Introduction to Machine Learning

What is Machine Learning?

"Learning is any process by which a system improves performance from experience."

- Herbert Simon

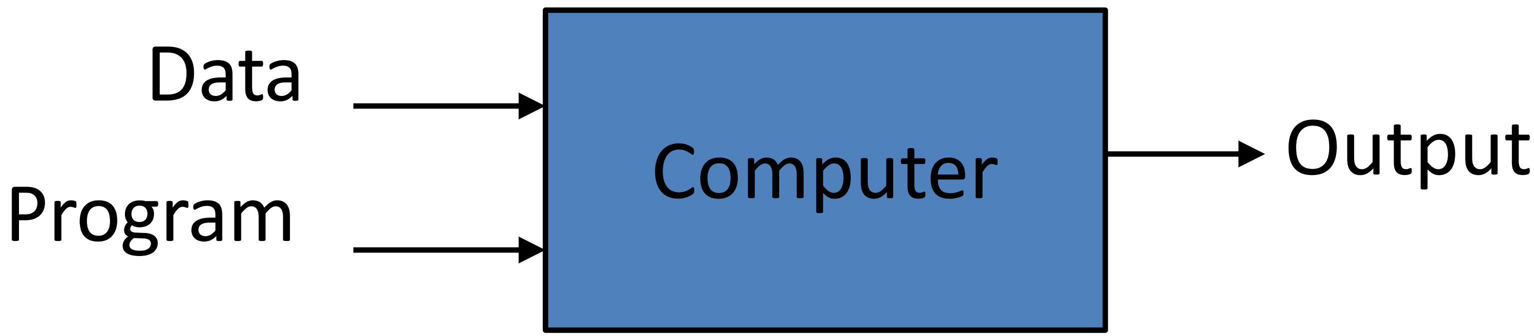
Definition by Tom Mitchell (1998):

Machine Learning is the study of algorithms that

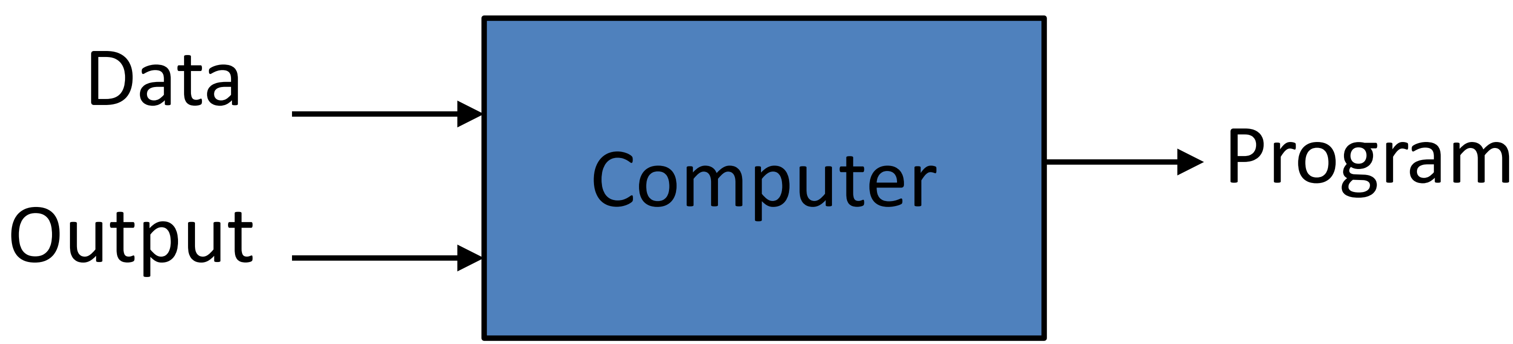
- improve their performance P
- at some task T
- with experience E.

A well-defined learning task is given by <P, T, E>.

Traditional Programming



Machine Learning

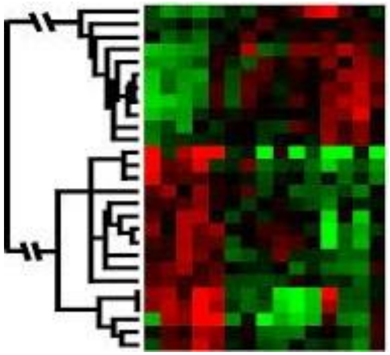


When Do We Use Machine Learning?

ML is used when:

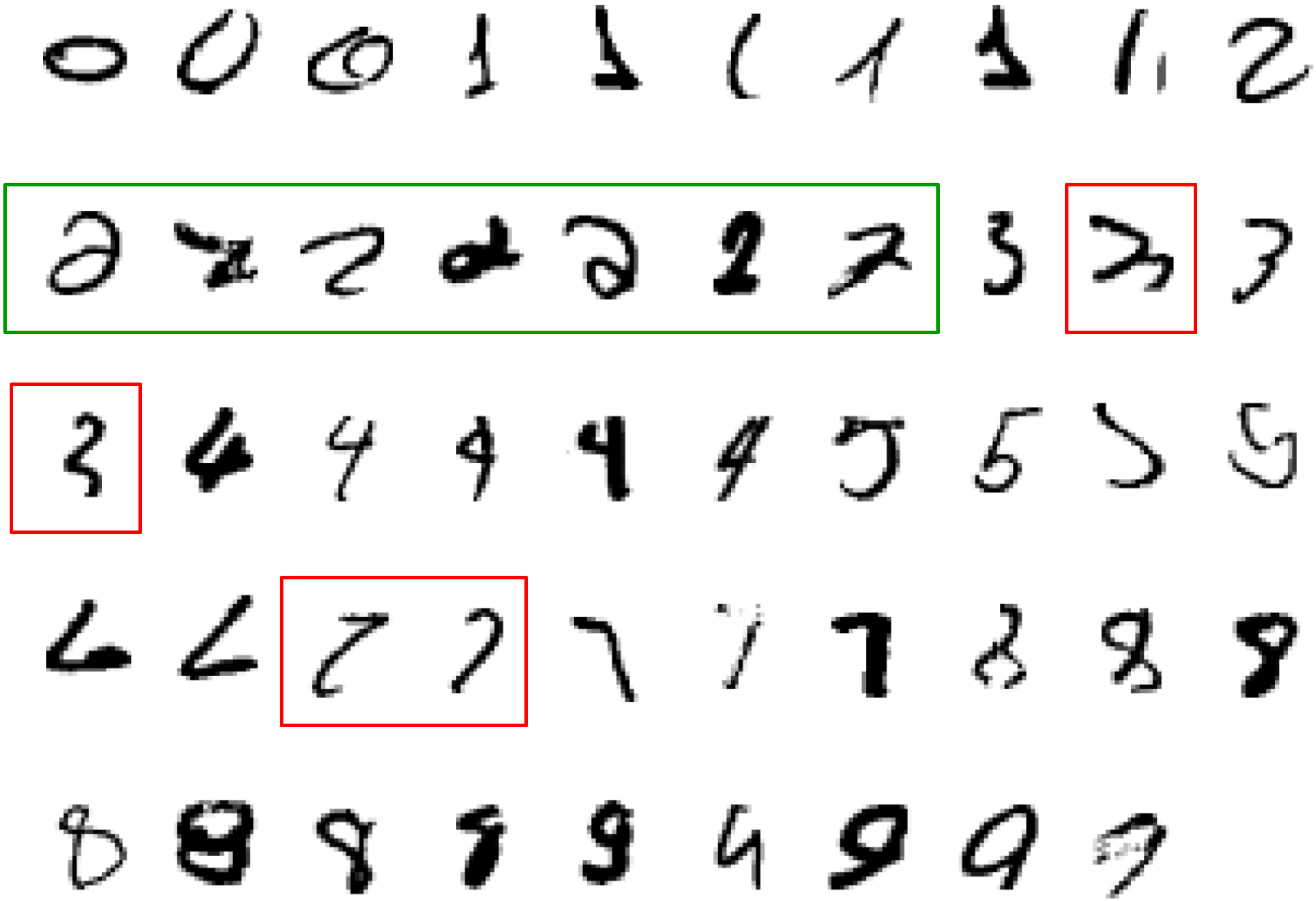
- Human expertise does not exist (navigating on Mars)
- Humans can't explain their expertise (speech recognition)
- Models must be customized (personalized medicine)
- Models are based on huge amounts of data (genomics)



A classic example of a task that requires machine learning:

It is very hard to say what makes a 2



Some more examples of tasks that are best solved by using a learning algorithm

- Recognizing patterns:
  - Facial identities or facial expressions
  - Handwritten or spoken words
  - Medical images
- Generating patterns:
  - Generating images or motion sequences
- Recognizing anomalies:
  - Unusual credit card transactions
  - Unusual patterns of sensor readings in a nuclear power plant
- Prediction:
  - Future stock prices or currency exchange rates

Sample Applications

- Web search
- Cybersecurity
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information Extraction
- Social networks
- Debugging software

Samuel's Checkers-Player

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

Defining the Learning Task

Improve on task T, with respect to performance metric P, based on experience E

T: Playing checkers
P: Percentage of games won against an arbitrary opponent
E: Playing practice games against itself

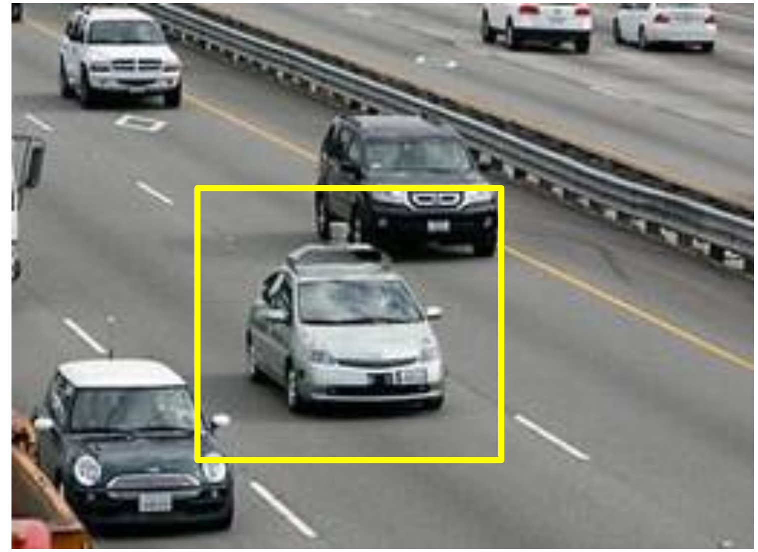
T: Recognizing hand-written words
P: Percentage of words correctly classified
E: Database of human-labeled images of handwritten words

T: Driving on four-lane highways using vision sensors
P: Average distance traveled before a human-judged error
E: A sequence of images and steering commands recorded while observing a human driver.

T: Categorize email messages as spam or legitimate.
P: Percentage of email messages correctly classified.
E: Database of emails, some with human-given labels

State of the Art Applications of Machine Learning

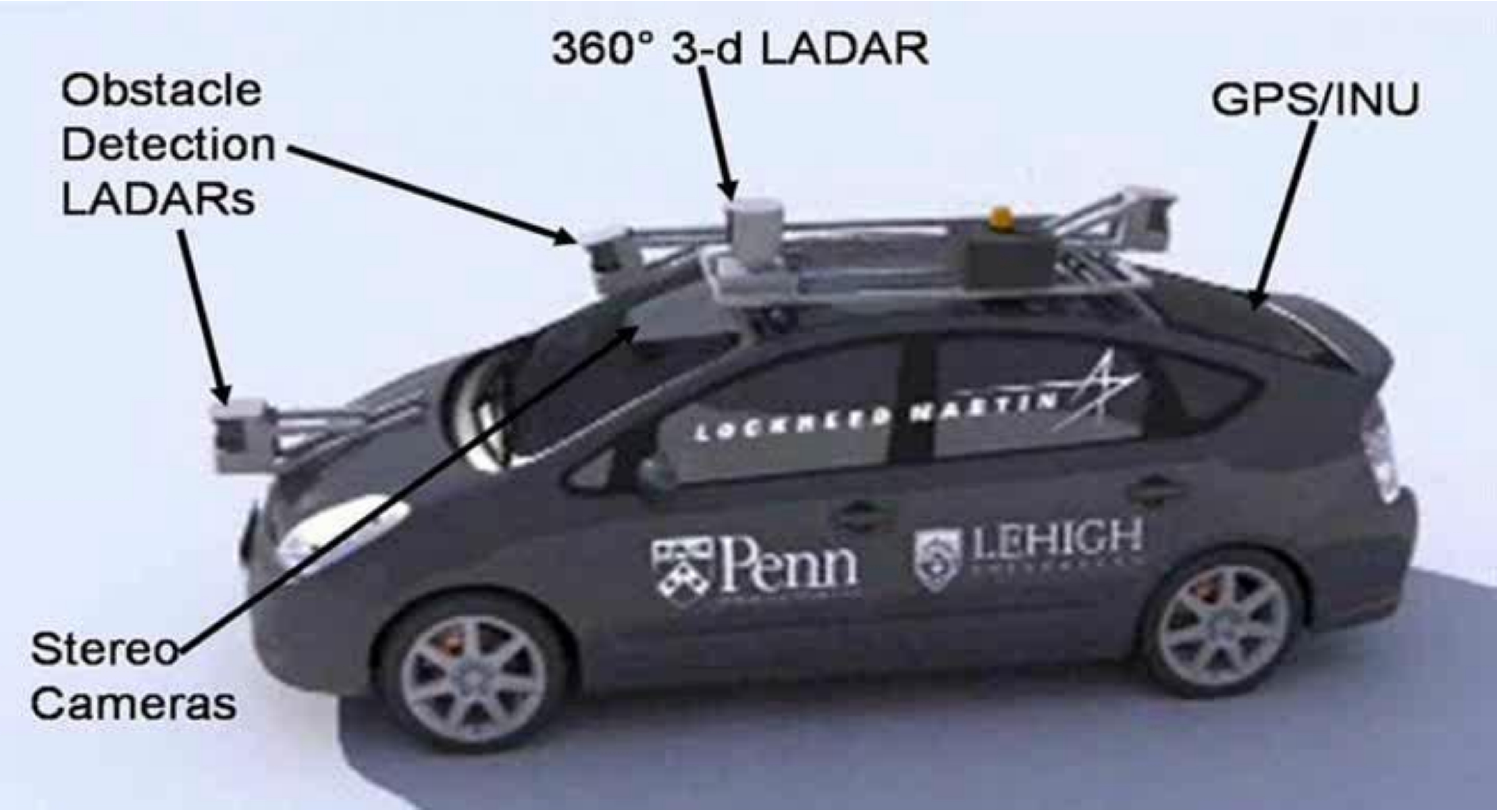
Autonomous Cars



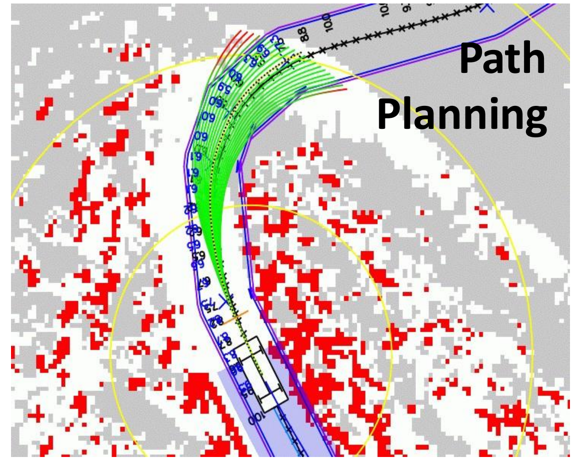
- Nevada made it legal for autonomous cars to drive on roads in June 2011
- As of 2013, four states (Nevada, Florida, California, and Michigan) have legalized autonomous cars

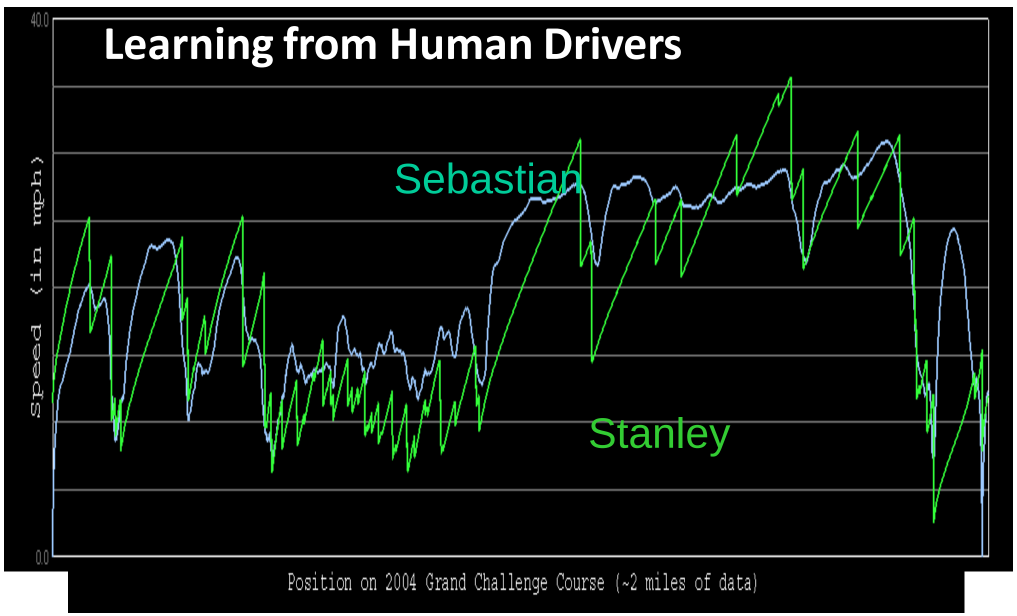
Penn's Autonomous Car →
(Ben Franklin Racing Team)

Autonomous Car Sensors



Autonomous Car Technology





Images and movies taken from Sebastian Thrun's multimedia website.

Deep Learning in the Headlines

BUSINESS NEWS

MIT Technology Review

Is Google Cornering the Market on Deep Learning?

A cutting-edge corner of science is being wooed by Silicon Valley, to the dismay of some academics.

By Antonio Regalado on January 29, 2014

How much are a dozen deep-learning researchers worth? Apparently, more than $400 million.

This week, Google reportedly paid that much to acquire DeepMind Technologies, a startup based in

BloombergBusinessweek Technology

Acquisitions

The Race to Buy the Human Brains Behind Deep Learning Machines

By Ashlee Vance | January 27, 2014

intelligence projects. "DeepMind is bona fide in terms of its research capabilities and depth," says Peter Lee, who heads Microsoft Research.

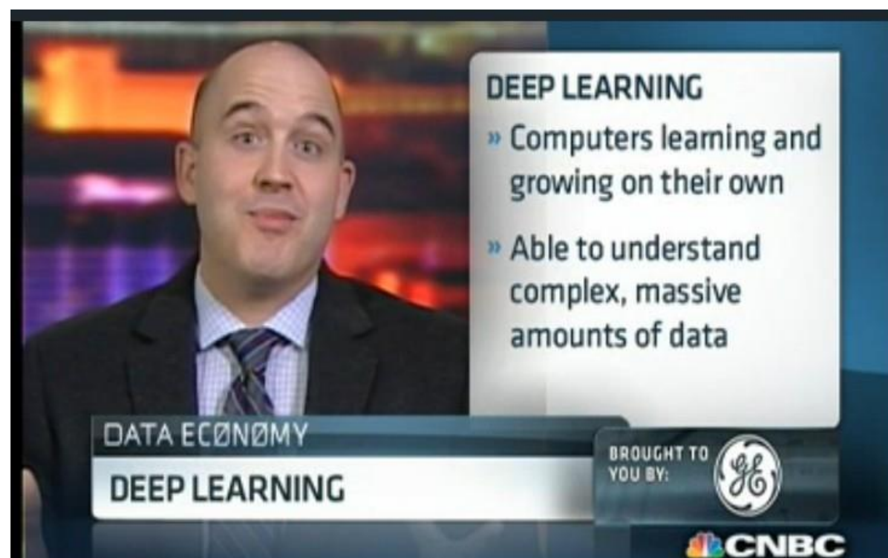
According to Lee, Microsoft, Facebook (FB), and Google find themselves in a battle for deep learning talent. Microsoft has gone from four full-time deep learning experts to 70 in the past three years. "We would have more if the talent was there to

WIRED GEAR SCIENCE ENTERTAINMENT BUSINESS SECURITY DESIG

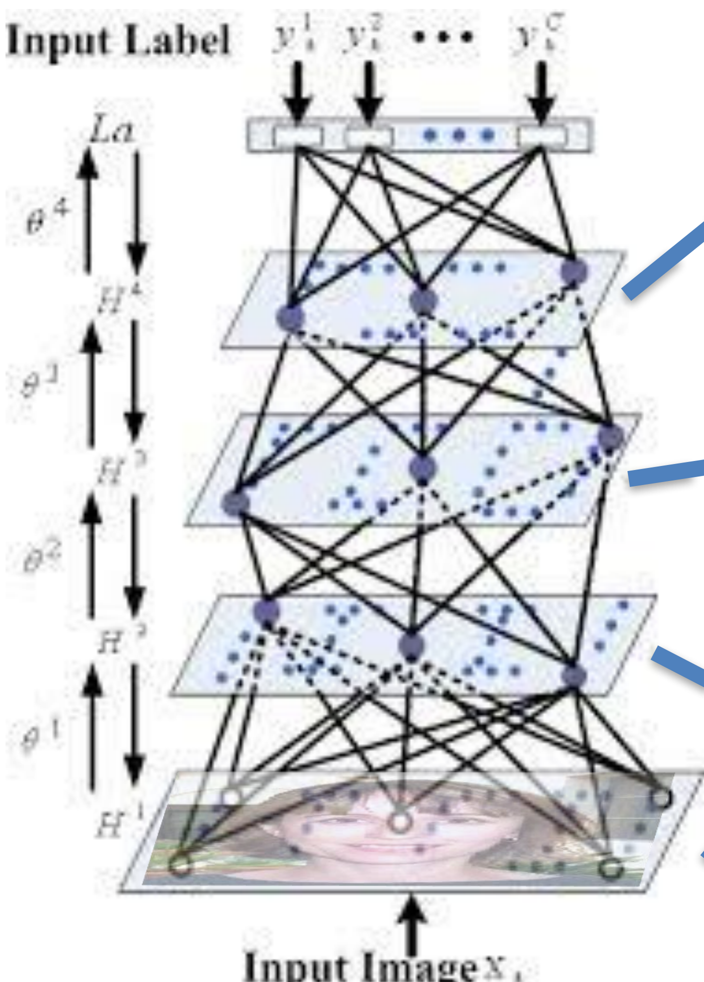
INNOVATION INSIGHTS | community content | featured

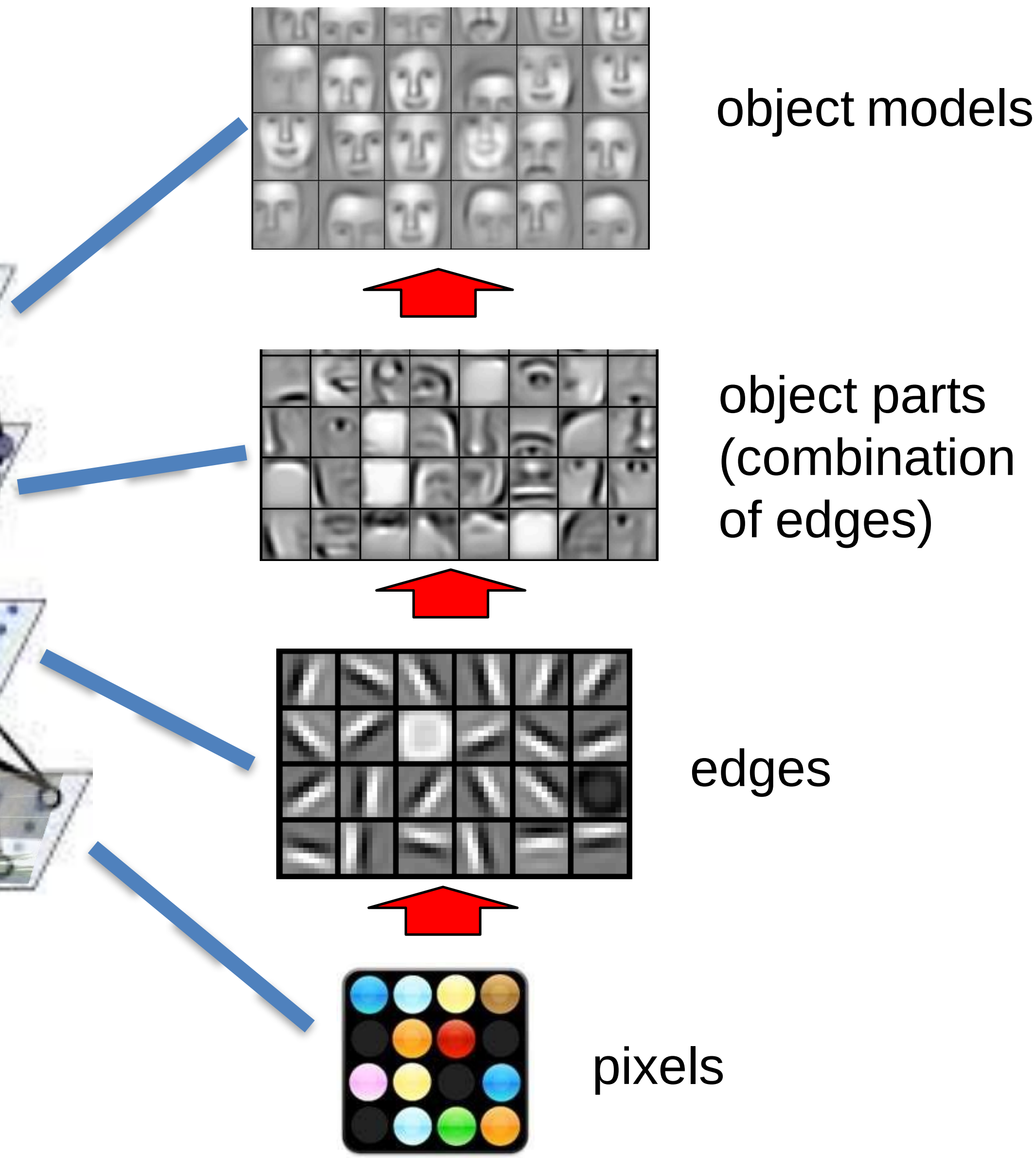
Deep Learning's Role in the Age of Robots

BY JULIAN GREEN, JETPAC 05.02.14 2:56 PM



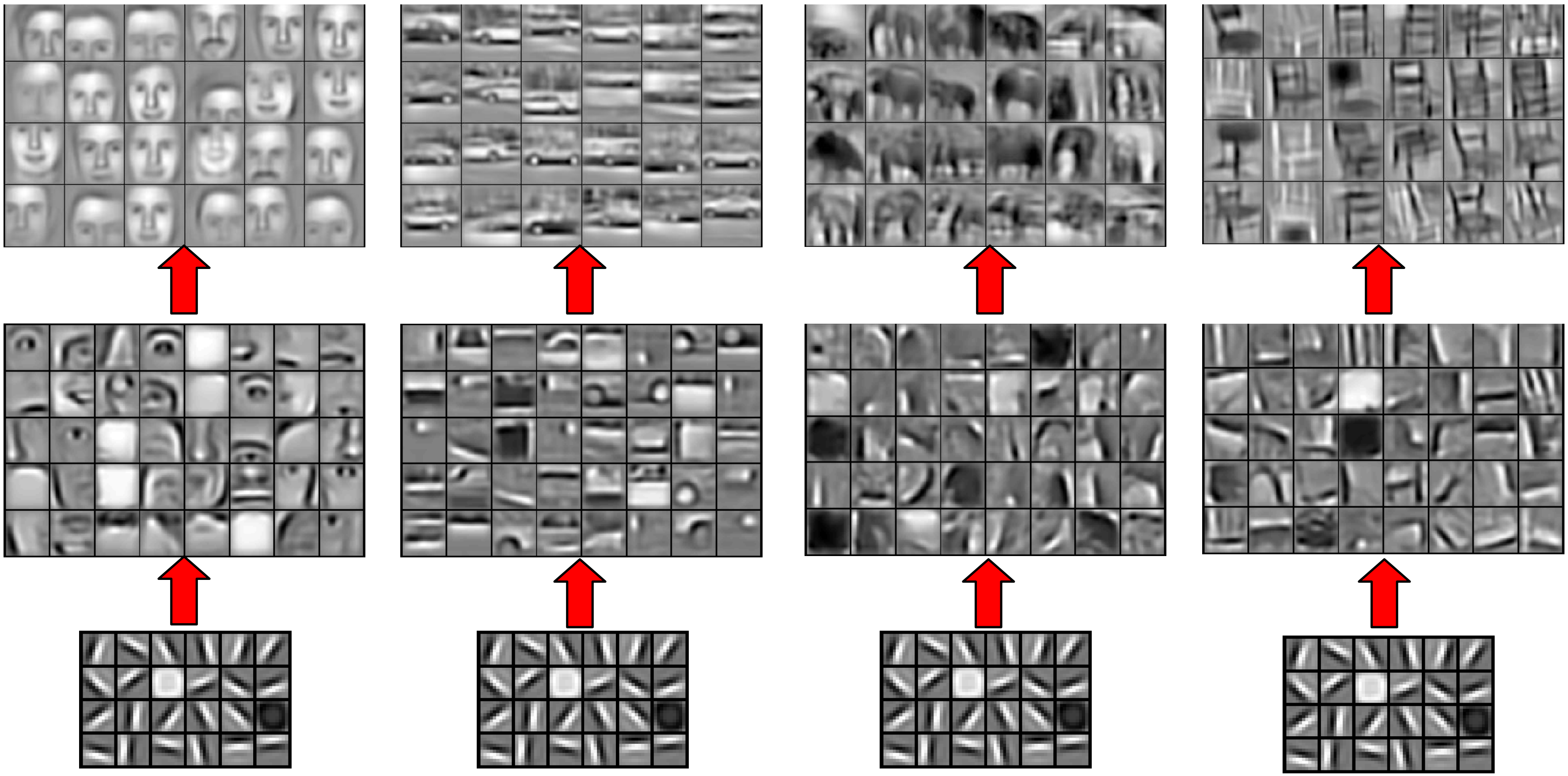
Deep Belief Net on Face Images





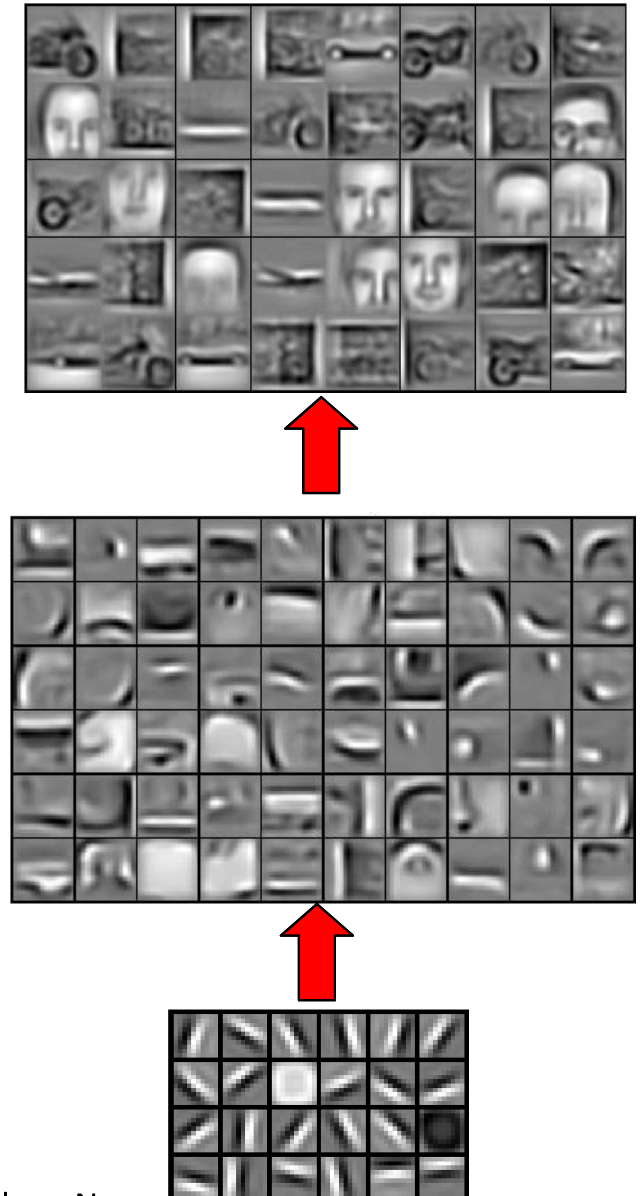
Based on materials
by Andrew Ng

Learning of Object Parts



Slide credit: Andrew Ng

Training on Multiple Objects



Trained on 4 classes (cars, faces, motorbikes, airplanes).

Second layer: Shared-features and object-specific features.

Third layer: More specific features.

Slide credit: Andrew Ng

Scene Labeling via Deep Learning

[Farabet et al. ICML 2012, PAMI 2013]

Inference from Deep Learned Models

Generating posterior samples from faces by “filling in” experiments
(cf. Lee and Mumford, 2003). Combine bottom-up and top-down inference.

Input images

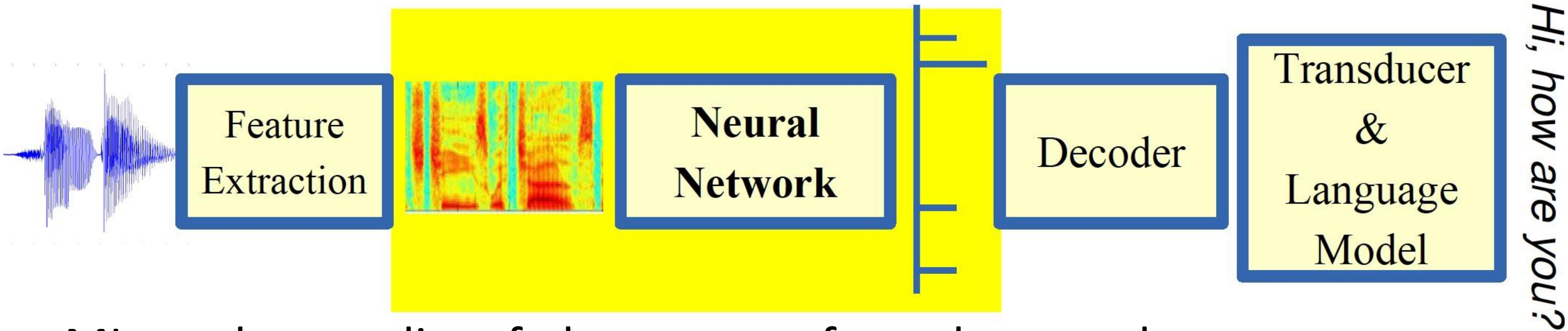
Samples from feedforward Inference (control)

Samples from Full posterior inference

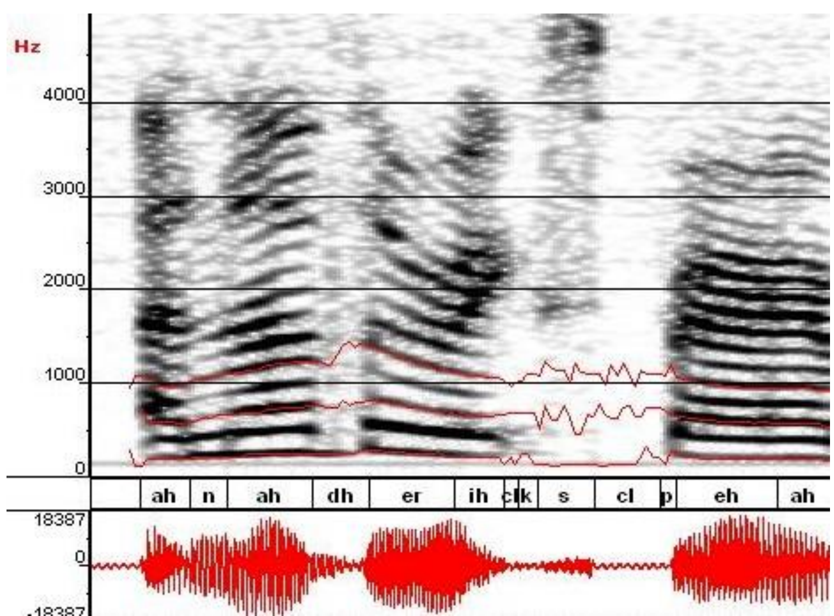
Slide credit: Andrew Ng

Machine Learning in Automatic Speech Recognition

A Typical Speech Recognition System



ML used to predict of phone states from the sound spectrogram



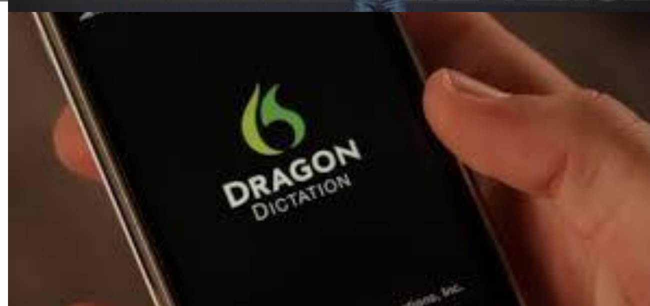
Deep learning has state-of-the-art results

# Hidden Layers	1	2	4	8	10	12
Word Error Rate %	16.0	12.8	11.4	10.9	11.0	11.1

Baseline GMM performance = 15.4%

[Zeiler et al. “On rectified linear units for speech recognition” ICASSP 2013]

Impact of Deep Learning in Speech Technology



Slide credit: Li Deng, MS Research

Types of Learning

Types of Learning

- **Supervised (inductive) learning**
  - Given: training data + desired outputs (labels)
- **Unsupervised learning**
  - Given: training data (without desired outputs)
- **Semi-supervised learning**
  - Given: training data + a few desired outputs
- **Reinforcement learning**
  - Rewards from sequence of actions

Tasks

Supervised Learning

x → Classification → y Discrete

x → Regression → y Continuous

Unsupervised Learning

x → Clustering → y Discrete ID

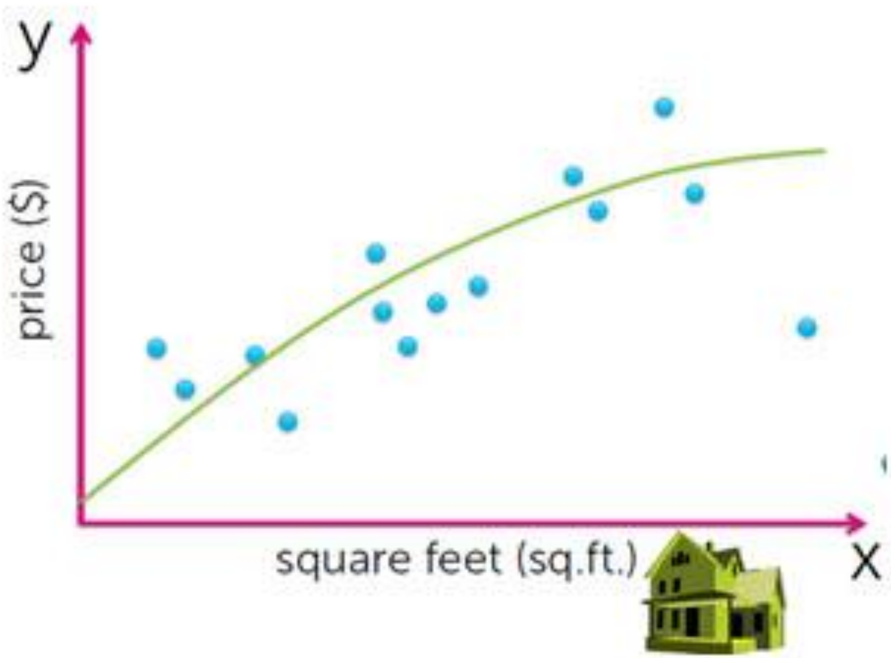
x → Dimensionality Reduction → y Continuous

Supervised Learning

Regression

x → Regression → y Continuous

House price prediction



Weather prediction

Temperature

27°C

(

Supervised Learning

Classification

Breast cancer (malignant, benign)

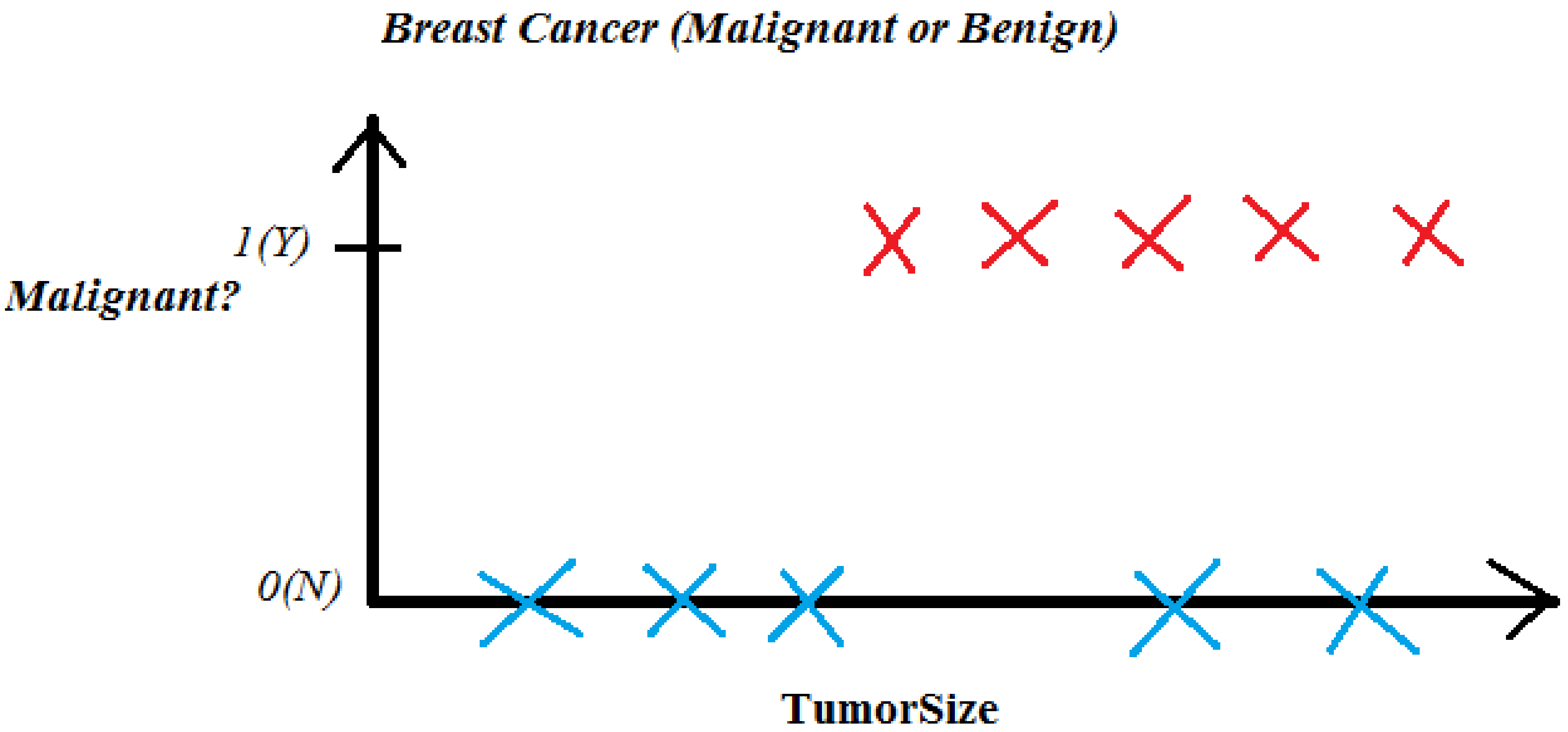


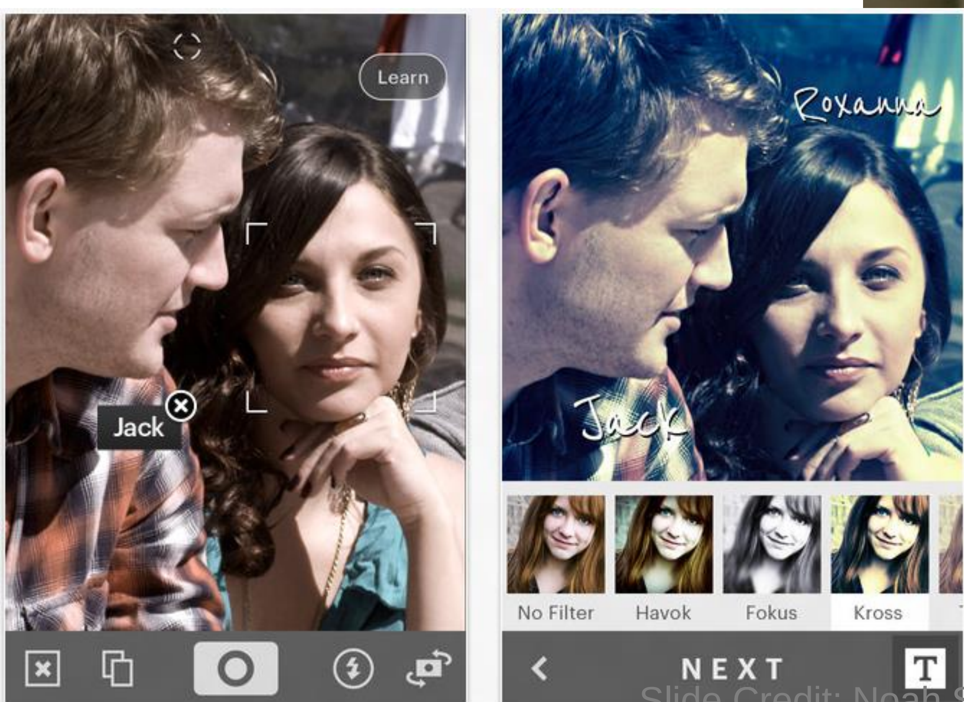
Image Classification

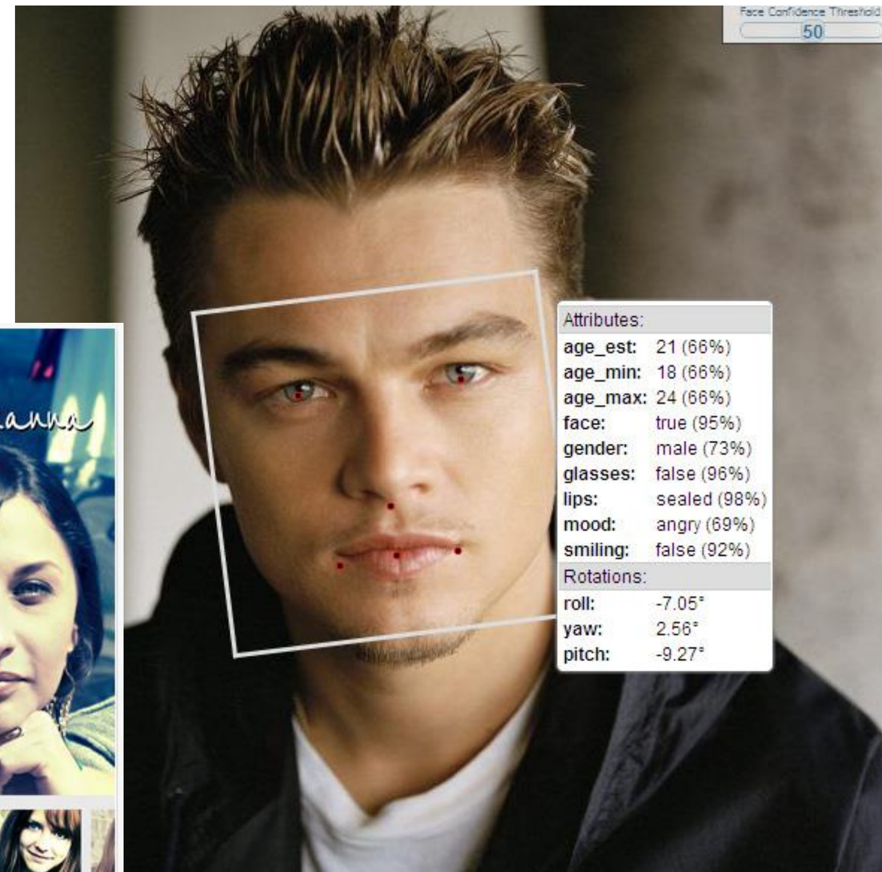
Im2tags; Im2text
http://deeplearning.cs.toronto.edu/

Pizza
Wine
Stove

Face Recognition

http://developers.face.com/tools/





Slide Credit: Noah Snavely

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.

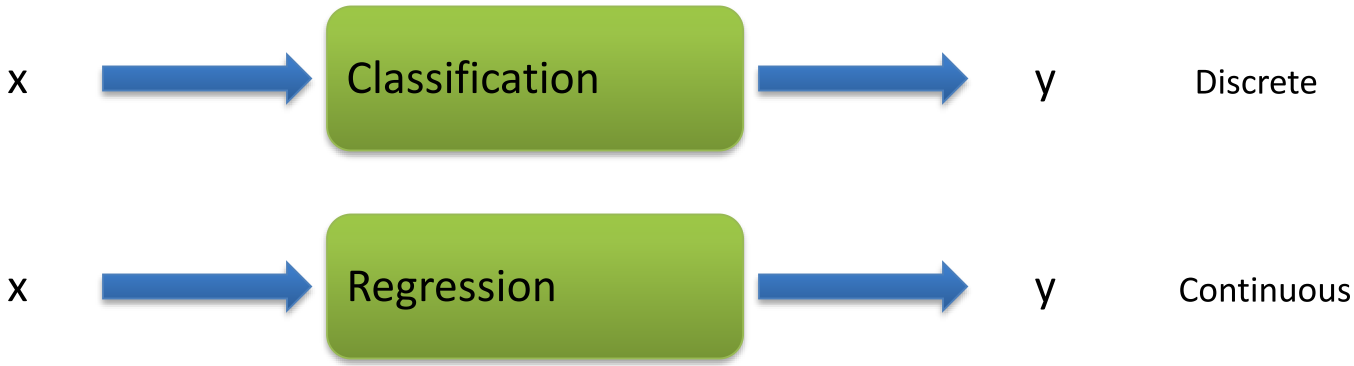
Problem 2: You’d like software to examine individual customer accounts, and for each account decide if it has been hacked/compromised.

Should you treat these as classification or as regression problems?

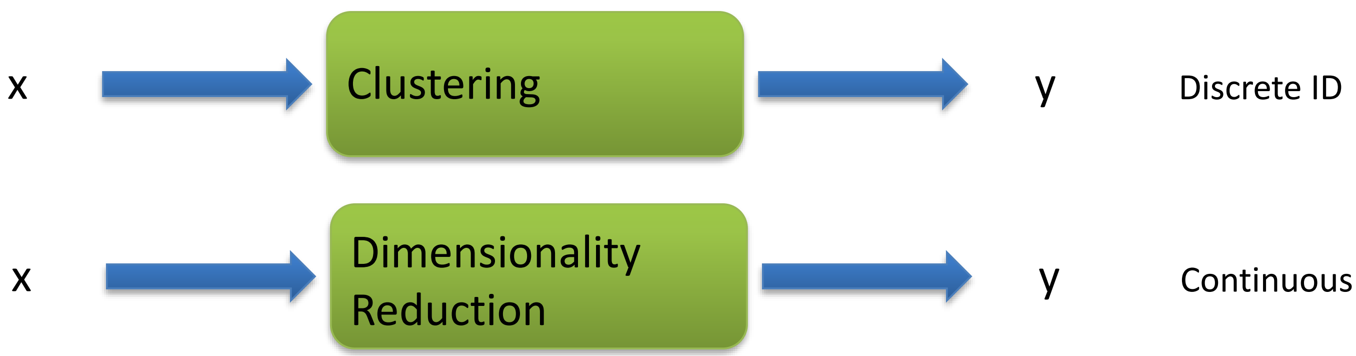
a) Treat both as classification problems.
b) Treat problem 1 as a classification problem, problem 2 as a regression problem.
c) Treat problem 1 as a regression problem, problem 2 as a classification problem.
d) Treat both as regression problems.

Tasks

Supervised Learning



Unsupervised Learning

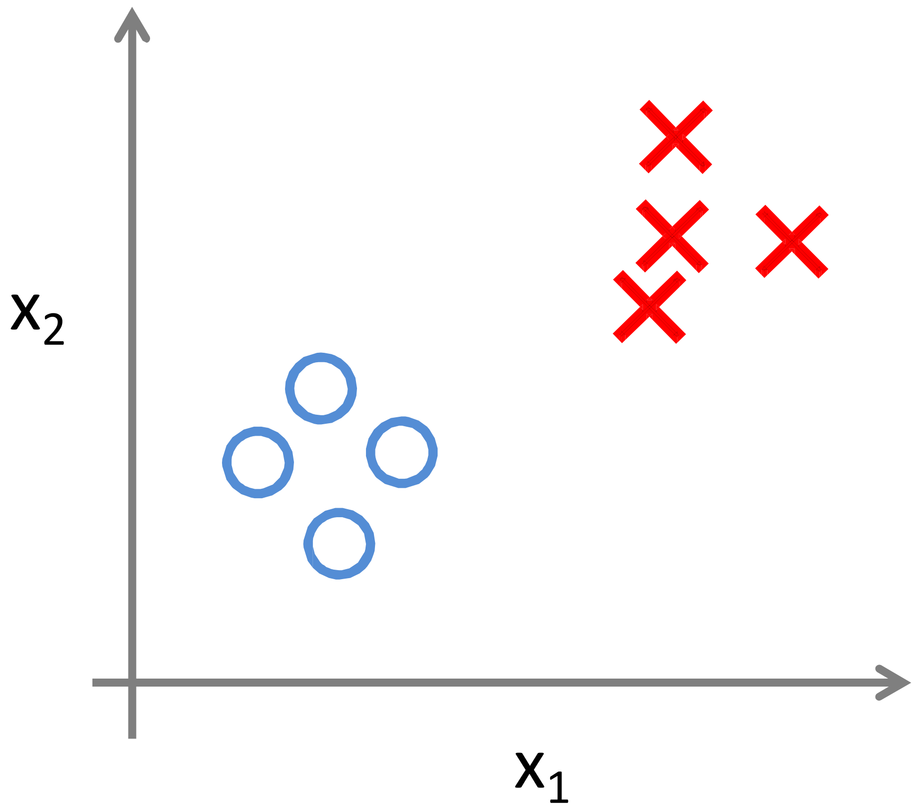


Unsupervised Learning

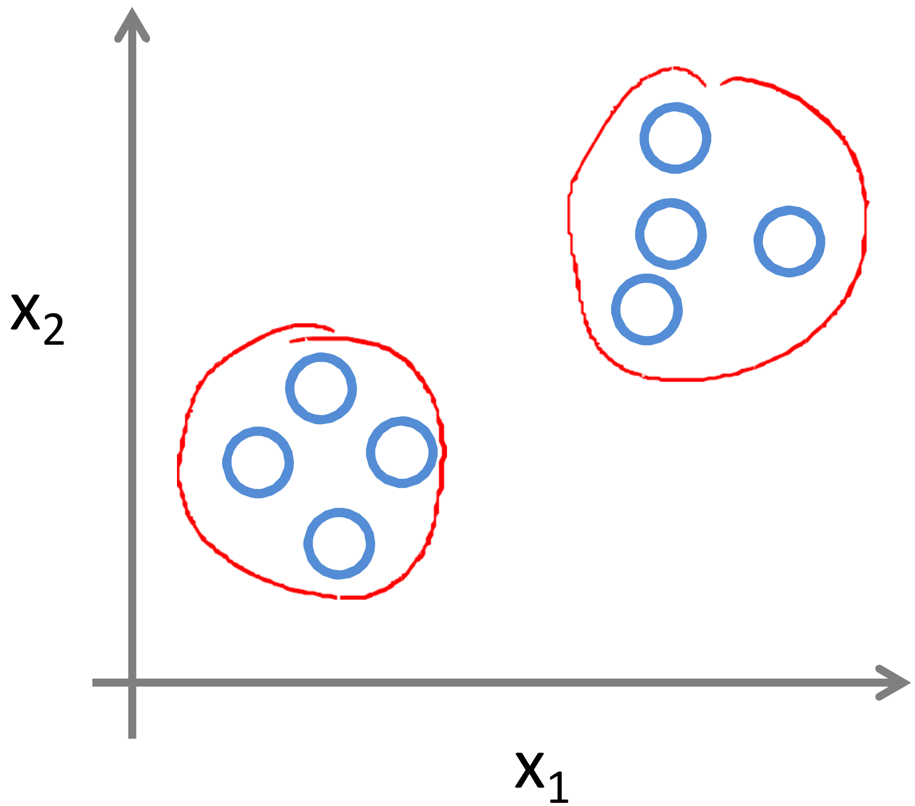
Clustering

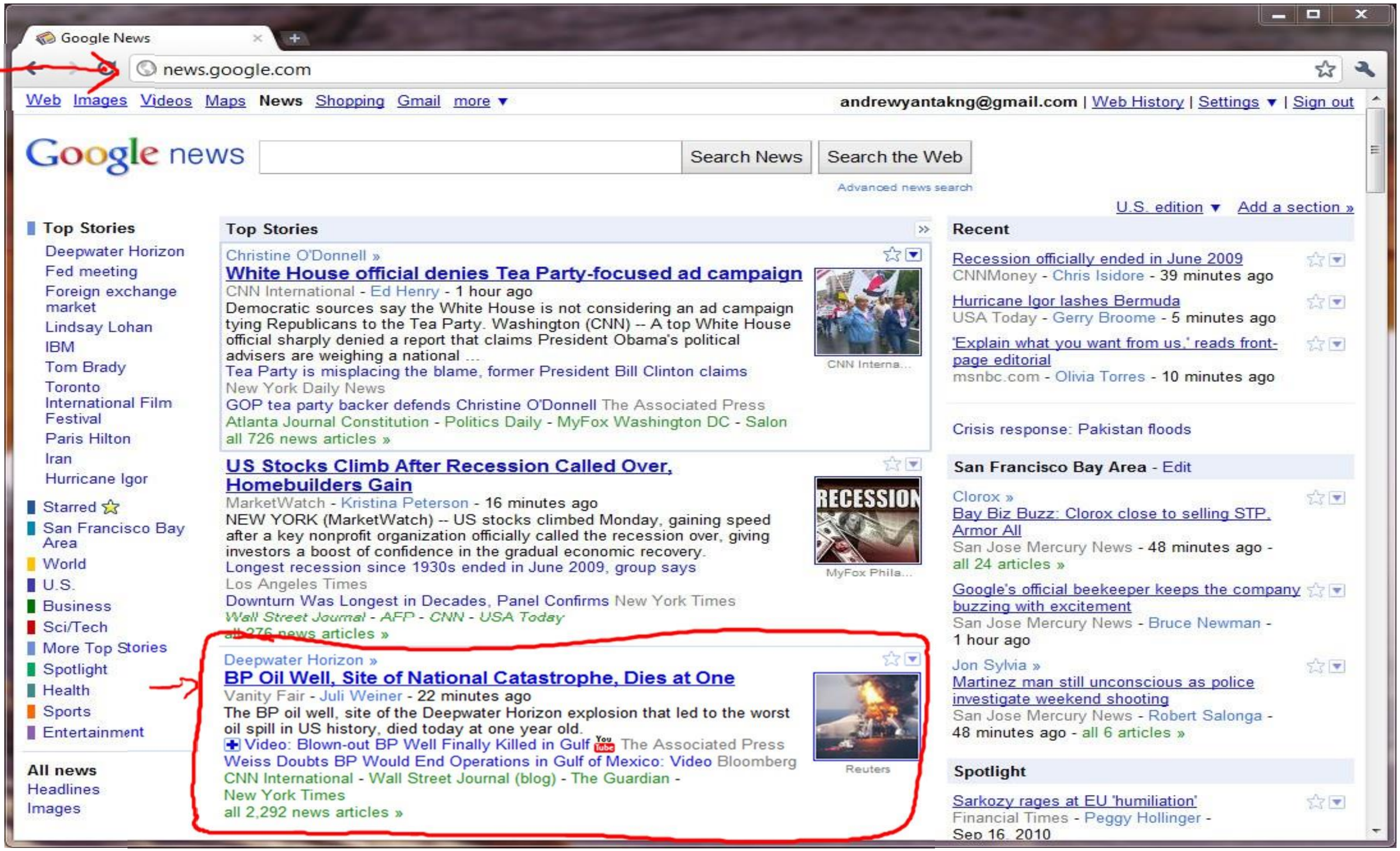
Unsupervised Learning
Y not provided

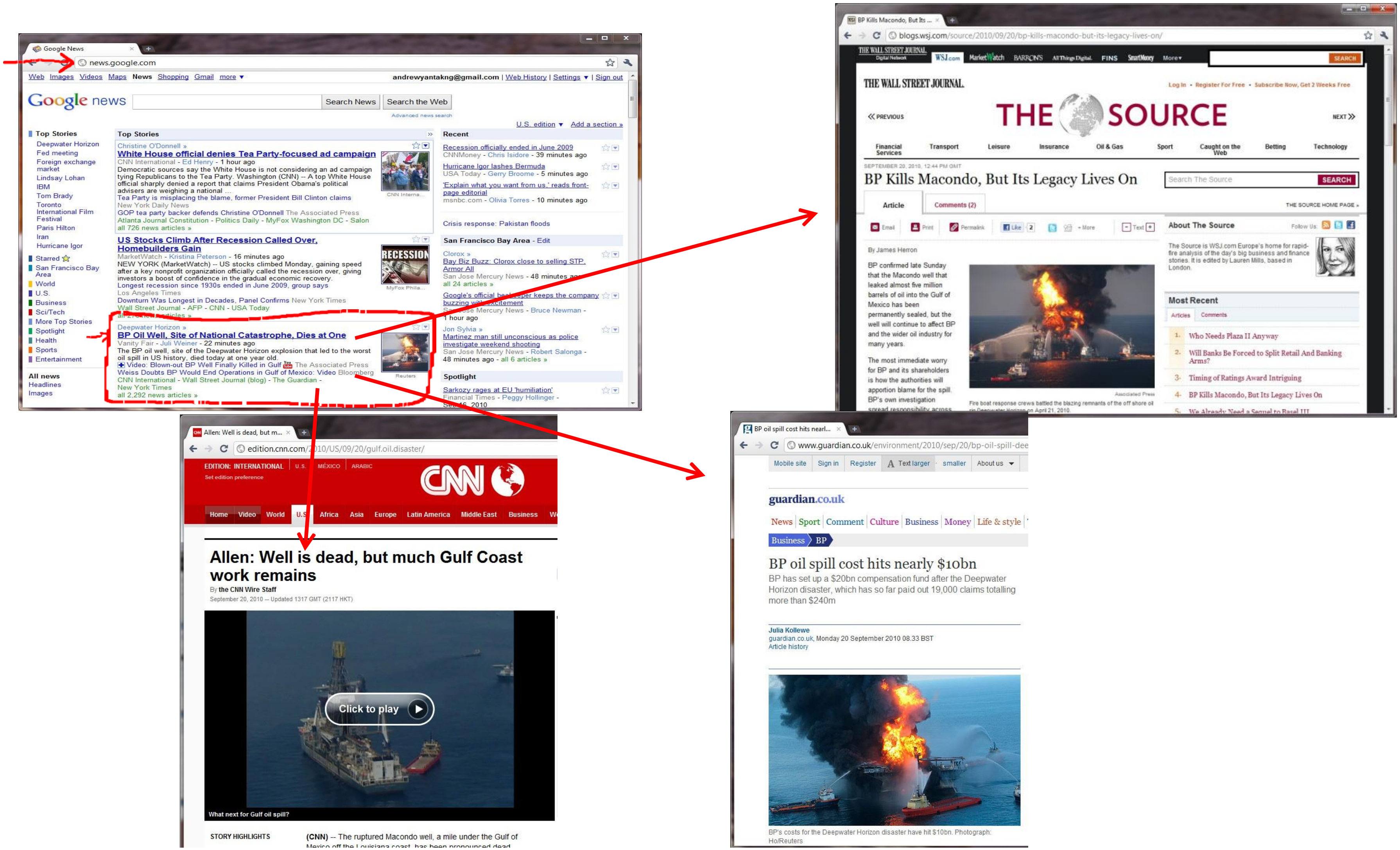
Supervised Learning



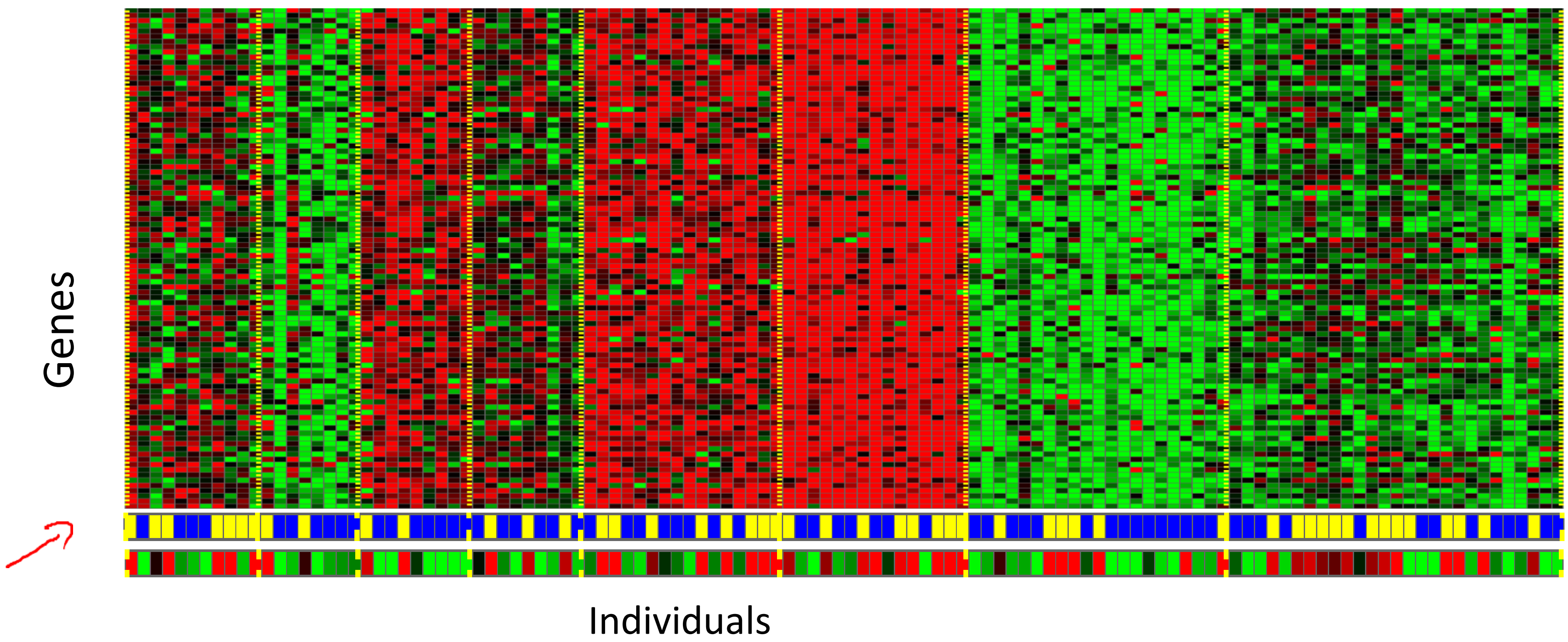
Unsupervised Learning

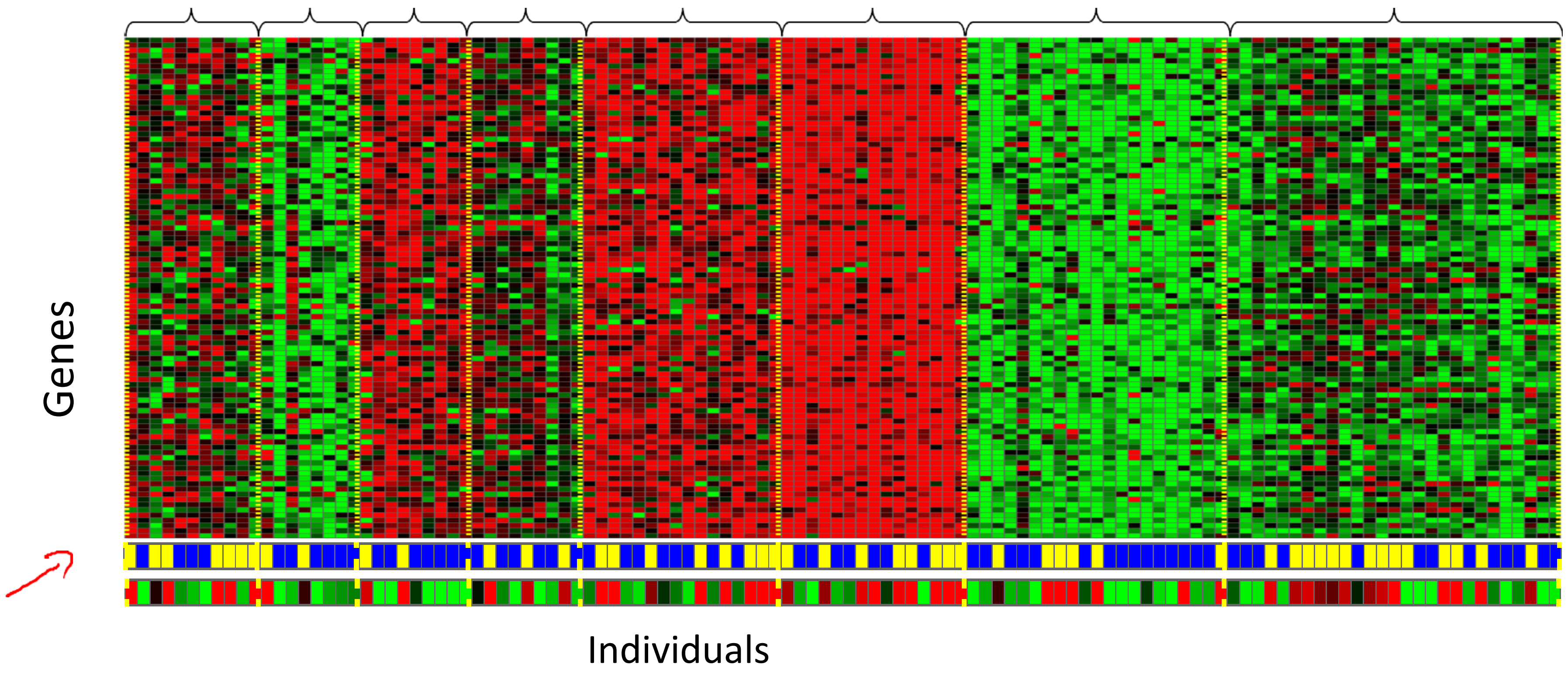



Google News
news.google.com
Web Images Videos Maps News Shopping Gmail more ▾
andrewyantakng@gmail.com | Web History | Settings ▾ | Sign out
Google news
Search News
Search the Web
Advanced news search
U.S. edition ▾ Add a section »
Top Stories
Deepwater Horizon
Fed meeting
Foreign exchange market
Lindsay Lohan
IBM
Tom Brady
Toronto International Film Festival
Paris Hilton
Iran
Hurricane Igor
Starred
San Francisco Bay Area
World
U.S.
Business
Sci/Tech
More Top Stories
Spotlight
Health
Sports
Entertainment
All news
Headlines
Images
Top Stories
Christine O'Donnell »
White House official denies Tea Party-focused ad campaign
CNN International - Ed Henry - 1 hour ago
Democratic sources say the White House is not considering an ad campaign tying Republicans to the Tea Party. Washington (CNN) -- A top White House official sharply denied a report that claims President Obama's political advisers are weighing a national ...
Tea Party is misplacing the blame, former President Bill Clinton claims
New York Daily News
GOP tea party backer defends Christine O'Donnell The Associated Press
Atlanta Journal Constitution - Politics Daily - MyFox Washington DC - Salon
all 726 news articles »
CNN Interna...
US Stocks Climb After Recession Called Over, Homebuilders Gain
MarketWatch - Kristina Peterson - 16 minutes ago
NEW YORK (MarketWatch) -- US stocks climbed Monday, gaining speed after a key nonprofit organization officially called the recession over, giving investors a boost of confidence in the gradual economic recovery.
Longest recession since 1930s ended in June 2009, group says
Los Angeles Times
Downturn Was Longest in Decades, Panel Confirms New York Times
Wall Street Journal - AFP - CNN - USA Today
all 276 news articles »
RECESSION
MyFox Phila...
Deepwater Horizon »
BP Oil Well, Site of National Catastrophe, Dies at One
Vanity Fair - Juli Weiner - 22 minutes ago
The BP oil well, site of the Deepwater Horizon explosion that led to the worst oil spill in US history, died today at one year old.
Video: Blown-out BP Well Finally Killed in Gulf The Associated Press
Weiss Doubts BP Would End Operations in Gulf of Mexico: Video Bloomberg
CNN International - Wall Street Journal (blog) - The Guardian - New York Times
all 2,292 news articles »
Reuters
Recent
Recession officially ended in June 2009
CNNMoney - Chris Isidore - 39 minutes ago
Hurricane Igor lashes Bermuda
USA Today - Gerry Broome - 5 minutes ago
'Explain what you want from us,' reads front-page editorial
msnbc.com - Olivia Torres - 10 minutes ago
Crisis response: Pakistan floods
San Francisco Bay Area - Edit
Clorox »
Bay Biz Buzz: Clorox close to selling STP, Armor All
San Jose Mercury News - 48 minutes ago - all 24 articles »
Google's official beekeeper keeps the company buzzing with excitement
San Jose Mercury News - Bruce Newman - 1 hour ago
Jon Sylvia »
Martinez man still unconscious as police investigate weekend shooting
San Jose Mercury News - Robert Salonga - 48 minutes ago - all 6 articles »
Spotlight
Sarkozy rages at EU 'humiliation'
Financial Times - Peggy Hollinger - Sep 16, 2010



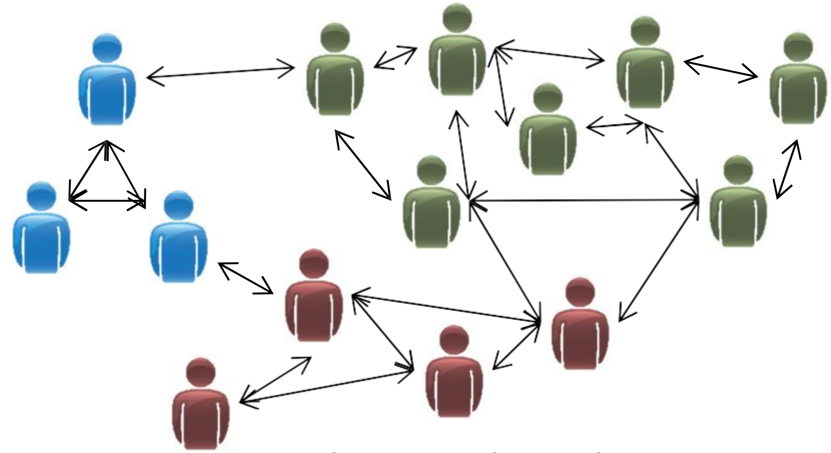
Google News
news.google.com
Google news
Search News
Search the Web
Top Stories
White House official denies Tea Party-focused ad campaign
US Stocks Climb After Recession Called Over, Homebuilders Gain
Deepwater Horizon
BP Oil Well, Site of National Catastrophe, Dies at One
Vanity Fair - Juli Weiner - 22 minutes ago
The BP oil well, site of the Deepwater Horizon explosion that led to the worst oil spill in US history, died today at one year old.
Weiss Doubts BP Would End Operations in Gulf of Mexico: Video Bloomberg
CNN International - Wall Street Journal (blog) - The Guardian - New York Times
all 2,292 news articles »
Recent
Recession officially ended in June 2009
Hurricane Igor lashes Bermuda
Spotlight
Sarkozy rages at EU 'humiliation'
blogs.wsj.com/source/2010/09/20/bp-kills-macondo-but-its-legacy-lives-on/
THE WALL STREET JOURNAL.
THE SOURCE
BP Kills Macondo, But Its Legacy Lives On
By James Herron
BP confirmed late Sunday that the Macondo well that leaked almost five million barrels of oil into the Gulf of Mexico has been permanently sealed, but the well will continue to affect BP and the wider oil industry for many years.
The most immediate worry for BP and its shareholders is how the authorities will apportion blame for the spill. BP's own investigation
About The Source
Most Recent
Who Needs Plaza II Anyway
Will Banks Be Forced to Split Retail And Banking Arms?
Timing of Ratings Award Intriguing
BP Kills Macondo, But Its Legacy Lives On
edition.cnn.com/2010/US/09/20/gulf.oil.disaster/
CNN
Allen: Well is dead, but much Gulf Coast work remains
By the CNN Wire Staff
September 20, 2010 -- Updated 1317 GMT (2117 HKT)
Click to play
What next for Gulf oil spill?
STORY HIGHLIGHTS
(CNN) -- The ruptured Macondo well, a mile under the Gulf of
www.guardian.co.uk/environment/2010/sep/20/bp-oil-spill-dee
guardian.co.uk
News Sport Comment Culture Business Money Life & style
Business BP
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
Article history
BP's costs for the Deepwater Horizon disaster have hit $10bn. Photograph: Ho/Reuters


Genes
Individuals

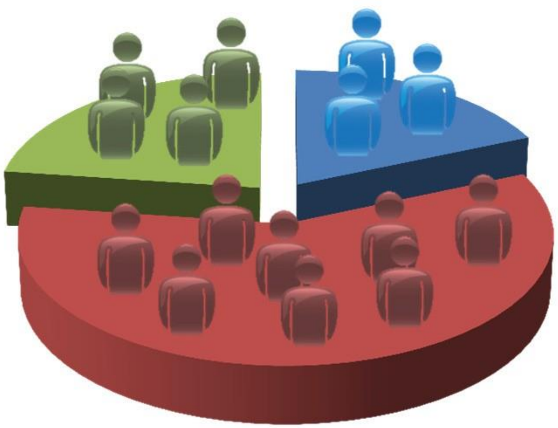


[Source: Daphne Koller]

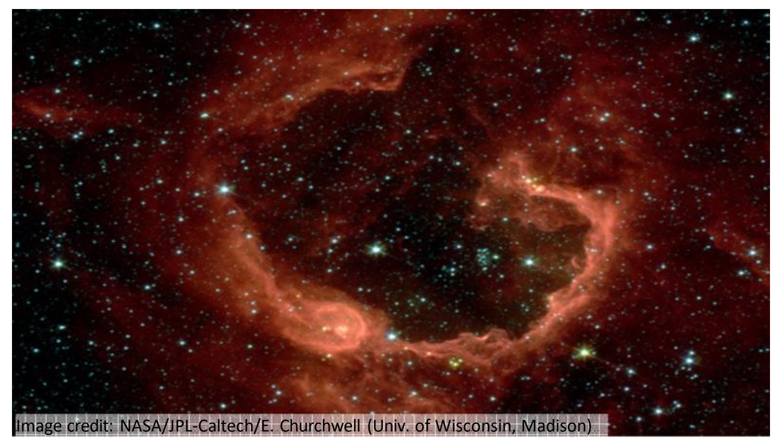
Organize computing clusters



Social network analysis



Market segmentation



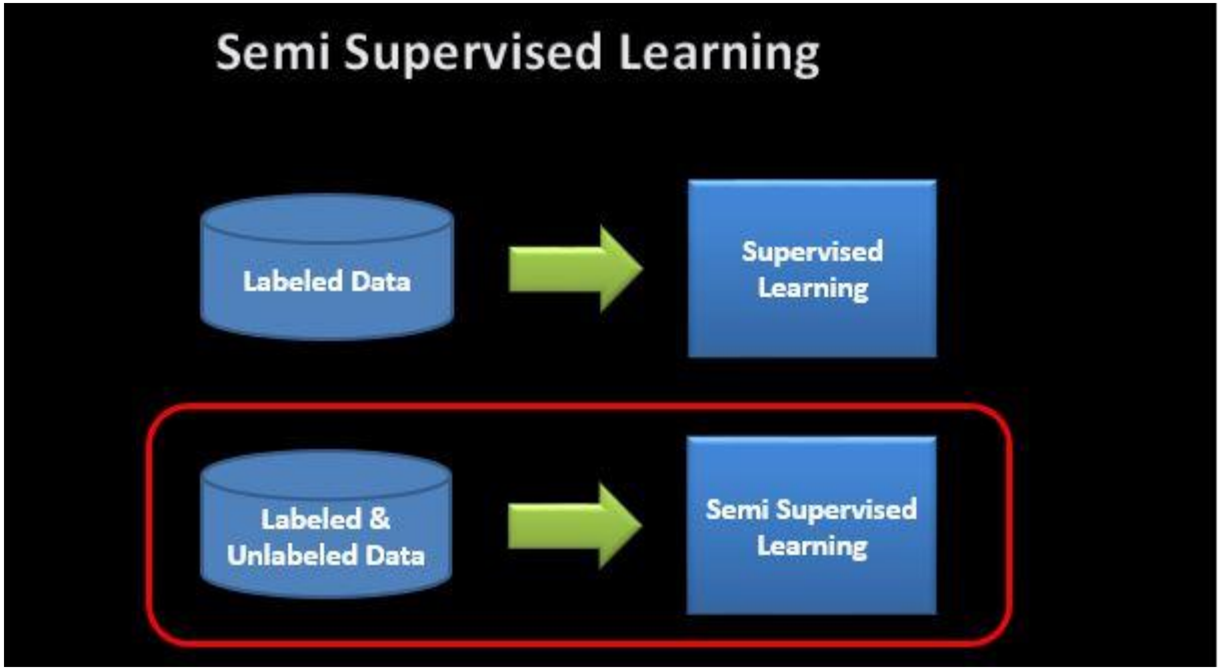
Astronomical data analysis

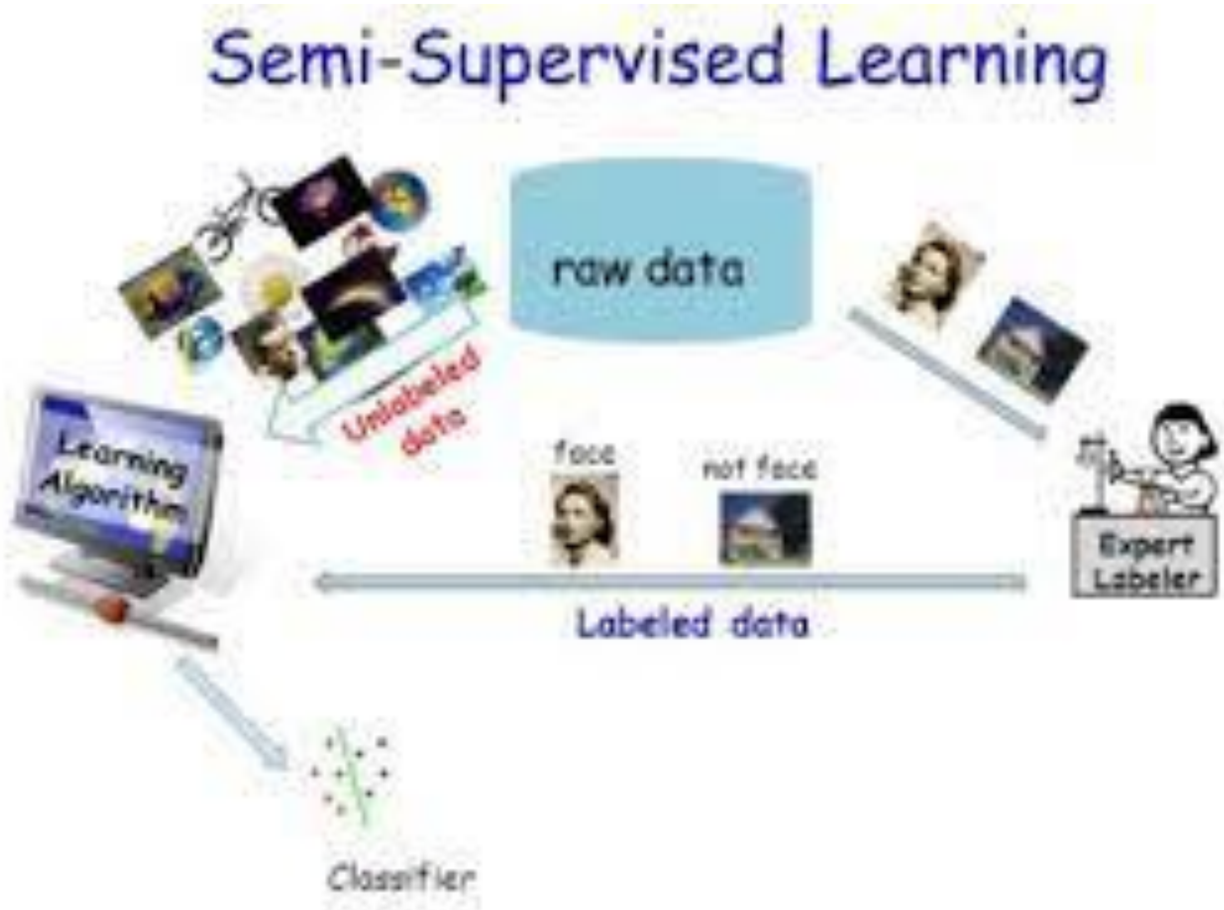
Of the following examples, which would you address using an <u>unsupervised</u> learning algorithm? (Check all that apply.)

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

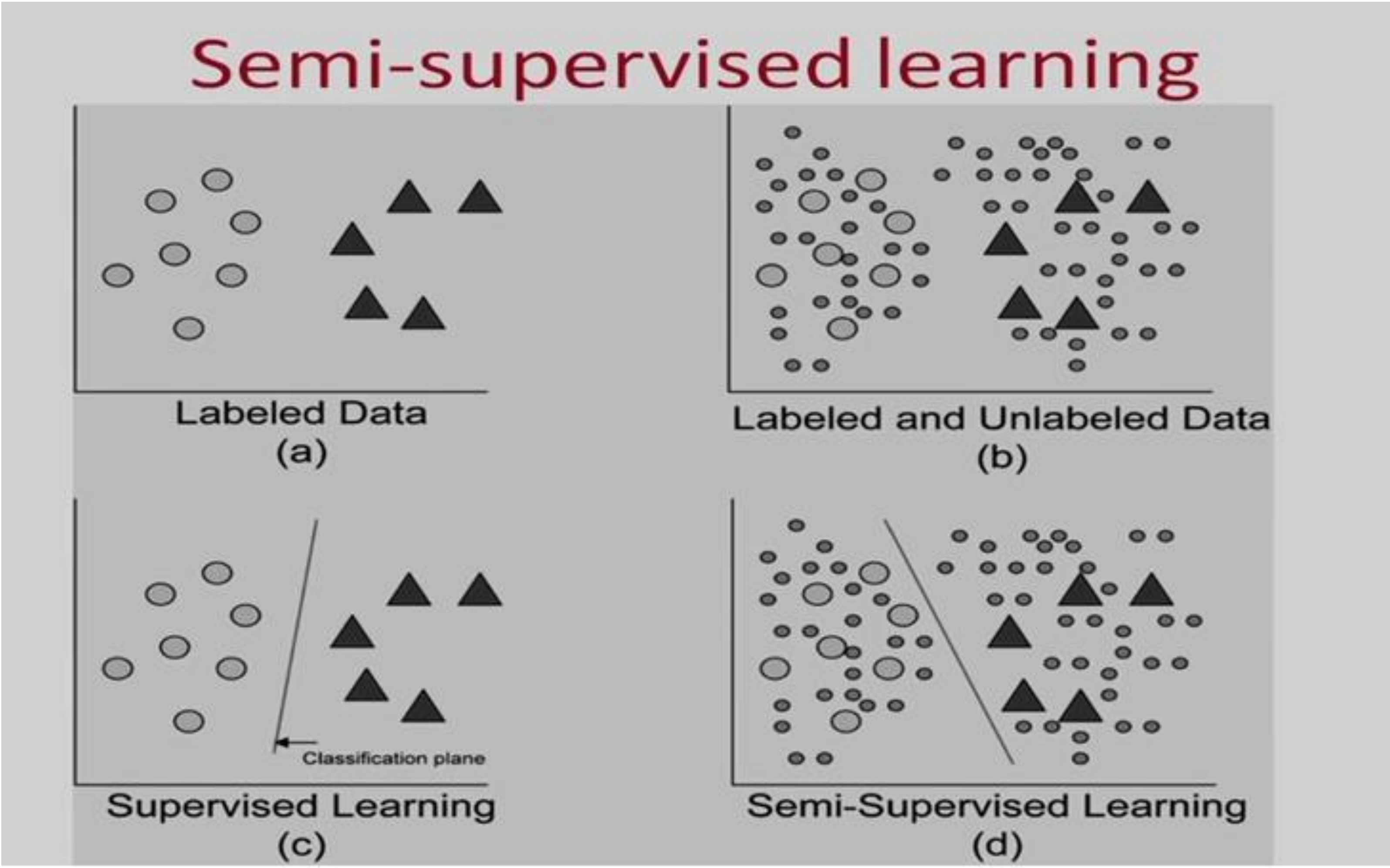
Semi-supervised learning

- Represents the intermediate ground between Supervised and Unsupervised learning algorithms
- It uses a combination of labeled and unlabeled datasets during the training period
- Training data is a combination of both labeled and unlabeled data
- Initially, similar data is clustered along with an unsupervised learning algorithm, and further, it helps to label the unlabeled data into labeled data





Semi-Supervised Learning
raw data
Unlabeled data
Learning Algorithm
face
not face
Expert Labeler
Labeled data
Classifier


Semi-supervised learning
Labeled Data
(a)
Labeled and Unlabeled Data
(b)
Classification plane
Supervised Learning
(c)
Semi-Supervised Learning
(d)

Reinforcement Learning

Learning from feedback

Reinforcement Learning: Learning to act

There is only one “supervised” signal at the end of the game.

But you need to make a move at every step

RL deals with “credit assignment”

ML in a Nutshell

- Tens of thousands of machine learning algorithms
  - Hundreds new every year

- Every ML algorithm has three components:
  - **Representation**
  - **Optimization**
  - **Evaluation**

Various Function Representations

- Numerical functions
  - Linear regression
  - Neural networks
  - Support vector machines
- Symbolic functions
  - Decision trees
  - Rules in propositional logic
  - Rules in first-order predicate logic
- Instance-based functions
  - Nearest-neighbor
  - Case-based
- Probabilistic Graphical Models
  - Naïve Bayes
  - Bayesian networks
  - Hidden-Markov Models (HMMs)
  - Probabilistic Context Free Grammars (PCFGs)
  - Markov networks

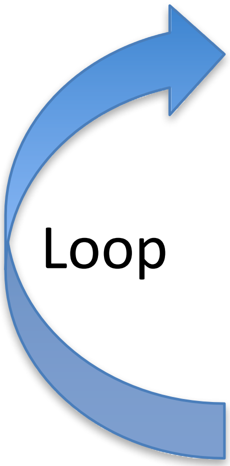
Various Search/Optimization Algorithms

- Gradient descent
  - Perceptron
  - Backpropagation
- Dynamic Programming
  - HMM Learning
  - PCFG Learning
- Divide and Conquer
  - Decision tree induction
  - Rule learning
- Evolutionary Computation
  - Genetic Algorithms (GAs)
  - Genetic Programming (GP)
  - Neuro-evolution

Evaluation

- Accuracy
- Precision and recall
- Squared error
- Likelihood
- Posterior probability
- Cost / Utility
- Margin
- Entropy
- K-L divergence
- etc.

ML in Practice



- Understand domain, prior knowledge, and goals
- Data integration, selection, cleaning, pre-processing, etc.
- Learn models
- Interpret results
- Consolidate and deploy discovered knowledge

A Brief History of Machine Learning

History of Machine Learning

- 1950s
  - Samuel’s checker player
  - Selfridge’s Pandemonium
- 1960s:
  - Neural networks: Perceptron
  - Pattern recognition
  - Learning in the limit theory
  - Minsky and Papert prove limitations of Perceptron
- 1970s:
  - Symbolic concept induction
  - Winston’s arch learner
  - Expert systems and the knowledge acquisition bottleneck
  - Quinlan’s ID3
  - Michalski’s AQ and soybean diagnosis
  - Scientific discovery with BACON
  - Mathematical discovery with AM

History of Machine Learning (cont.)

- 1980s:
  - Advanced decision tree and rule learning
  - Explanation-based Learning (EBL)
  - Learning and planning and problem solving
  - Utility problem
  - Analogy
  - Cognitive architectures
  - Resurgence of neural networks (connectionism, backpropagation)
  - Valiant's PAC Learning Theory
  - Focus on experimental methodology
- 1990s
  - Data mining
  - Adaptive software agents and web applications
  - Text learning
  - Reinforcement learning (RL)
  - Inductive Logic Programming (ILP)
  - Ensembles: Bagging, Boosting, and Stacking
  - Bayes Net learning

History of Machine Learning (cont.)

- 2000s
  - Support vector machines & kernel methods
  - Graphical models
  - Statistical relational learning
  - Transfer learning
  - Sequence labeling
  - Collective classification and structured outputs
  - Computer Systems Applications (Compilers, Debugging, Graphics, Security)
  - E-mail management
  - Personalized assistants that learn
  - Learning in robotics and vision
- 2010s
  - Deep learning systems
  - Learning for big data
  - Bayesian methods
  - Multi-task & lifelong learning
  - Applications to vision, speech, social networks, learning to read, etc.