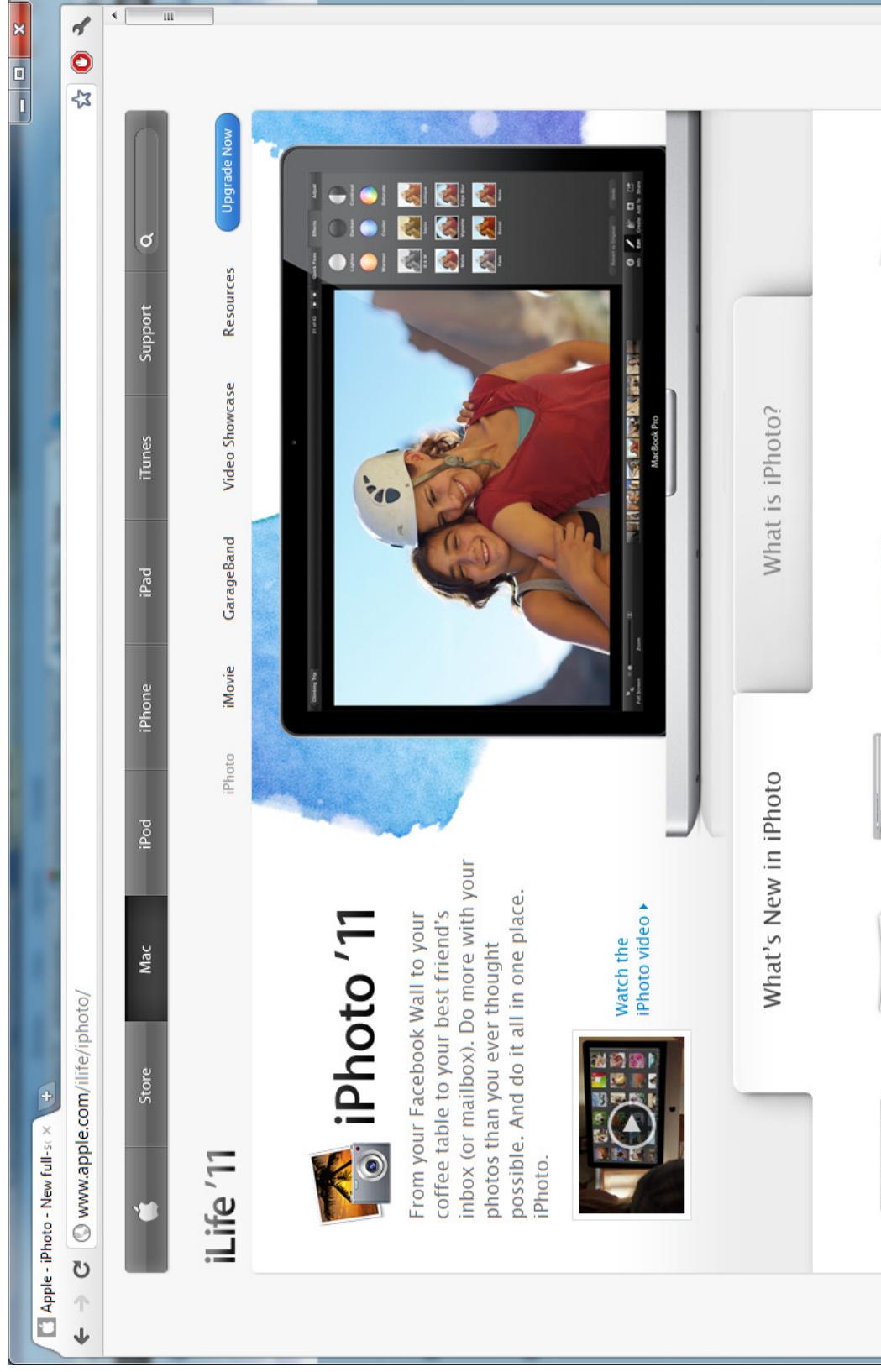


Machine Learning

Introduction

Welcome





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

- |

Exam

- |

- |



ig

lost of

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

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.

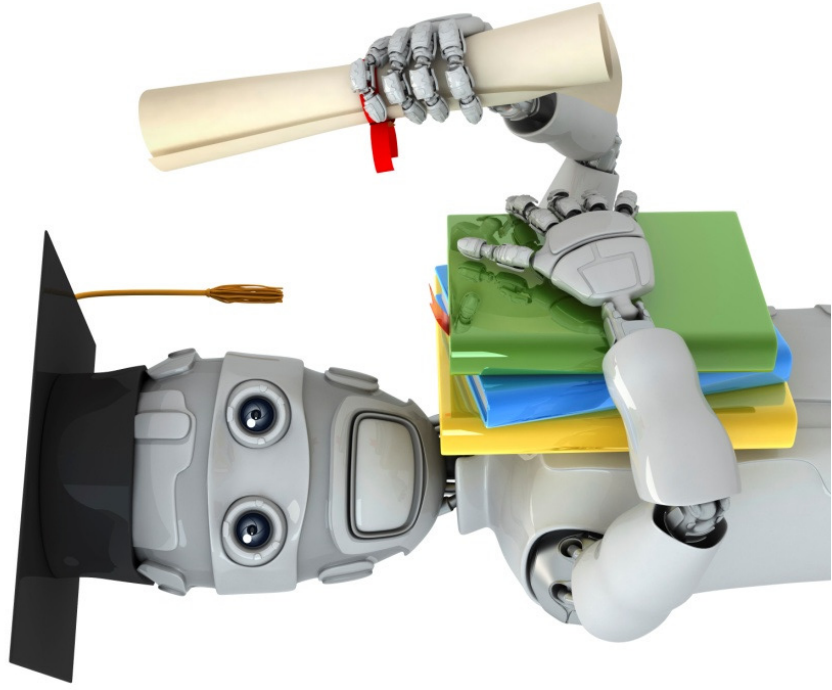
E.g., Autonomous helicopter, handwriting recognition, most of

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

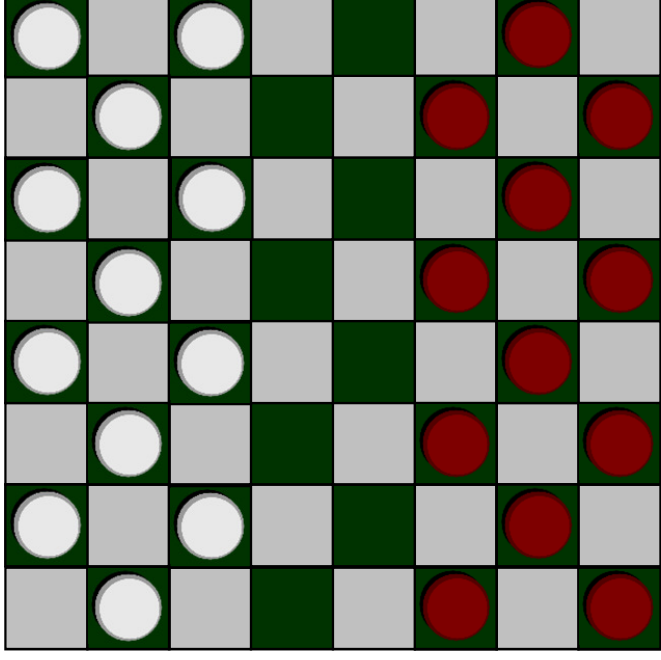
Machine Learning definition

Machine Learning definition

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

Machine Learning definition

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



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?

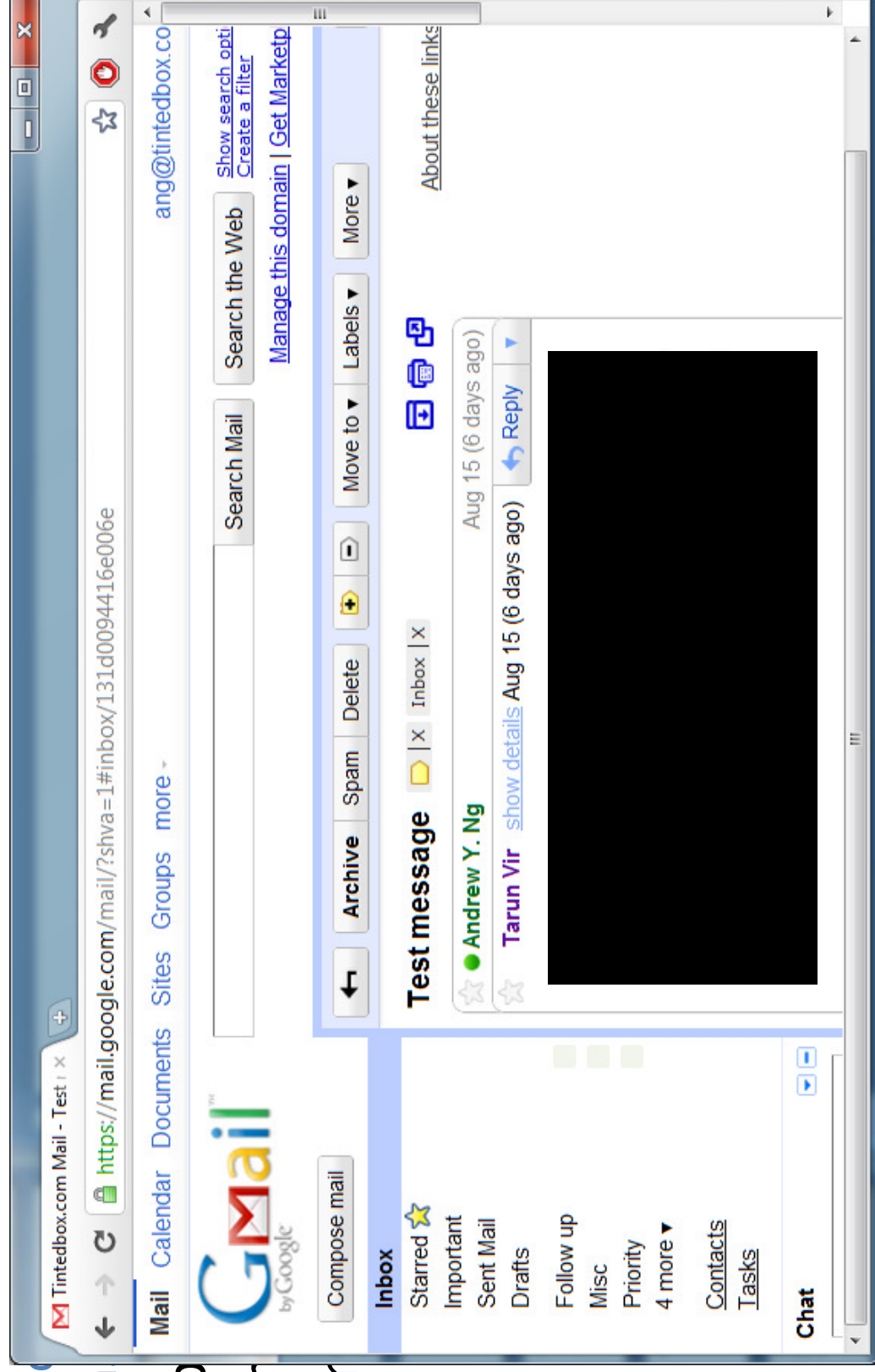
- ☒ Classifying emails as spam or not spam. T ←
- ☐ Watching you label emails as spam or not spam. E ←
- ☐ The number (or fraction) of emails correctly classified as spam/not spam.
- ☐ None of the above—this is not a machine learning problem.

✓ P ←

“A computer program is said to *learn* from experience E with respect to
on T,

Support not spam or do filter

pam.




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

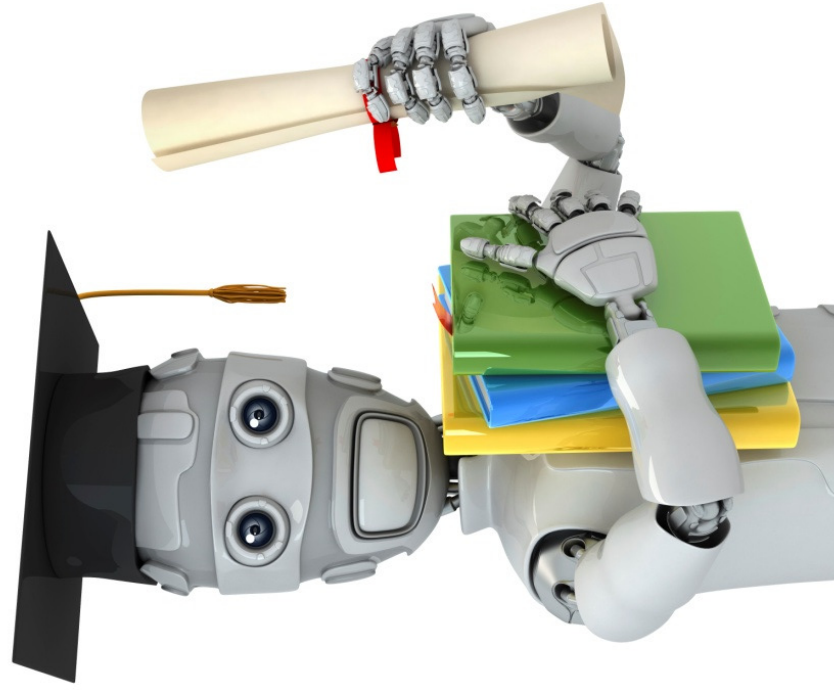
✓ P ←

Machine learning algorithms:

- Supervised learning
 - Unsupervised learning
- 

Others: Reinforcement learning, recommender systems.

Also talk about: Practical advice for applying learning algorithms. 

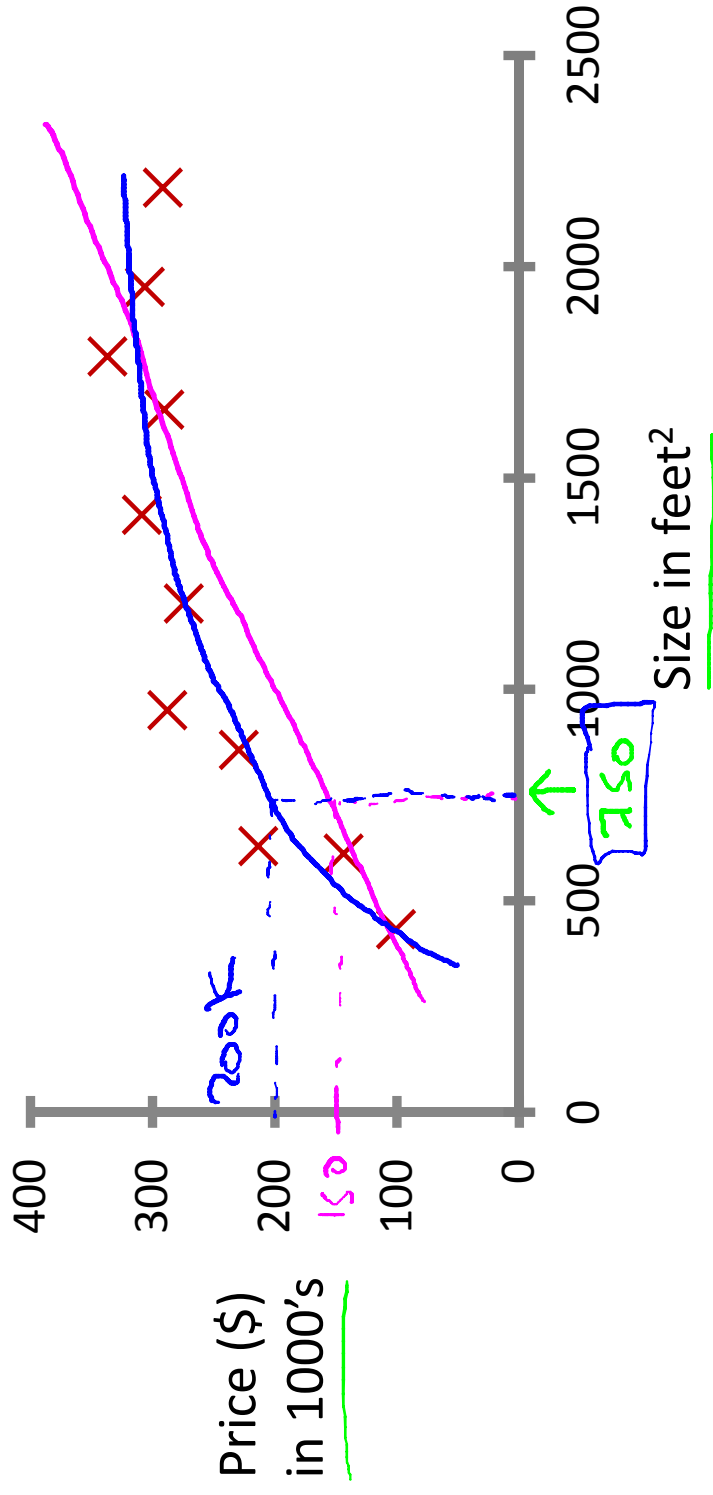


Machine Learning

Introduction

Supervised Learning

Housing price prediction.

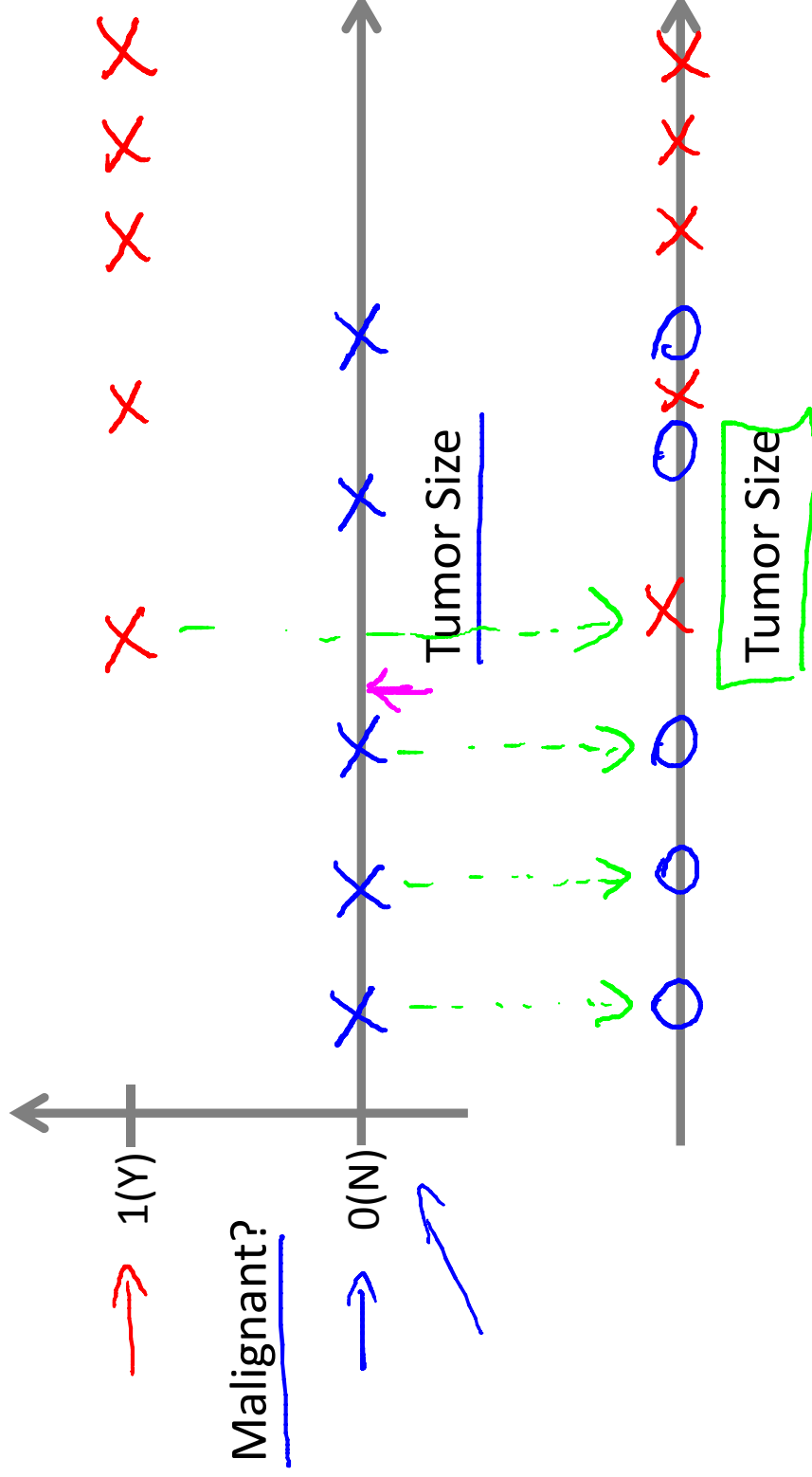


Supervised Learning

"right answers" given

Regression: Predict continuous
valued output (price)

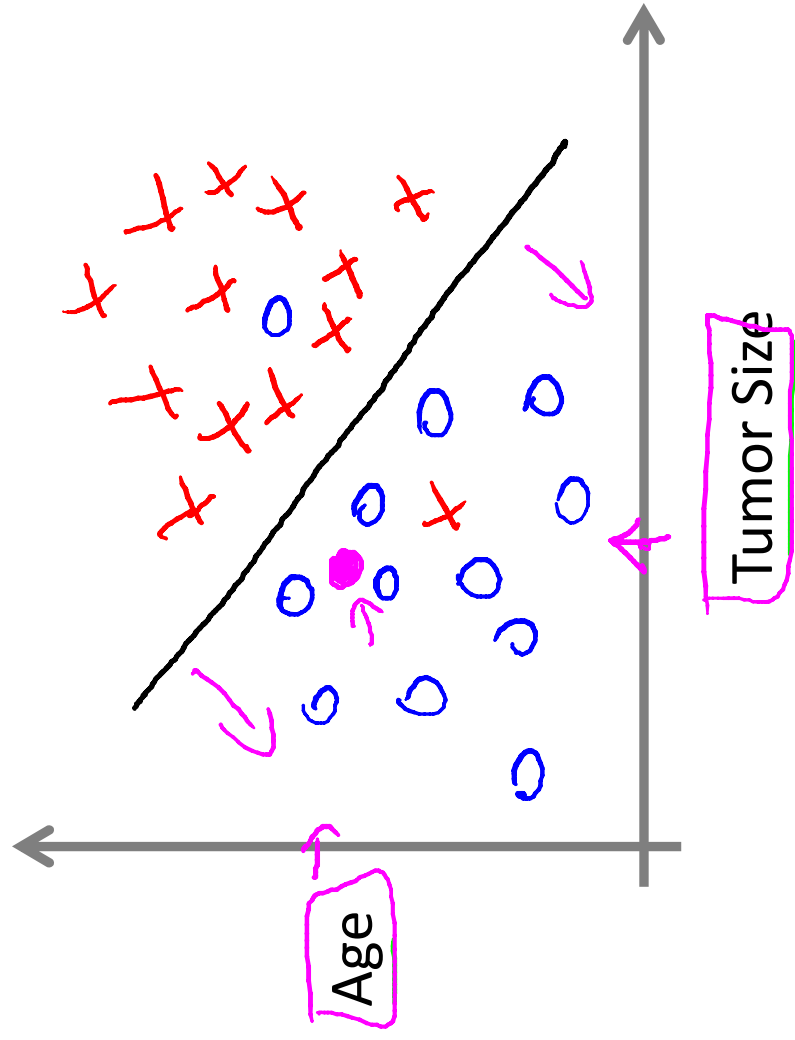
Breast cancer (malignant, benign)



Classification

Discrete valued
output (0 or 1)

0, 1, 2, 3
benign type 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.

1000's

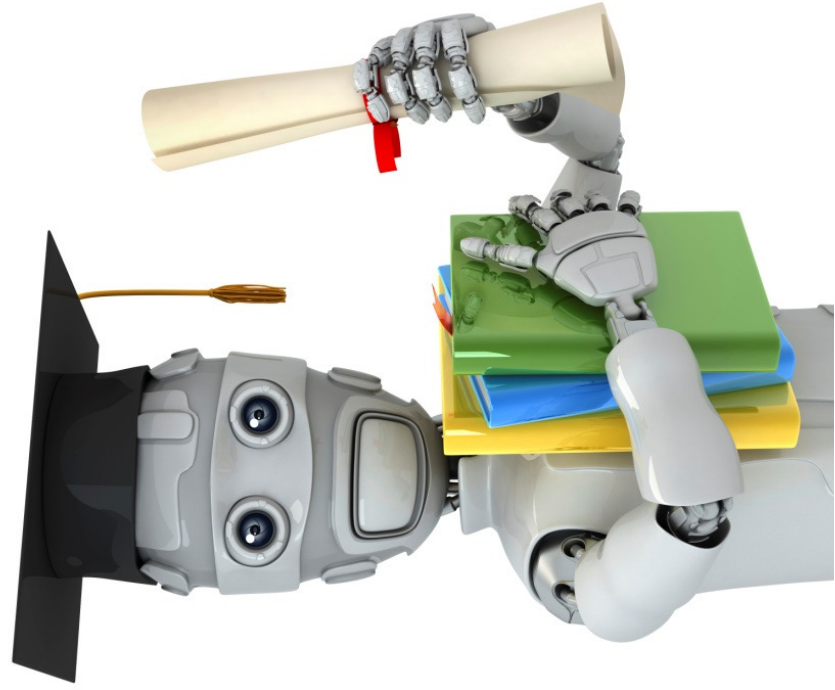
Problem 1: You have a large inventory of identical items. You want to predict how many of these items will sell over the next 3 months.

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

→ 0 - not hacked
→ 1 - hacked

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.

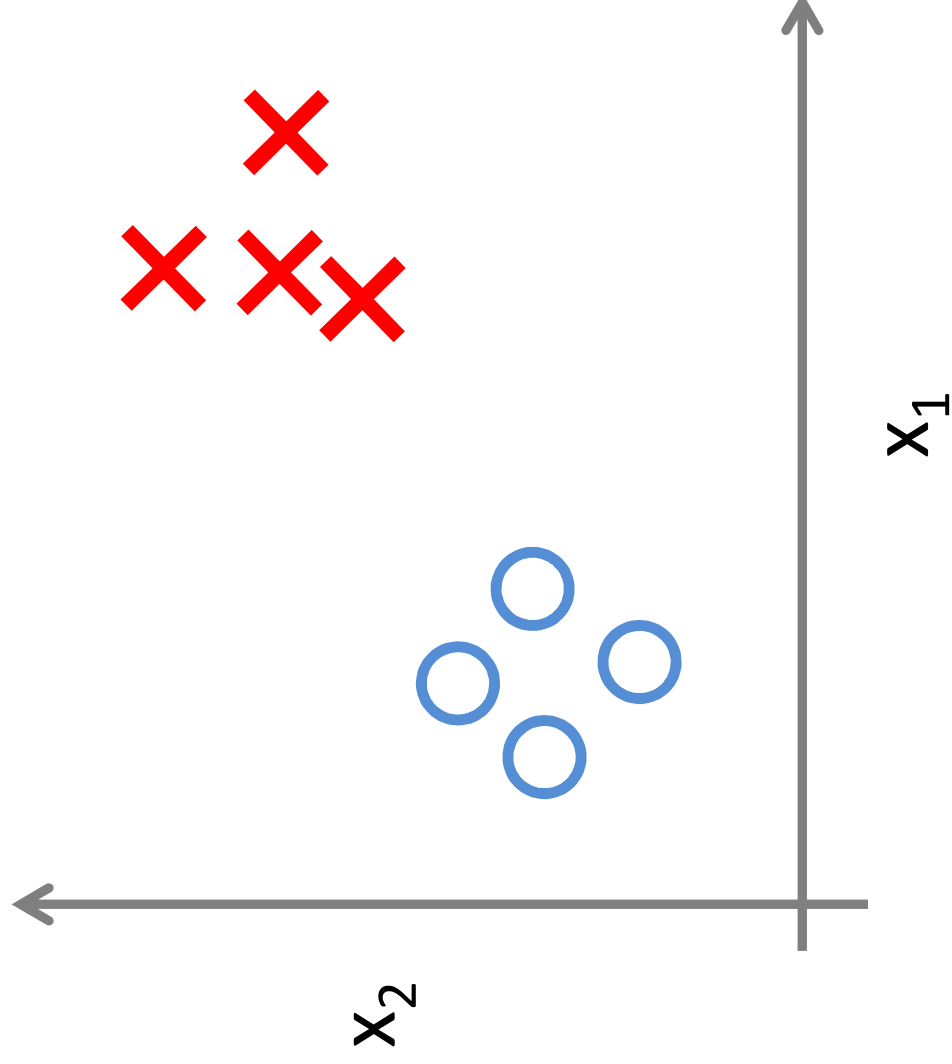


Machine Learning

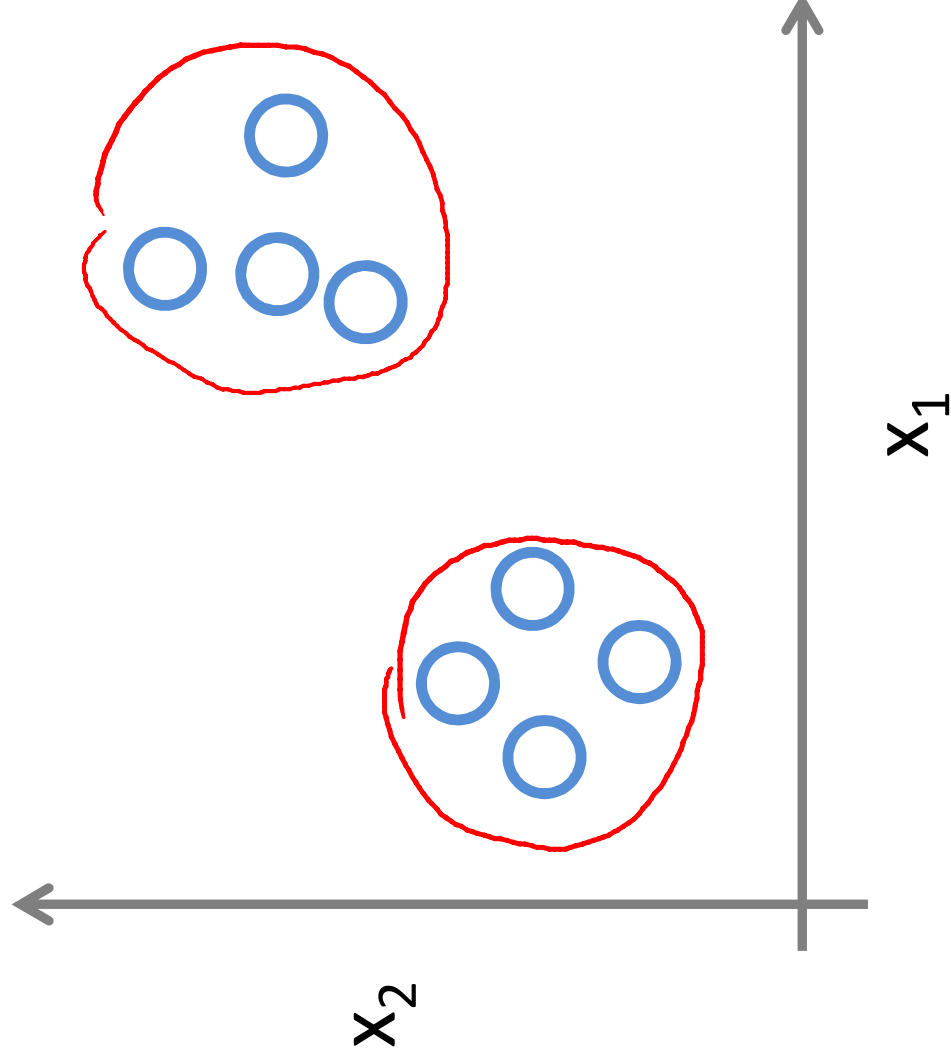
Introduction

Unsupervised Learning



Supervised Learning



Unsupervised Learning




[illegible]

 BP oil spill cost hits next... >  www.guardian.co.uk/environment/2010/sep/20/bp-oil-spill-dee

Mobile site Sign in Register A Text larger smaller About us ▼

guardian.co.uk


News Sport Comment Culture Business Money Life & style

Business 

BP oil spill cost hits nearly \$10bn

BP has set up a \$20bn compensation fund after the Deepwater Horizon disaster, which has so far paid out 19,000 claims totalling more than \$240m

Julia Kollewe
guardian.co.uk Monday, 20 September 2010 08:33 BST
Africa history



SPEX coast for the Deepwater Horizon disaster have hit \$10bn. Photograph: Ho'Reuters

[illegible]

Allen: Well is dead, but m... >

EDITOR: INTERNATIONAL | U.S. | MEXICO | ARABIC

Set edition preference

Home Video World U.S. Africa Asia Europe Latin America Middle East Business

CNN

Allen: Well is dead, but much Gulf Coast work remains

By the CNN Wire Staff
September 20, 2010 — Updated 1317 GMT (2:17 HKT)

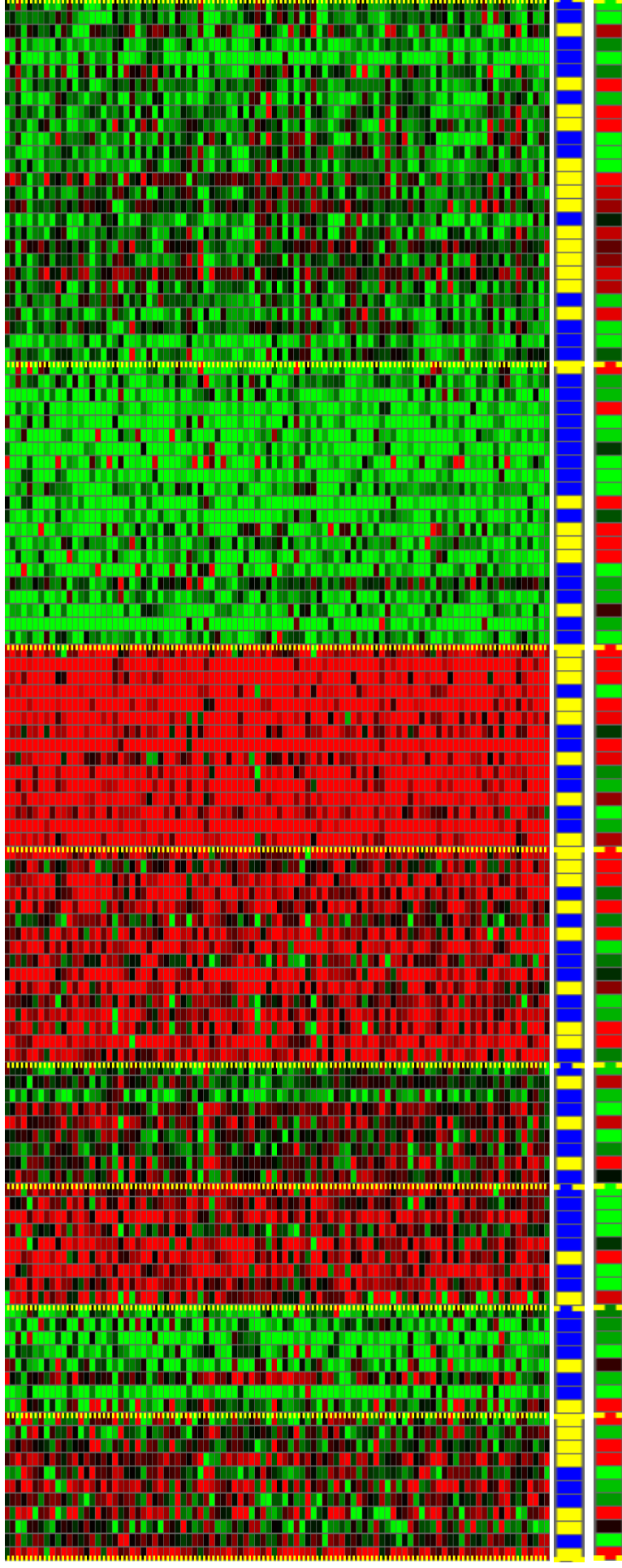
Click to play

What most for Gulf oil spill?

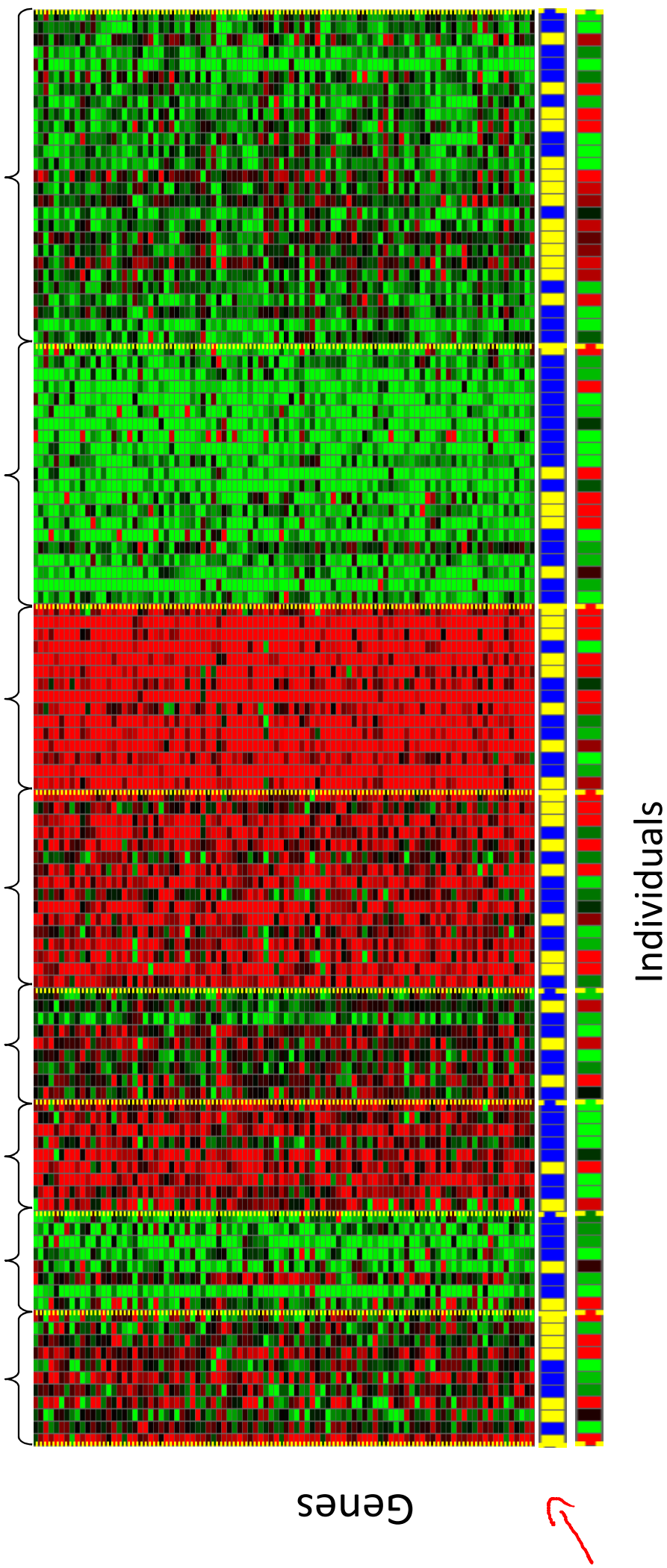
STORY HIGHLIGHTS

(CNN) — The ruptured Macondo well, a mile under the Gulf of

Genes



Individuals

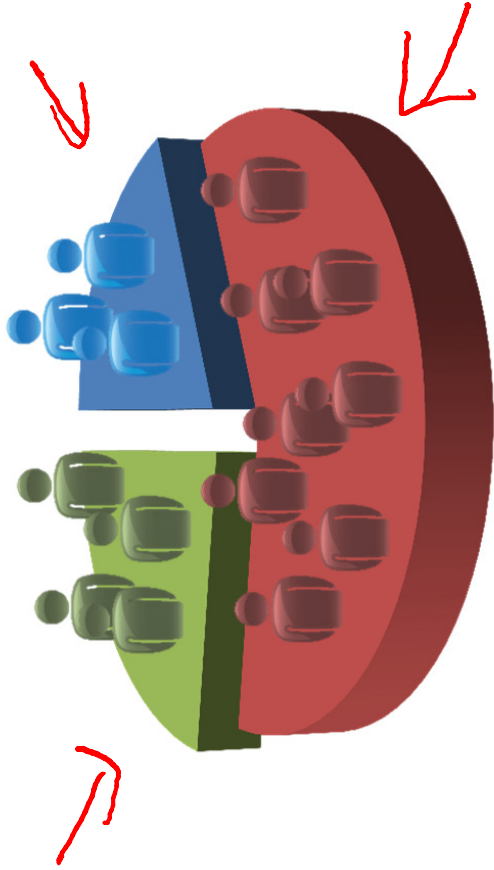


[Source: Daphne Koller]

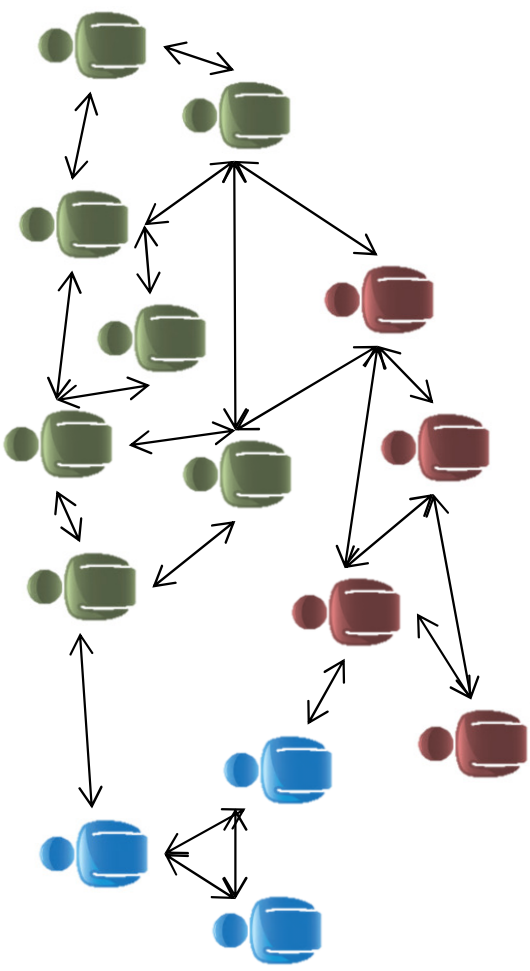
Andrew Ng



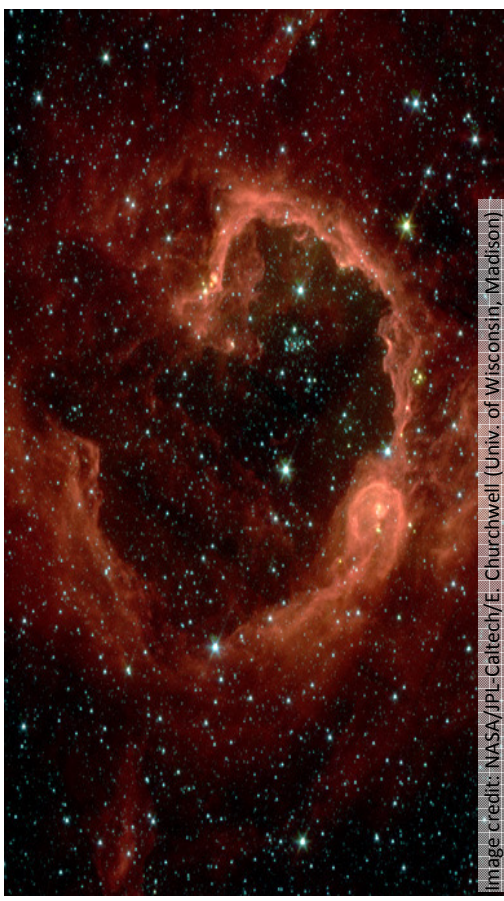
Organize computing clusters



Market segmentation

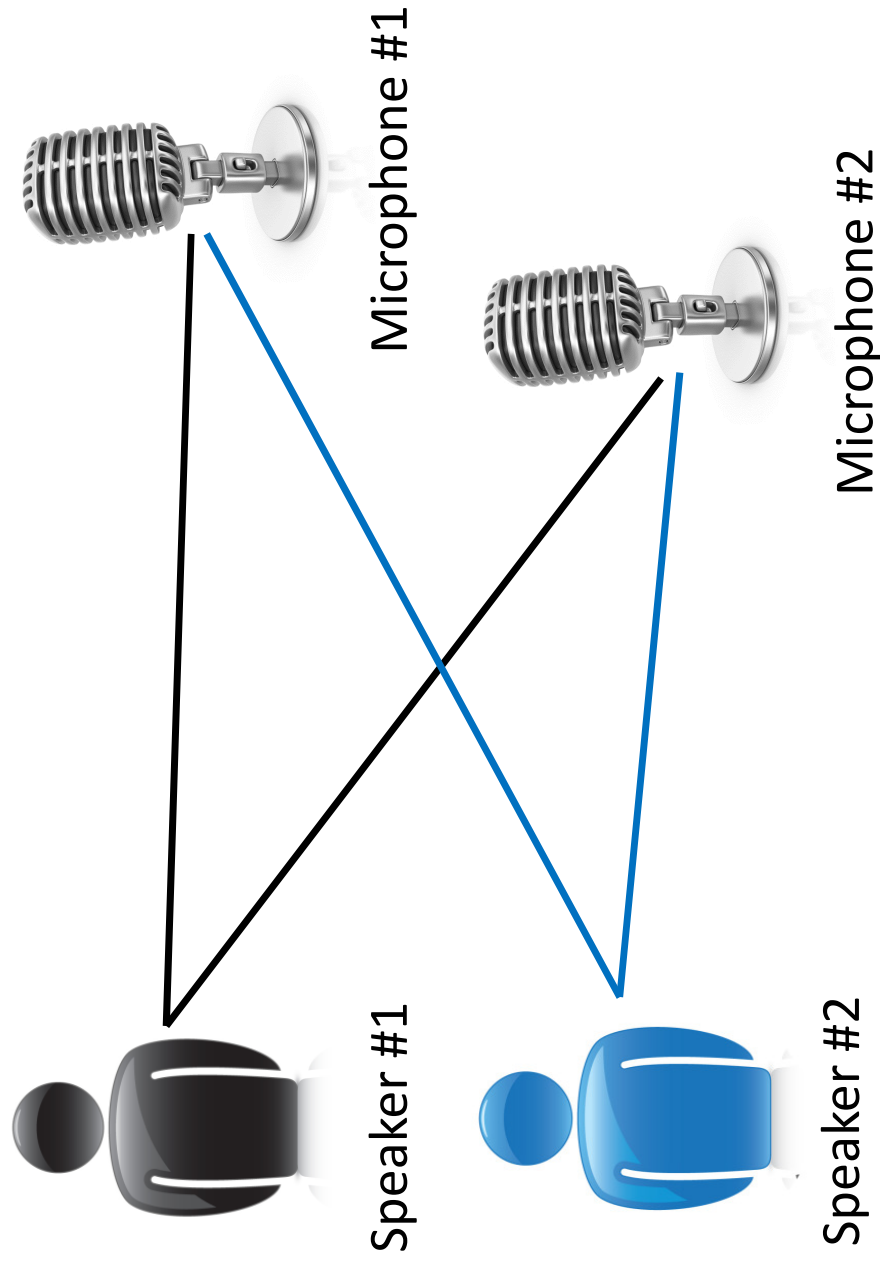


Social network analysis



Astronomical data analysis

Cocktail party problem



Microphone #1: 💡

Output #1: 💡

Microphone #2: 💡

Output #2: 💡

Microphone #1: 💡

Output #1: 💡

Microphone #2: 💡

Output #2: 💡

[Audio clips courtesy of Te-Won Lee.]

Cocktail party problem algorithm

$$[W,s,v] = \text{svd}(\text{repmat}(\text{sum}(x.^*x,1),\text{size}(x,1),1).^*x).^*x');$$

[Source: Sam Roweis, Yair Weiss & Eero Simoncelli]

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.

