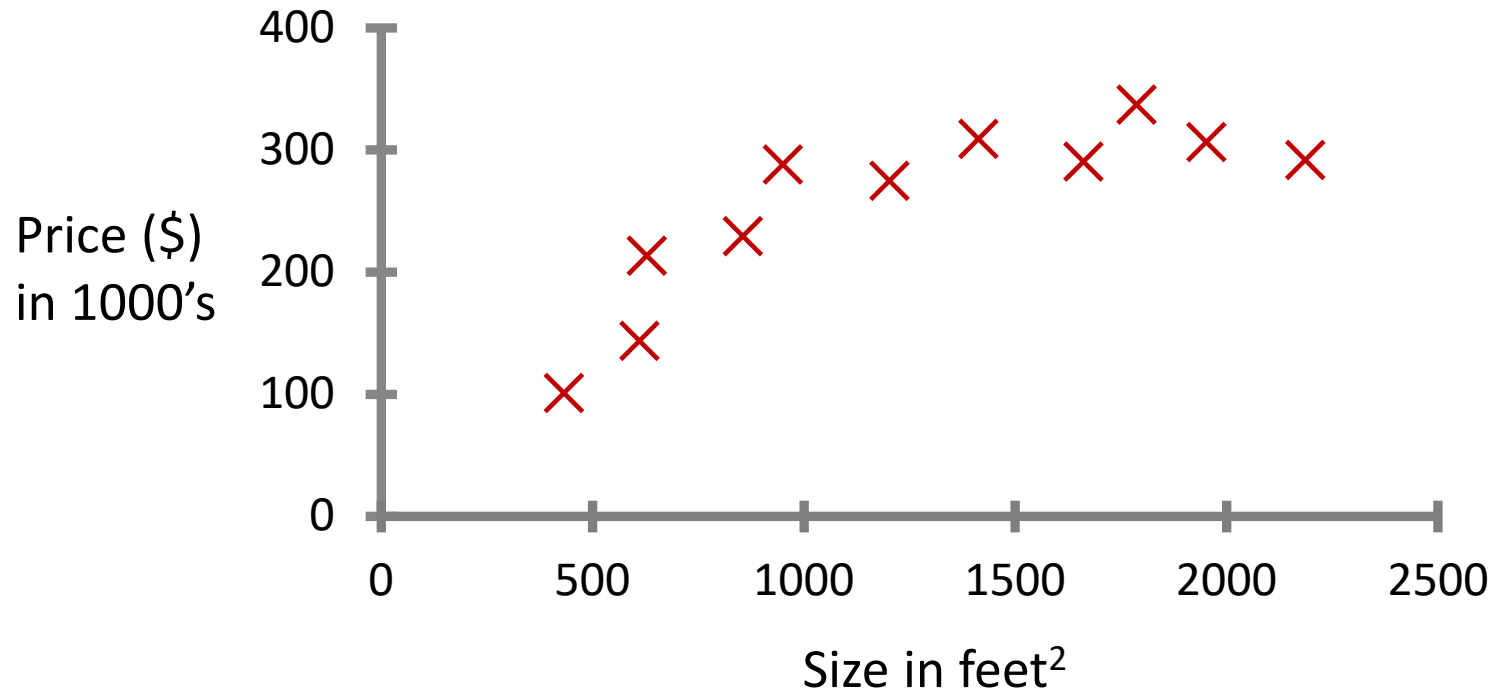
Machine Learning

# Introduction

---

# Supervised Learning

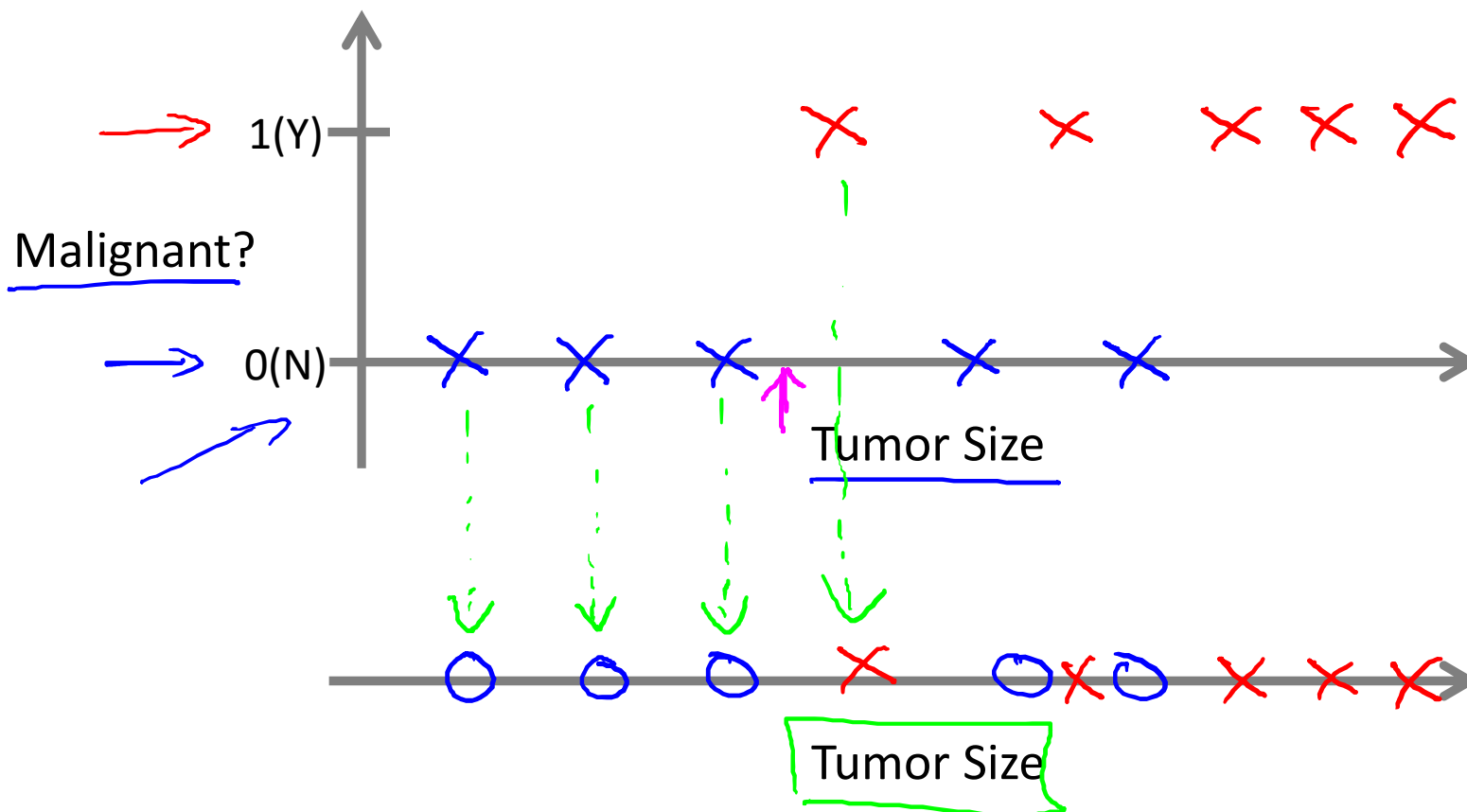
# Housing price prediction.



Supervised Learning  
“right answers” given

Regression: Predict continuous  
valued output (price)

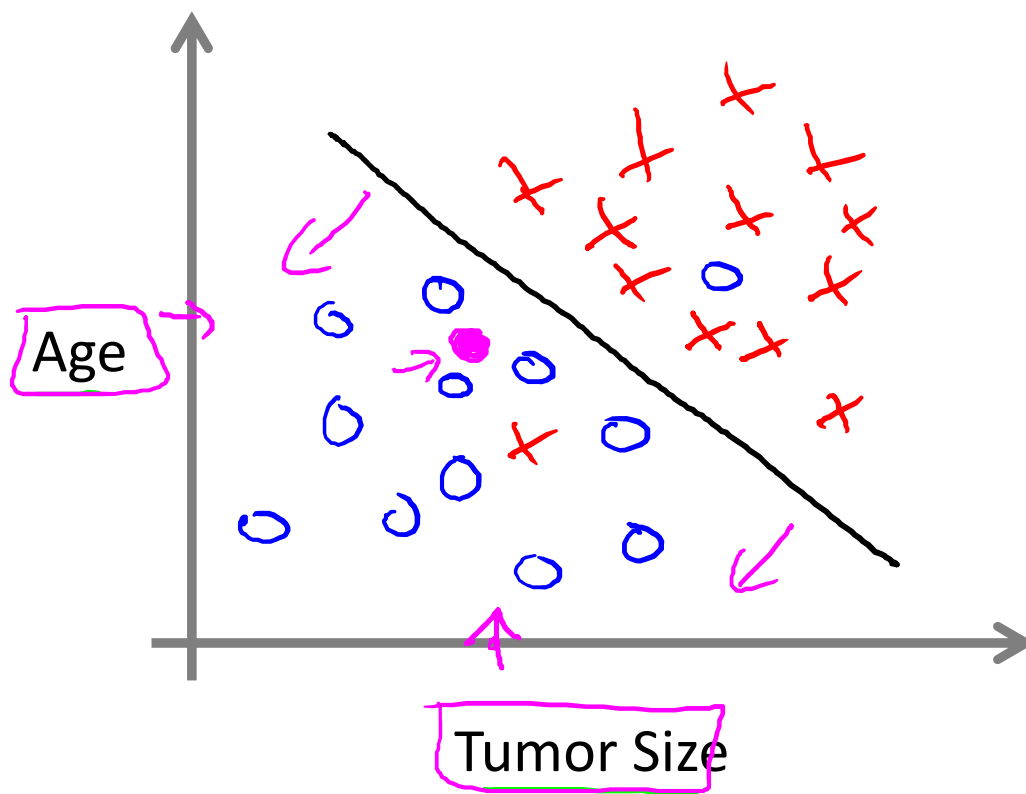
# Breast cancer (malignant, benign)



## Classification

Discrete valued  
output (0 or 1)

0, 1, 2, 3  
↓  
benign type 1  
cancer



- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape
- ...

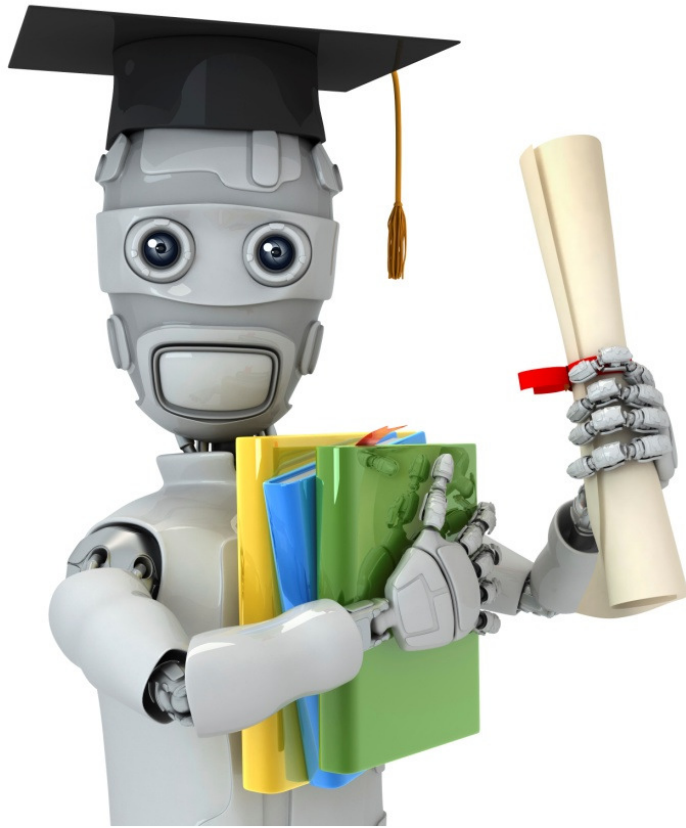
You're running a company, and you want to develop learning algorithms to address each of two problems.

Problem 1: You have a large inventory of identical items. You want to predict how many of these items will sell over the next 3 months.

Problem 2: You'd like software to examine individual customer accounts, and for each account decide if it has been hacked/compromised.

Should you treat these as classification or as regression problems?

- ☐ Treat both as classification problems.
- ☐ Treat problem 1 as a classification problem, problem 2 as a regression problem.
- ☐ Treat problem 1 as a regression problem, problem 2 as a classification problem.
- ☐ Treat both as regression problems.



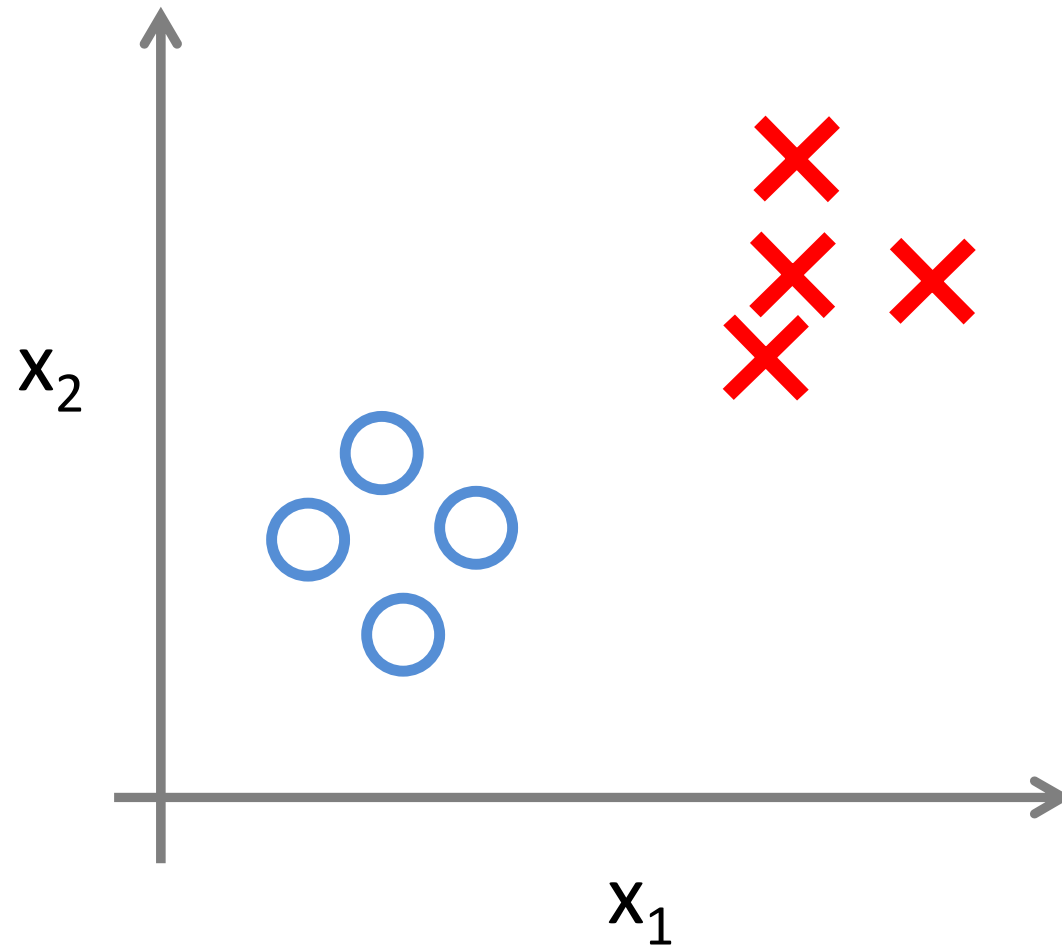
Machine Learning

# Introduction

---

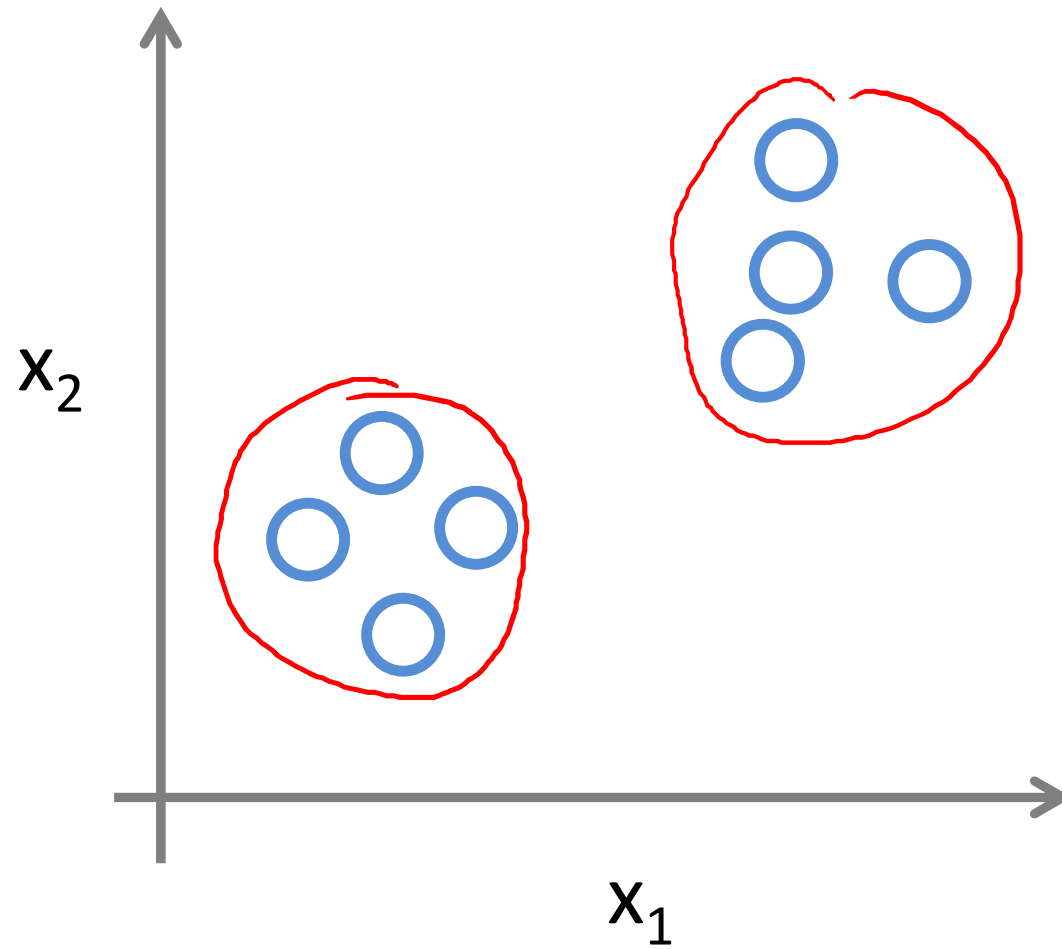
# Unsupervised Learning

# Supervised Learning



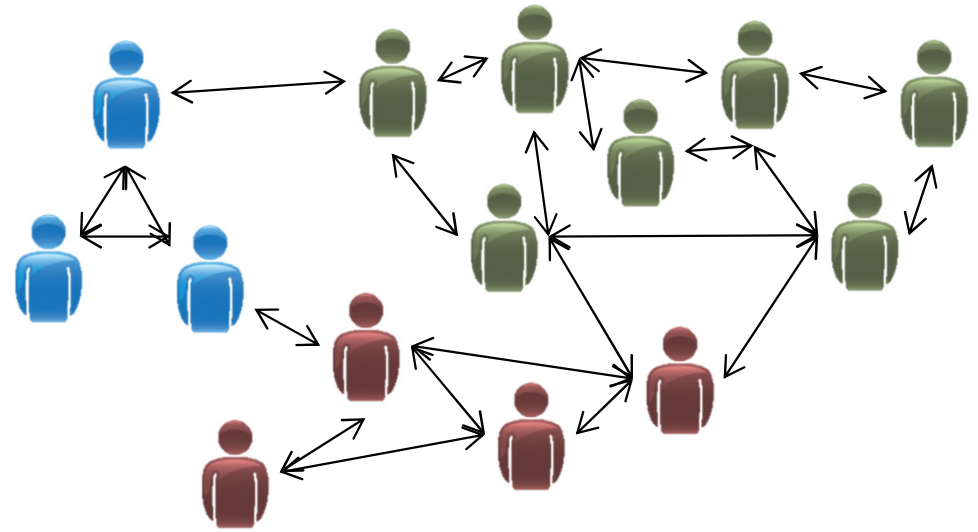


# Unsupervised Learning

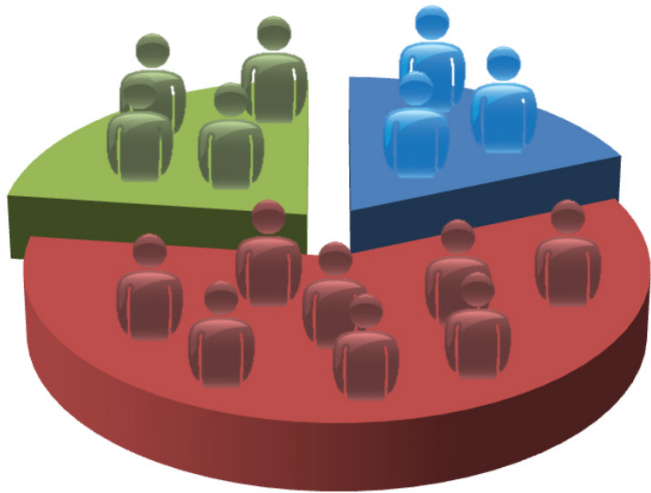




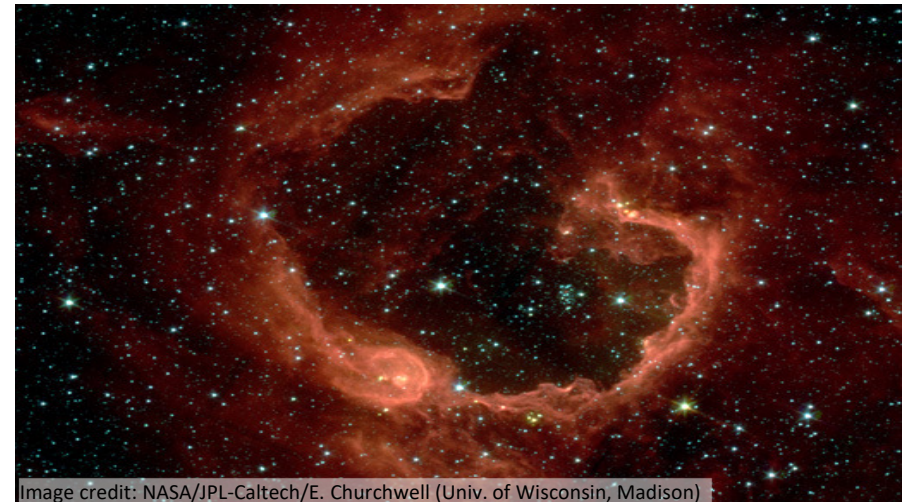
Organize computing clusters



Social network analysis



Market segmentation



Astronomical data analysis

Of the following examples, which would you address using an unsupervised learning algorithm? (Check all that apply.)

- ☐ Given email labeled as spam/not spam, learn a spam filter.
- ☐ Given a set of news articles found on the web, group them into set of articles about the same story.
- ☐ Given a database of customer data, automatically discover market segments and group customers into different market segments.
- ☐ Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.