

Machine Learning

BCSE3066

Instructor: Pratyush Kumar Deka

These slides have been adapted from [Eric Eaton](#), with grateful acknowledgement of the many others who made their course materials freely available online. Feel free to reuse or adapt these slides for your own academic purposes, provided that you include proper attribution. Please send comments and corrections..

Machine Learning is Everywhere

Mail thinks this message is Junk Mail.

? Load Images Not Junk

From: IEEE World Congress on Multimedia <iccsa2013@yahoo.com> **Hide**

Subject: First CFP Submission :15 August, 2013 World Congress on Multimedia & Computer science: October 04-06, 2013, Hammamet, Tunisia

Date: July 17, 2013 9:01:20 AM CDT

To: Julia Hockenmaier

Reply-To: iccsa2013@yahoo.com

7 Attachments, 374 KB Save Quick Look

First- Call For Papers Submission : 15th of



World Congress on Multimedia and Computer science (WCMCS' 2013) October 04-06, 2013, Hammamet, Tunisia

Azure Machine Learning - Microsoft.com
Ad www.microsoft.com/MachineLearning Powerful Machine Learning Service With all the Benefits of the Cloud.

Predictive Sales - 6sense.com
Ad www.6sense.com GSense Tells You Who Will Buy, When They'll Buy, and How Much. Request a demo - GSense for Sales

Machine Learning Mitchell - Amazon.com
Ad www.amazon.com/Books Books to Satisfy Anyone from New Learners to Computer Geeks.

Amazon.com: Machine Learning: Books
www.amazon.com/Machine_Learning... - Amazon.com Results 1 - 12 of 2112 - Online shopping for Machine Learning from a great selection at Books Store. Kindle Edition - JH Friedman - Last 30 days - German

PDF INTRODUCTION TO MACHINE LEARNING - Alex Smola
alex.smola@rafts.thebook.pdf Results 1 - 10 - A catalogue record for this book is available from the British Library ... This page has pointers to my draft book on Machine Learning and to its ...

Machine Learning Books : MachineLearning - Reddit
www.reddit.com/r/MachineLearning/_/machine_learning_books Reddit Jul 30, 2013 - I have been collecting machine learning books over the past couple months. It seems that machine learning professors are good about posting ...

FAQ: What machine learning book should I start with? - Benmbaby.com.../faq-what-machine-learning-book-should-start-with... Oct 7, 2011 - What makes a good starting book is dependent on your background (specifically ... It gives the reader an intuition on several machine learning ...

What are some good books on machine learning? - Quora
www.quora.com/What-are-some-good-books-on-machine-learning Quora Answer 1 of 15: This answer attempts the very ambitious problem of producing an approximately complete list. Please leave comments and tell me what's wrong ...

Amazon.com: Why is this recommended for you?

amazon.com Help | Close window

Recommended for You

Semantics, Second Edition (Modern Linguistics)
by Kate Kearns (May 15, 2011)
In Stock
List Price: \$40.00
Price: \$37.31
68 used & new from \$14.00

Add to Cart Add to Wish List

Because you purchased...

Meaning: A Slim Guide to Semantics (Oxford Linguistics) (Paperback)
by Paul Elbourne (Author)

This was a gift
 Don't use for recommendations

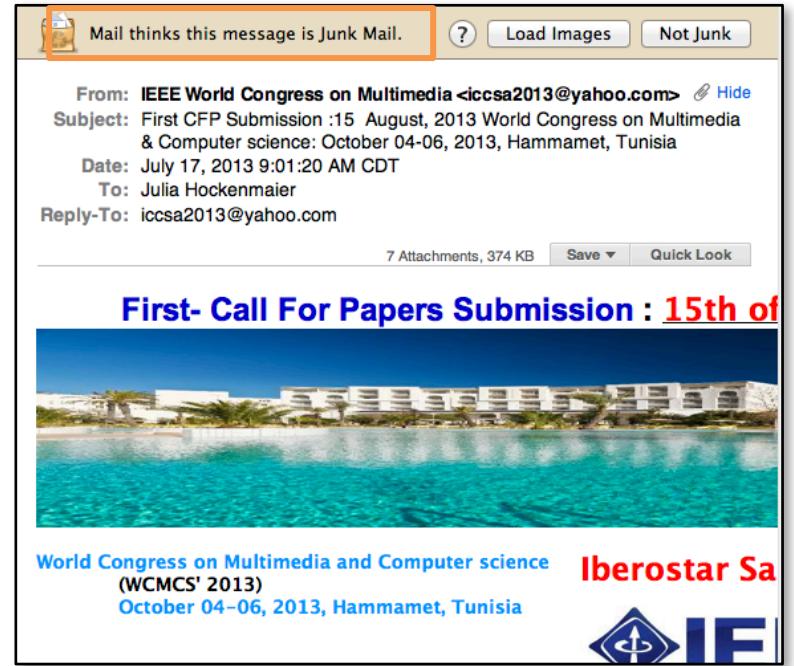
Google Translate From: English - detected To: Chinese (Simplified) Translate

English Spanish French English - detected Chinese (Simplified) English Spanish

The blue fox jumps over the hedge 蓝狐跨越对冲

Applications: Spam Detection

- This is a **binary classification task**: Assign **one of two labels** (i.e. yes/no) to the input (here, an email message)
- Classification requires a model (**a classifier**) to determine which label to assign to items.
- In this class, we study **algorithms and techniques** to learn such models from data.



<u>Data</u>	<u>Labels</u>
Documents	Politics, Sports, Finance
Sentences	Positive, Negative
Phrases	Person, Location
Images	cats, dogs, snakes
Medical records	Re-admit soon/Not
.....	

Applications: Object Recognition

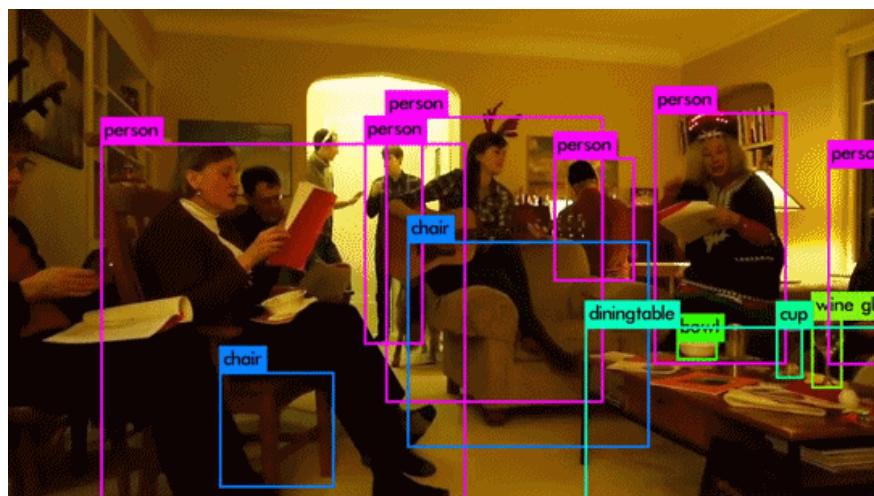
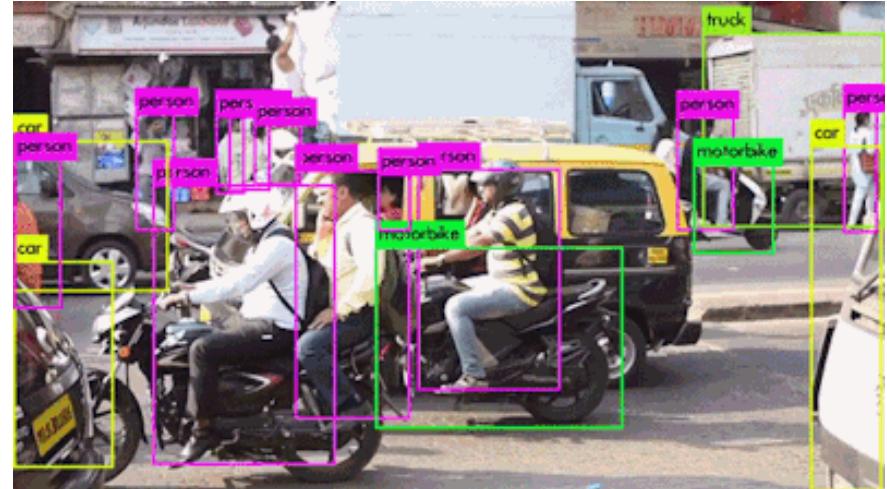
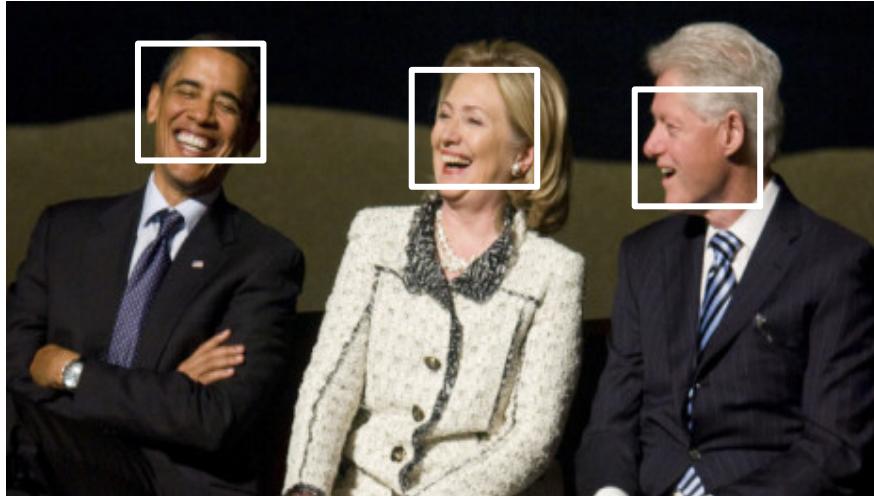


Image credits: <http://prostheticknowledge.tumblr.com/post/172309765681/yolo-v3new-release-of-object-recognition-framework>
<https://junshengfu.github.io/vehicle-detection/>

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

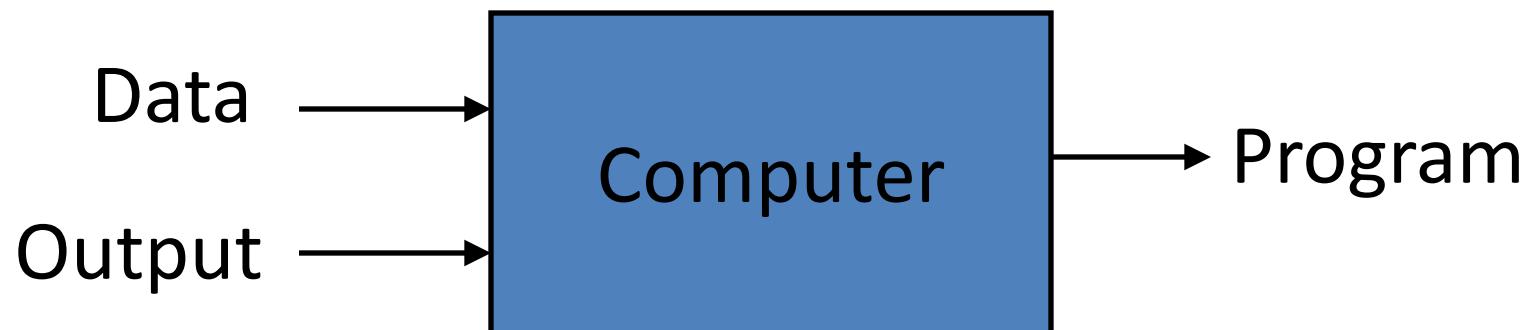
- improve their performance P
- at some task T
- with experience E .

A well-defined learning task is given by $\langle P, T, E \rangle$.

Traditional Programming



Machine Learning



Why Study Machine Learning?

“A breakthrough in machine learning would be worth ten Microsofts”

-Bill Gates, Chairman, Microsoft

“Machine learning is the next Internet”

-Tony Tether, Director, DARPA

Machine learning is the hot new thing”

-John Hennessy, President, Stanford

“Web rankings today are mostly a matter of machine learning”

-Prabhakar Raghavan, Dir. Research, Yahoo

“Machine learning is going to result in a real revolution”

-Greg Papadopoulos, CTO, Sun

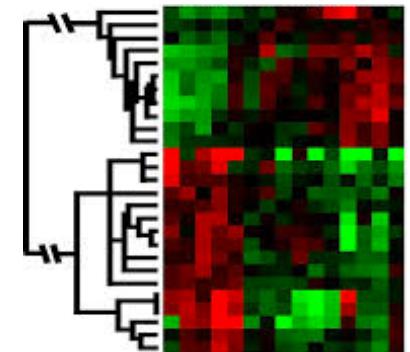
“Machine learning is today’s discontinuity”

-Jerry Yang, CEO, Yahoo

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)



Learning isn't always useful:

- There is no need to “learn” to calculate payroll

A classic example of a task that requires machine learning:
It is very hard to say what makes a 2

0 0 0 1 1 (1 1 1 2

2 2 2 2 2 2 2 3 3 3

3 4 4 4 4 4 5 5 5 5

6 6 7 7 7 7 7 8 8 8

8 8 8 7 9 4 9 9 9

Learning = Generalization

H. Simon -

“Learning denotes changes in the system that are adaptive in the sense that they enable the system to do the task or tasks drawn from the same population more efficiently and more effectively the next time.”

The ability to perform a task in a situation which has never been encountered before

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
- Computational biology
- Finance
- E-commerce
- Space exploration
- Robotics
- Information extraction
- Social networks
- Debugging software
- [Your favorite area]

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.

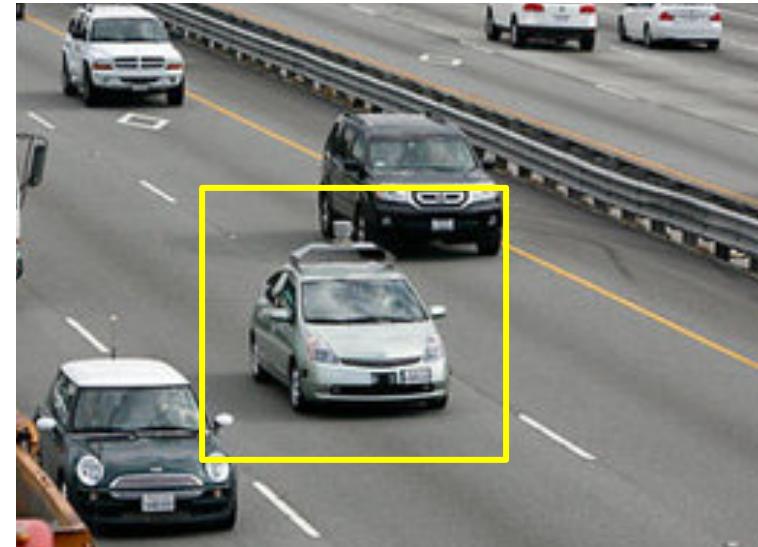
T: Categorize email messages as spam or legitimate.

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

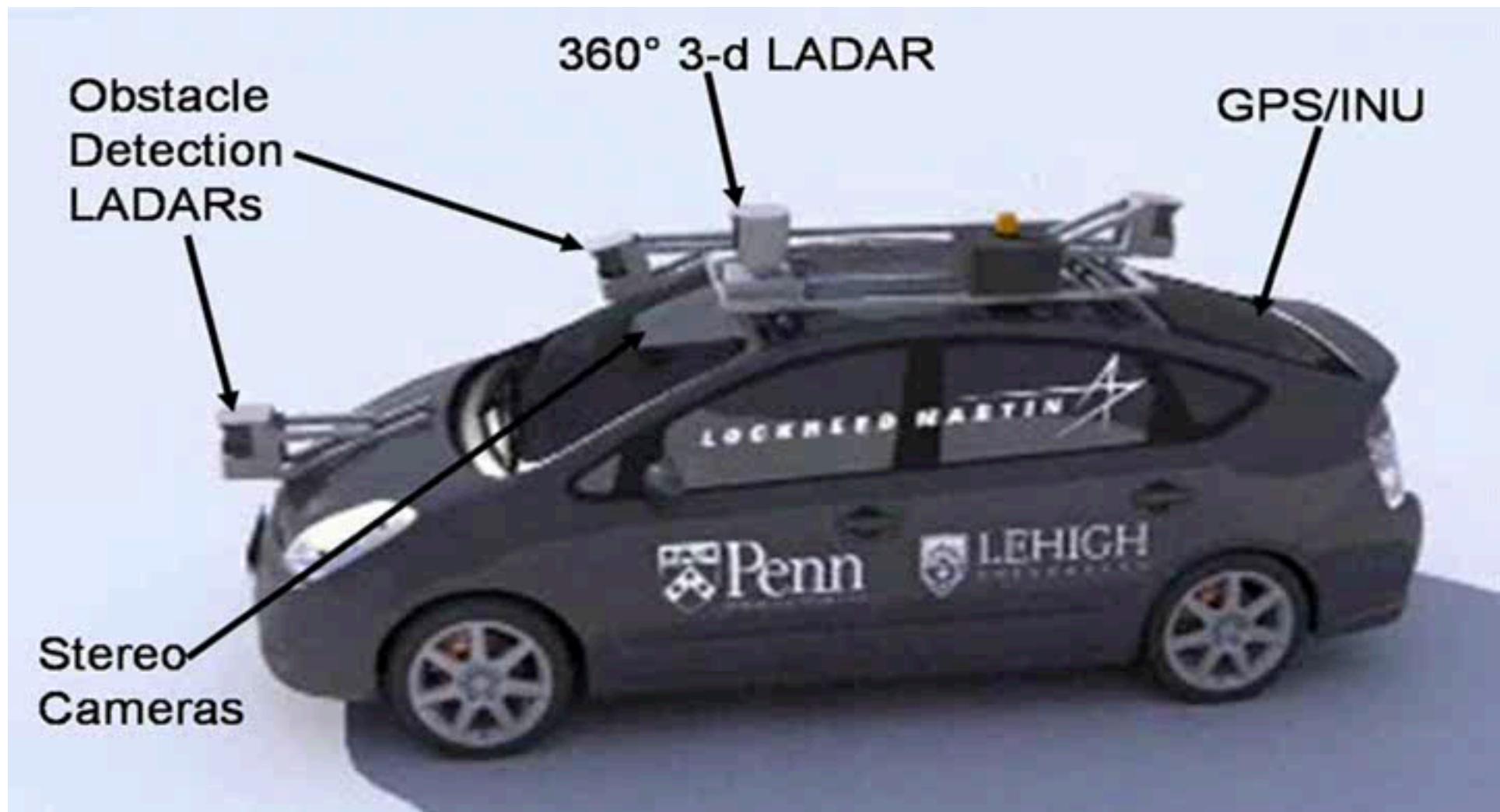


- Nevada made it legal for autonomous cars to drive on roads in June 2011
- As of 2017, 29 states have enacted legislation regarding autonomous cars

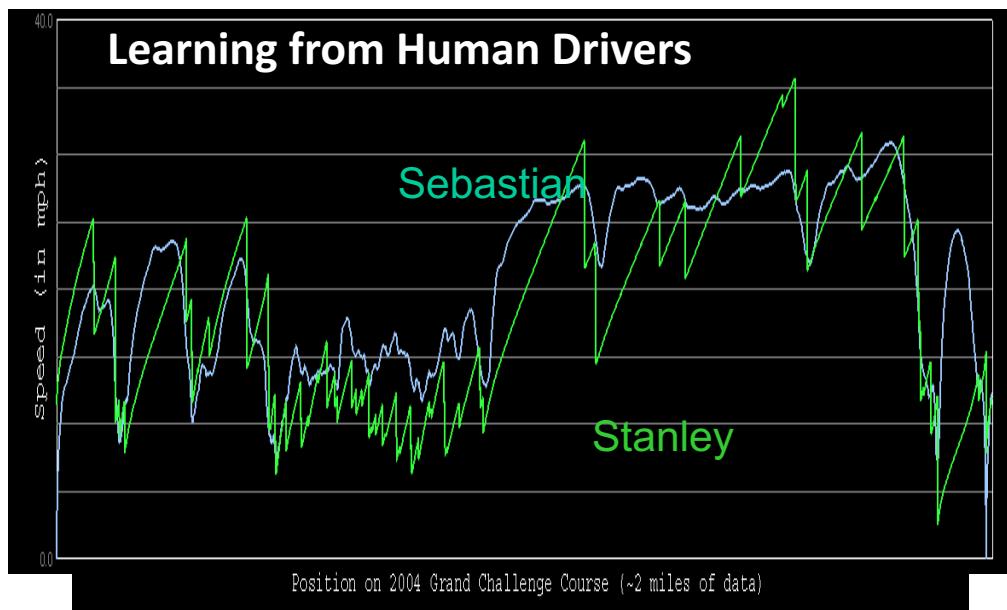
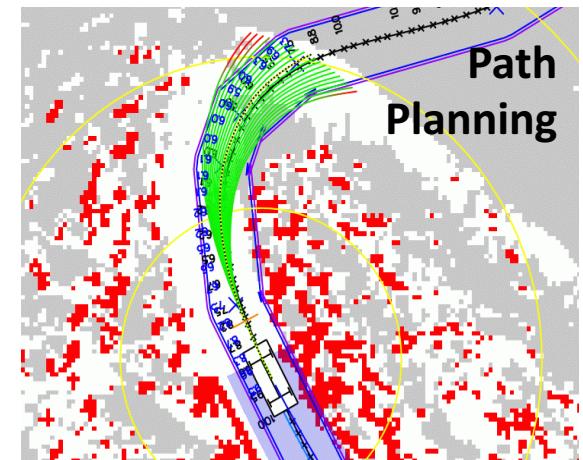


Penn's Autonomous Car →
(Ben Franklin Racing Team)

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

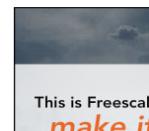
Is Google Cornering the Market on Deep Learning?

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.

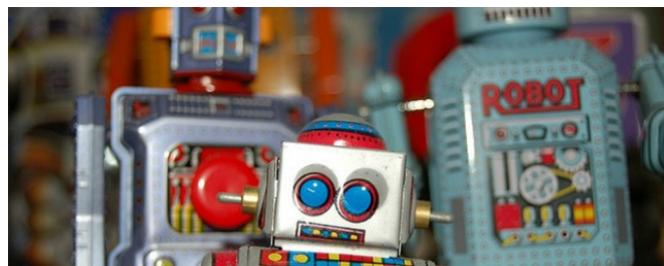


This week, Google [reportedly paid that much](#) to acquire [DeepMind Technologies](#), a startup based in

WIRED GEAR SCIENCE ENTERTAINMENT BUSINESS SECURITY DESIGN
INNOVATION INSIGHTS | [community content](#) | ▾ — featured

Deep Learning's Role in the Age of Robots

BY JULIAN GREEN, JETPAC 05.02.14 2:56 PM



BloombergBusinessweek
Technology

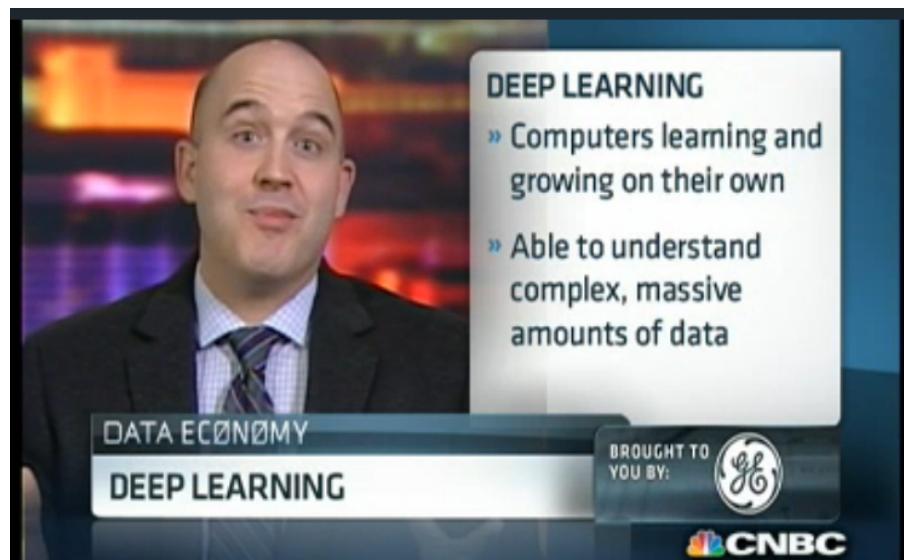
Acquisitions

The Race to Buy the Human Brains Behind Deep Learning Machines

By Ashlee Vance [Twitter](#) | 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



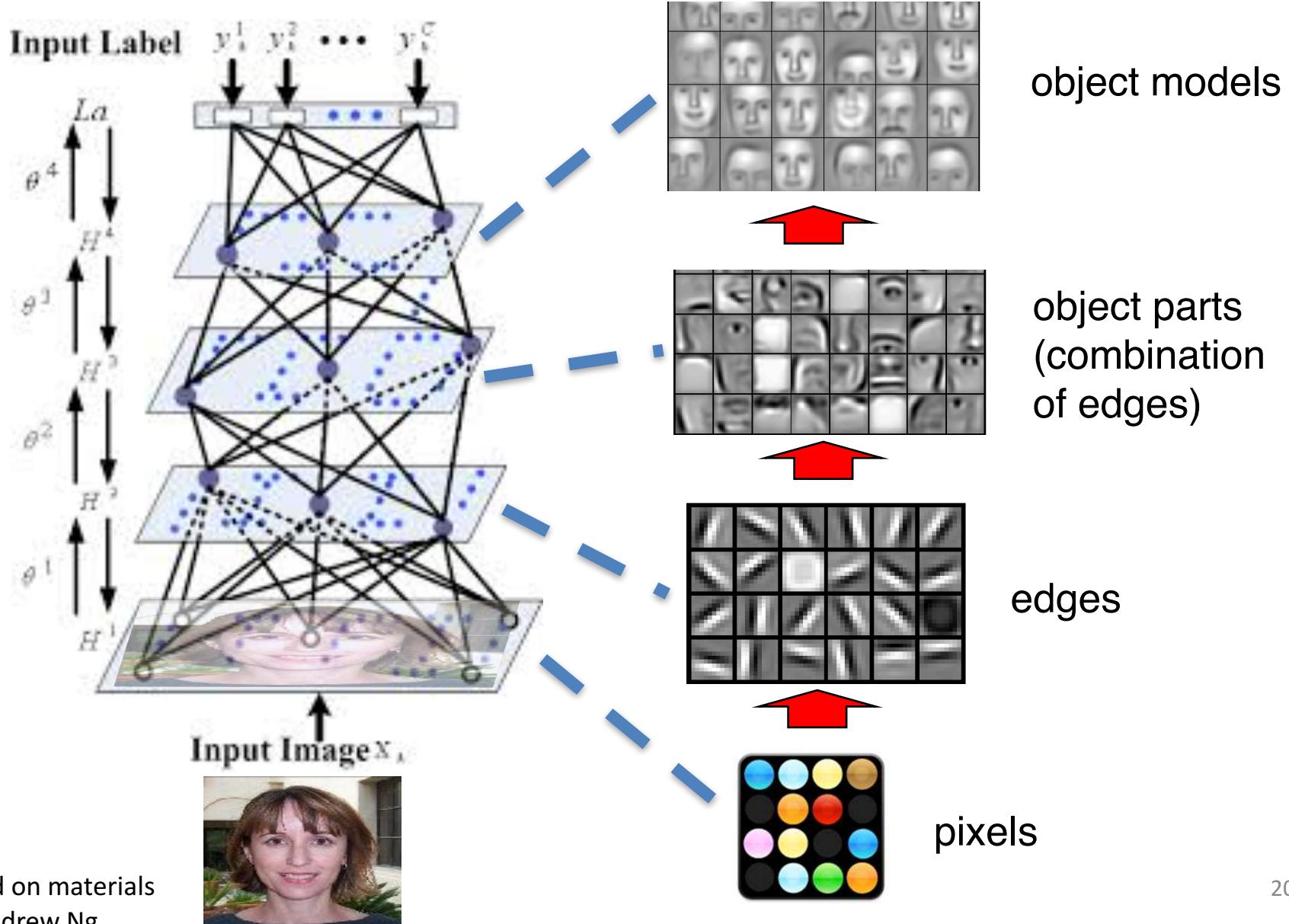
DEEP LEARNING

- » Computers learning and growing on their own
- » Able to understand complex, massive amounts of data

DATA ECONOMY
DEEP LEARNING

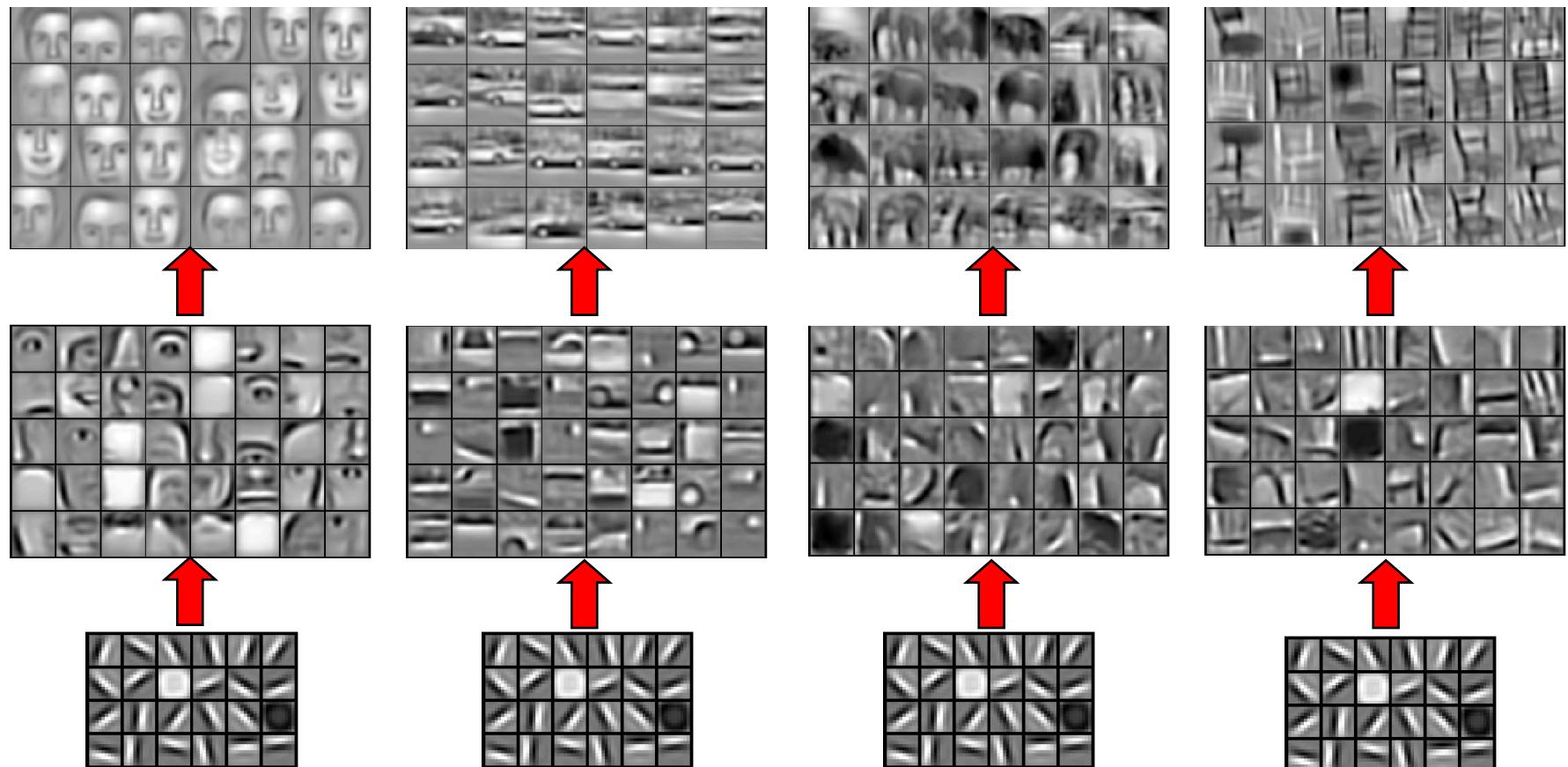
BROUGHT TO YOU BY:  

Deep Belief Net on Face Images

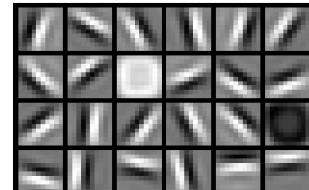
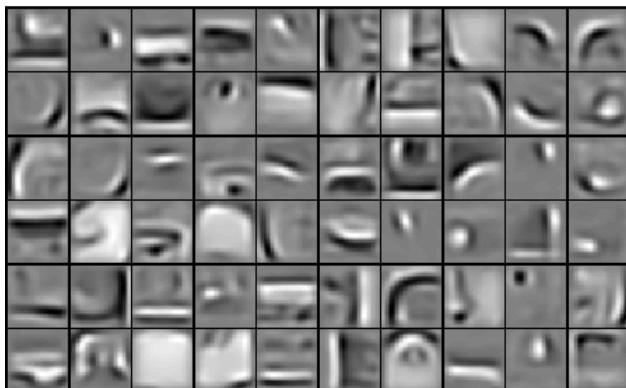


Based on materials
by Andrew Ng

Learning of Object Parts



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.

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

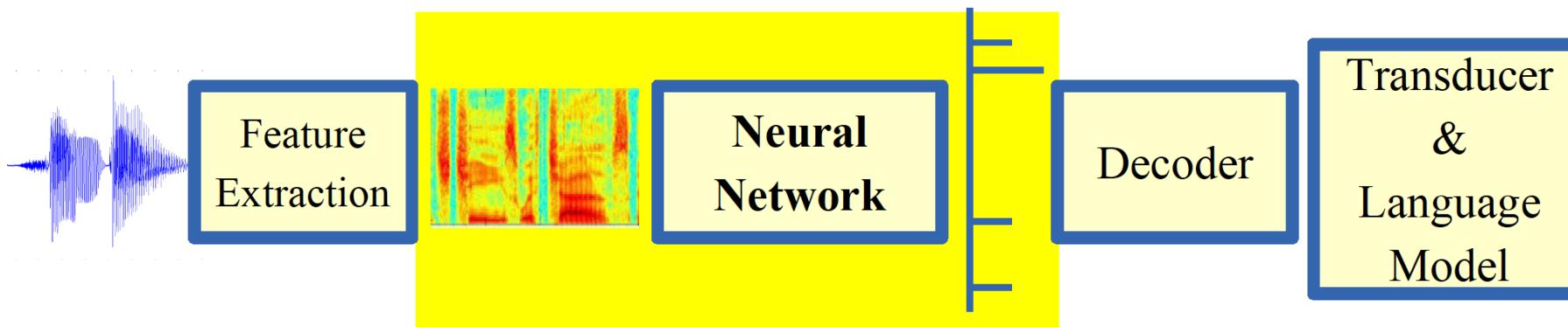


Samples from
feedforward
Inference
(control)

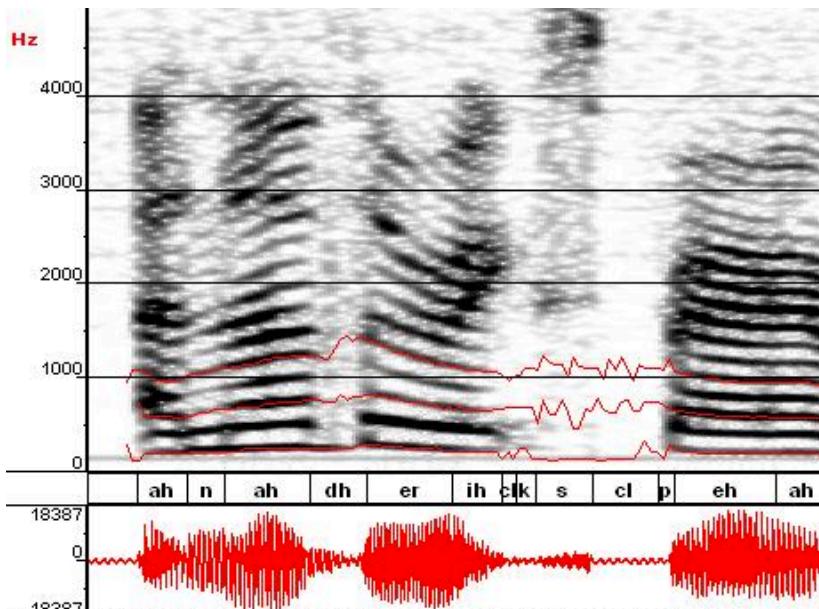
Samples from
Full posterior
inference

Machine Learning in Automatic Speech Recognition

A Typical Speech Recognition System



ML used to predict 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

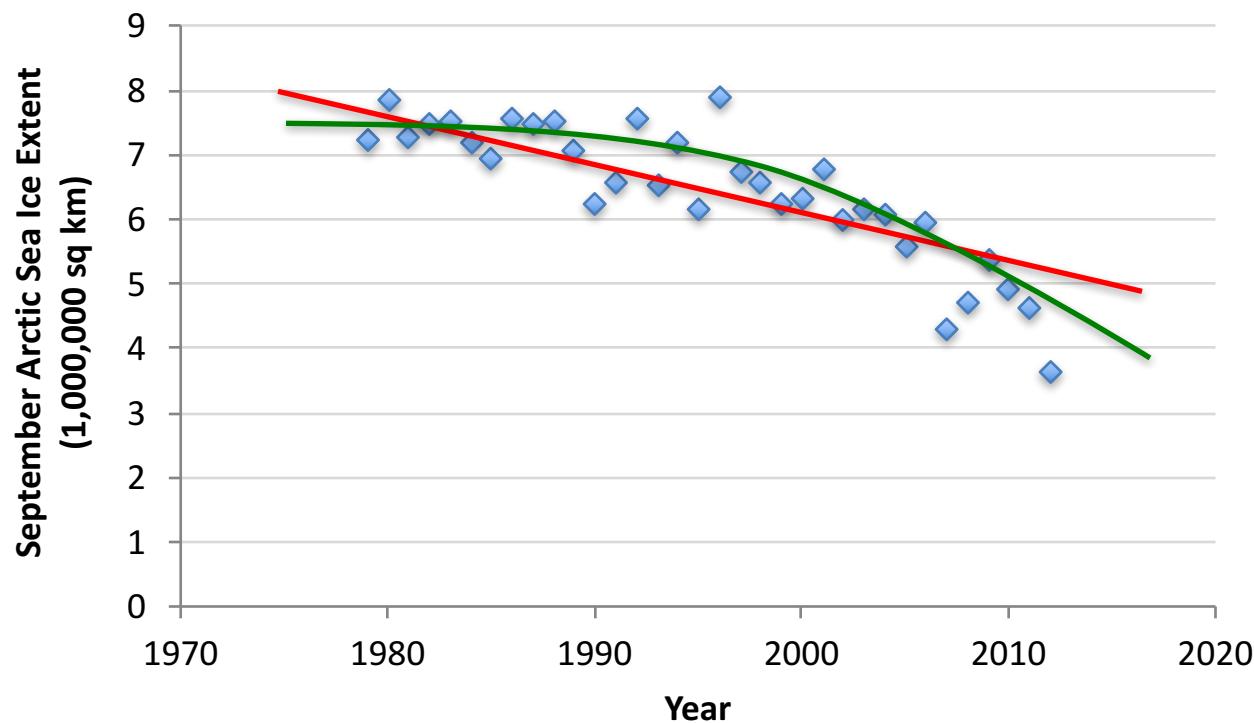
Types of Learning

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

Supervised Learning: Regression

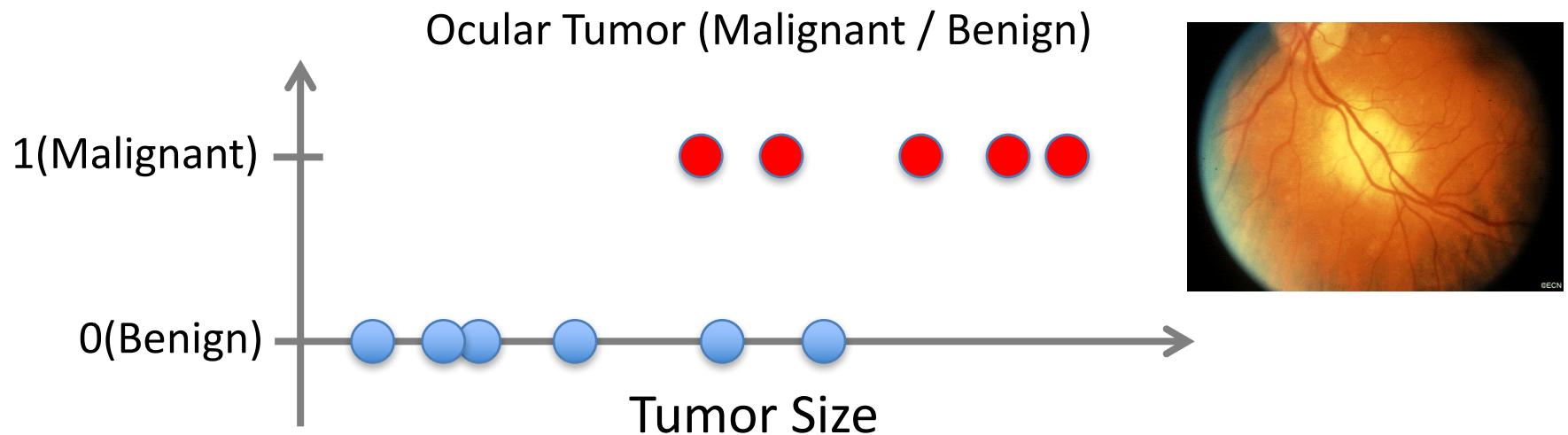
- Given $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
 - y is real-valued == regression



Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013)

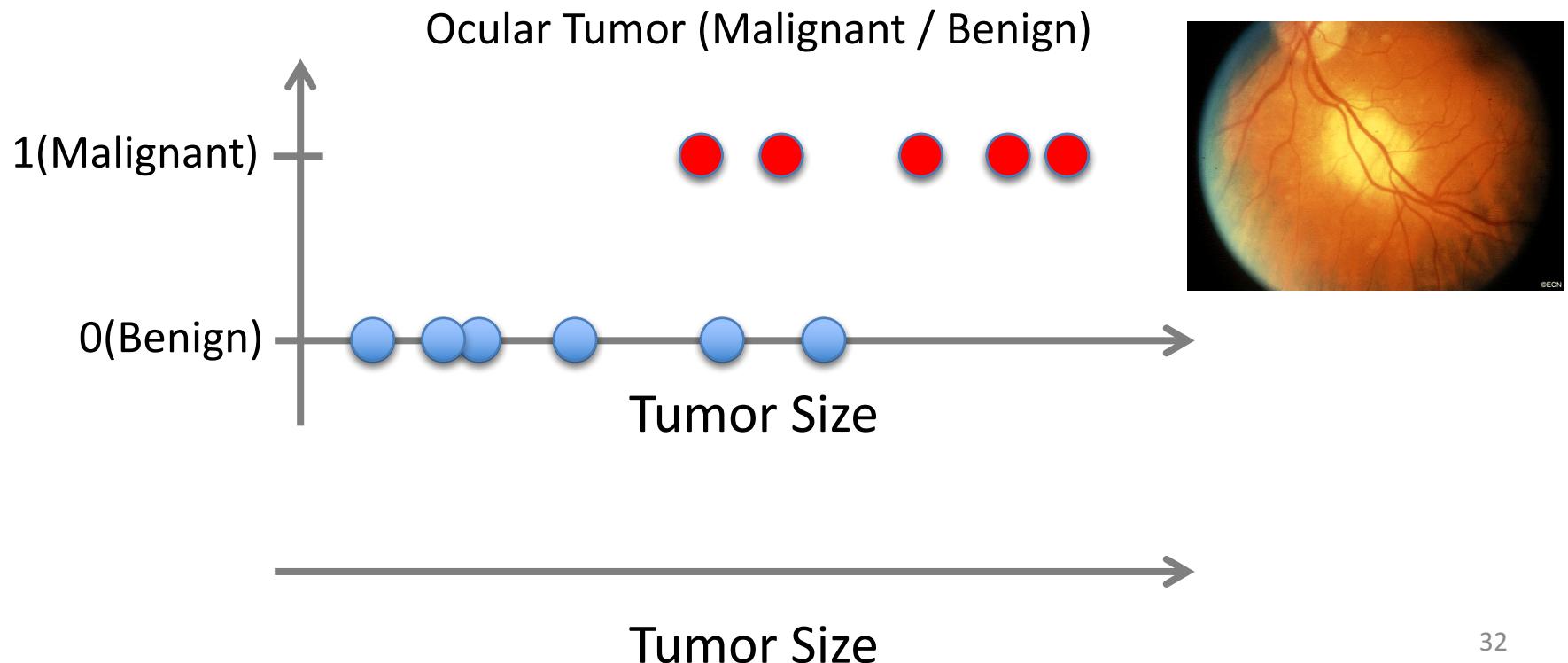
Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
 - y is categorical == classification



Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
 - y is categorical == classification



Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
 - y is categorical == classification

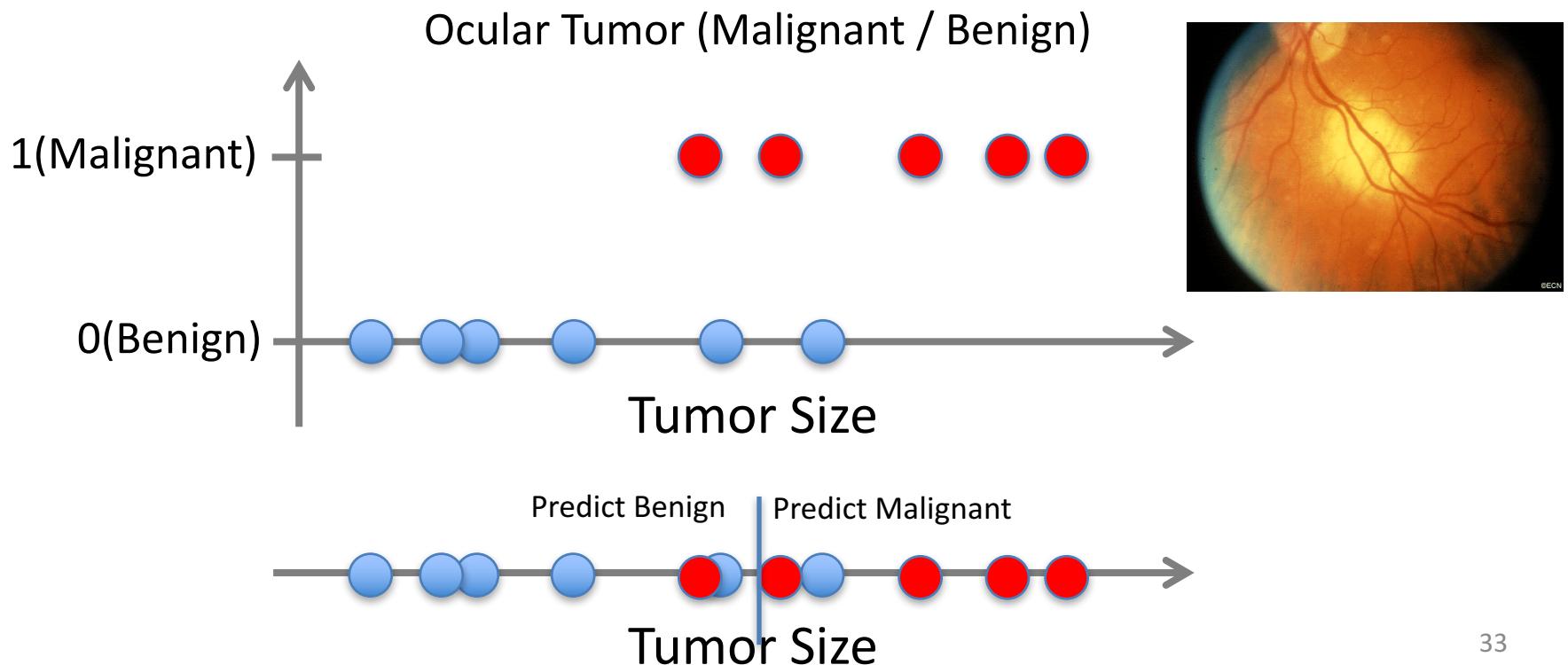
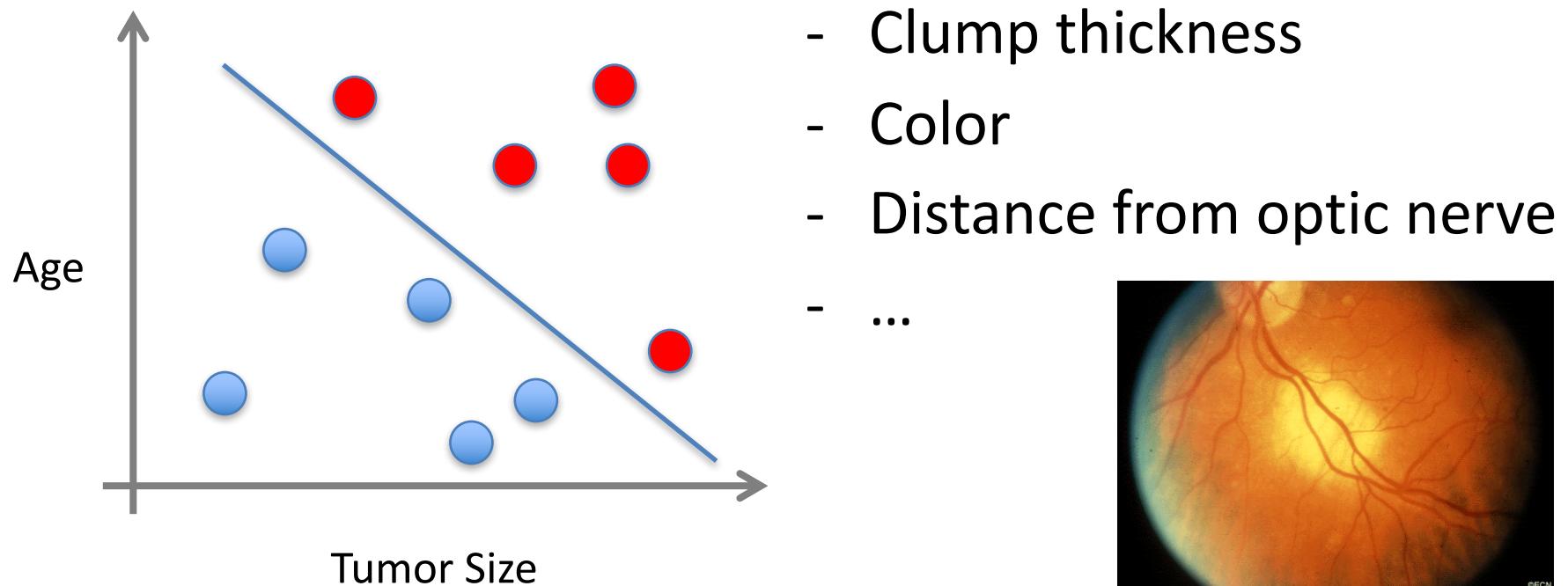


Image source: <https://eyecancer.com/uncategorized/choroidal-metastasis-test/>

Supervised Learning

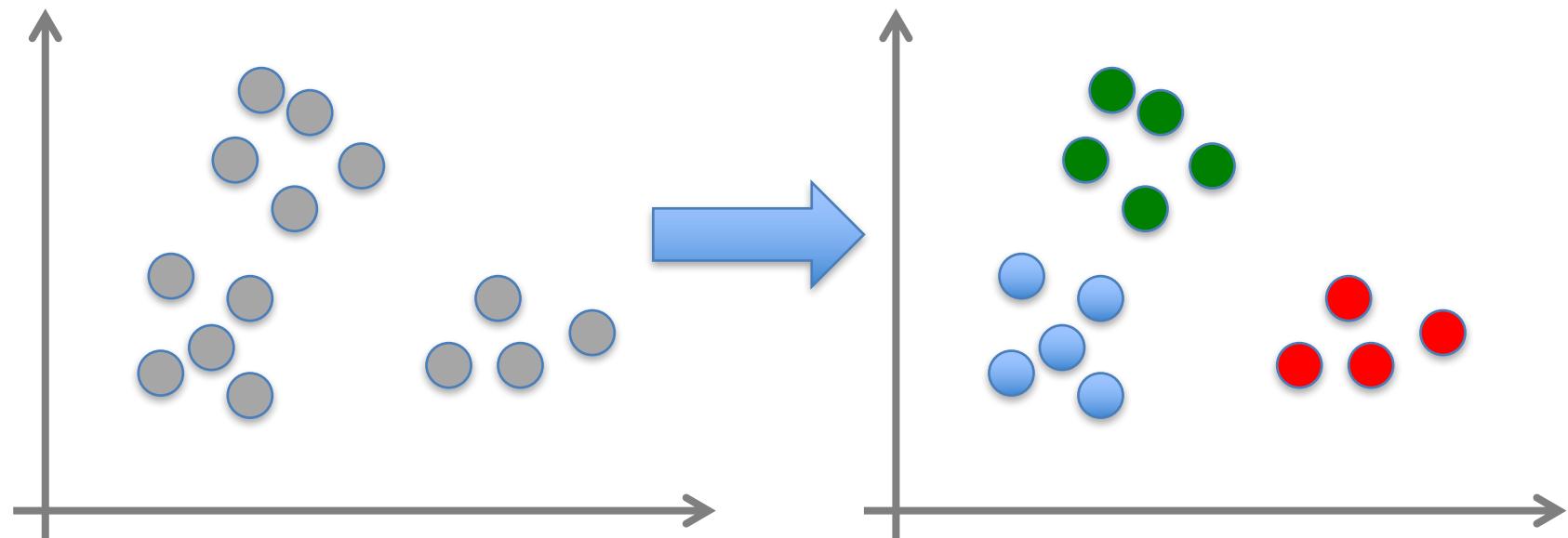
- x can be multi-dimensional
 - Each dimension corresponds to an attribute:



- Cell type is the most telling feature, but it's risky to do a biopsy of the eye
 - ML can help determine *when* a feature is needed

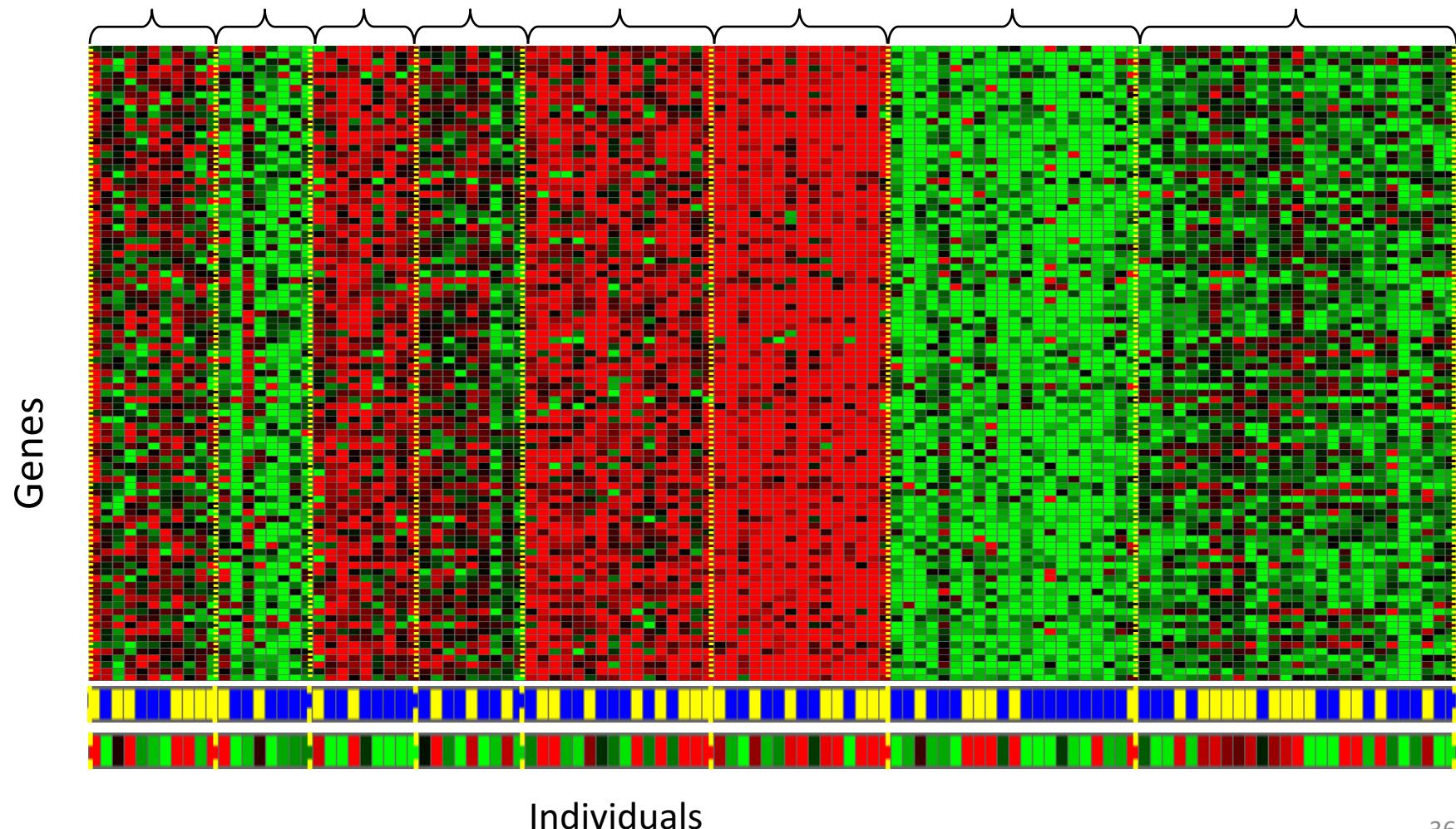
Unsupervised Learning

- Given x_1, x_2, \dots, x_n (without labels)
- Output hidden structure behind the x 's
 - E.g., clustering



Unsupervised Learning

Genomics application: group individuals by genetic similarity

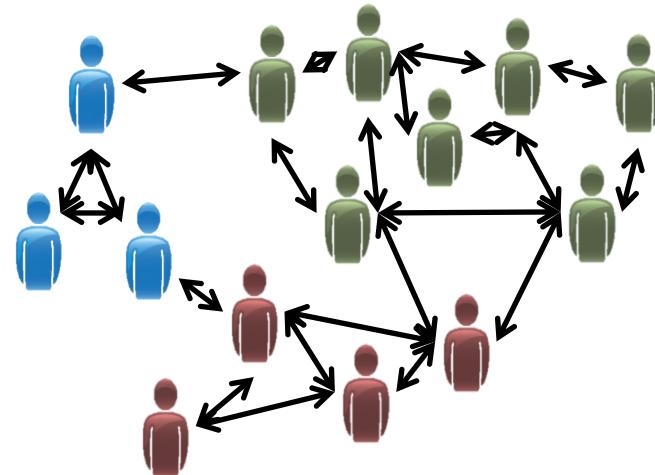


[Source: Daphne Koller]

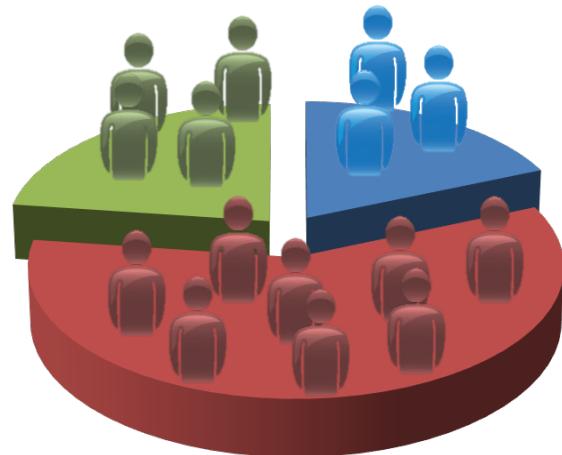
Unsupervised Learning



Organize computing clusters



Social network analysis



Market segmentation

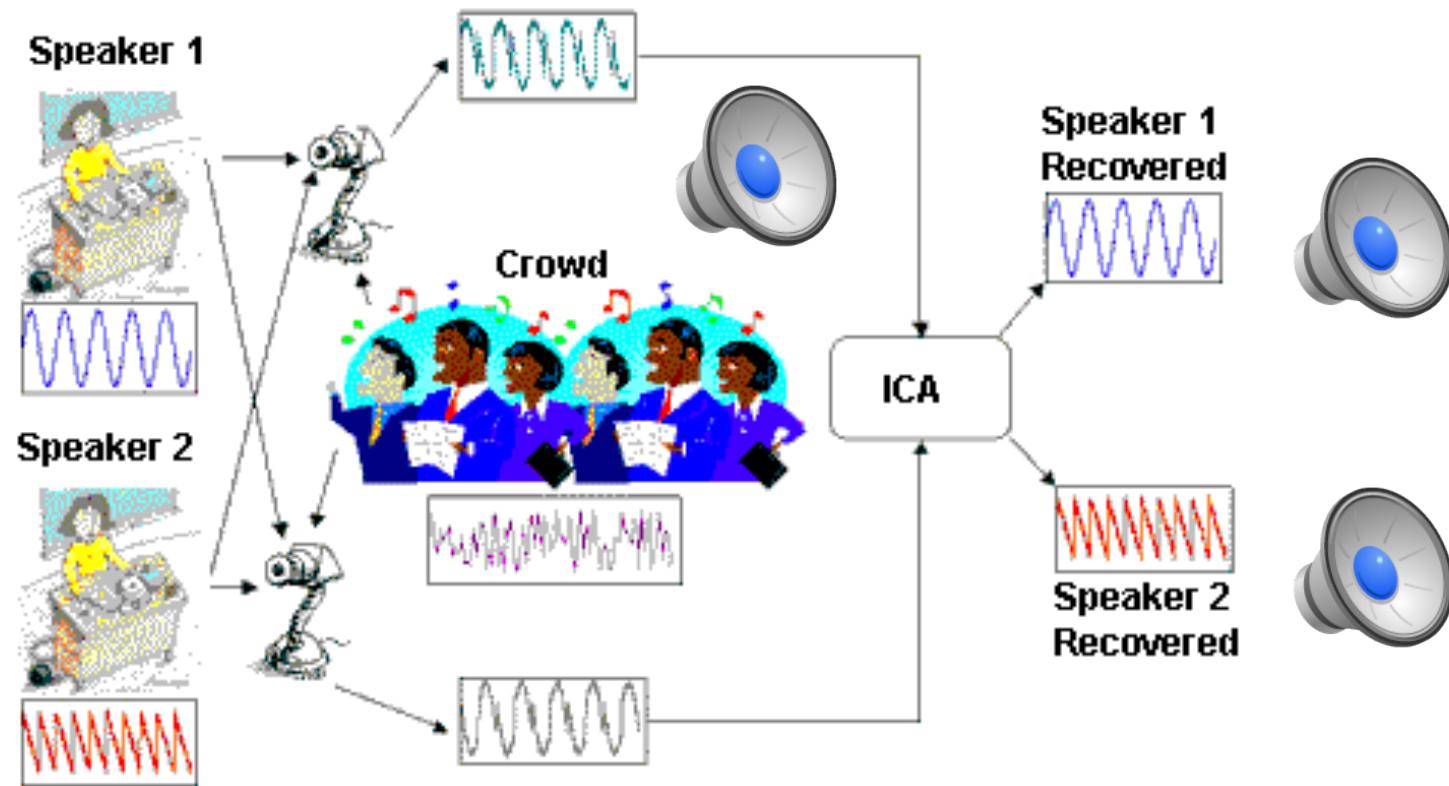


Image credit: NASA/JPL-Caltech/E. Churchwell (Univ. of Wisconsin, Madison)

Astronomical data analysis

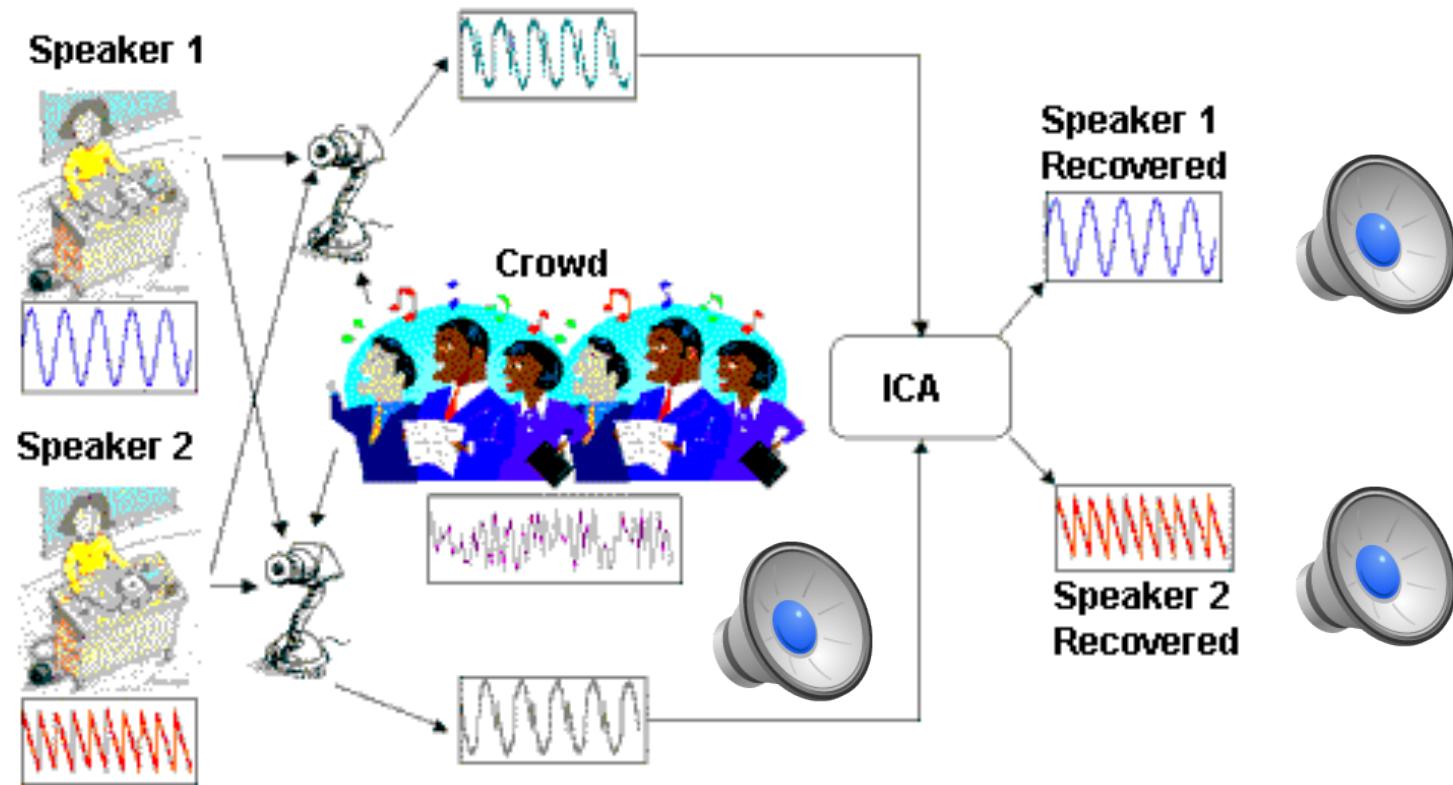
Unsupervised Learning

- Independent component analysis – separate a combined signal into its original sources



Unsupervised Learning

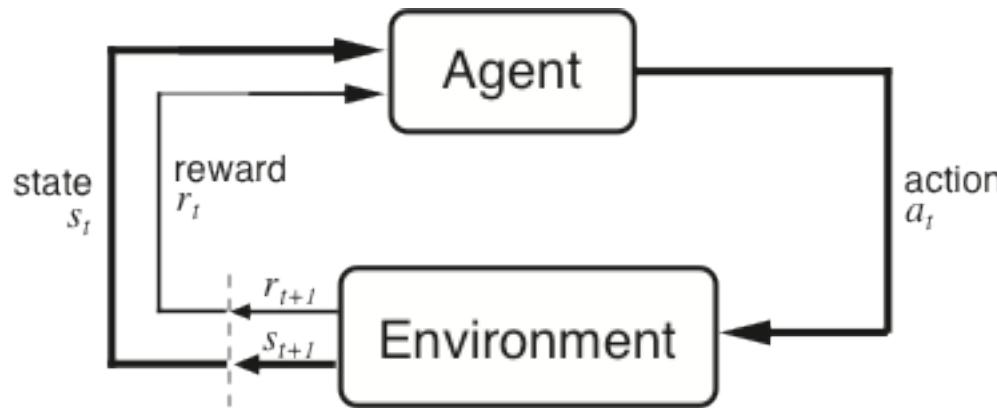
- Independent component analysis – separate a combined signal into its original sources



Reinforcement Learning

- Given a sequence of states and actions with (delayed) rewards, output a policy
 - Policy is a mapping from states → actions that tells you what to do in a given state
- Examples:
 - Credit assignment problem
 - Game playing
 - Robot in a maze
 - Balance a pole on your hand

The Agent-Environment Interface



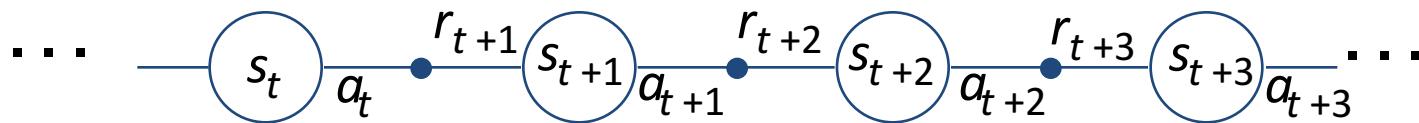
Agent and environment interact at discrete time steps : $t = 0, 1, 2, K$

Agent observes state at step t : $s_t \in S$

produces action at step t : $a_t \in A(s_t)$

gets resulting reward : $r_{t+1} \in \Re$

and resulting next state : s_{t+1}



Reinforcement Learning



<https://www.youtube.com/watch?v=4cgWya-wjgY>

Inverse Reinforcement Learning

- Learn policy from user demonstrations



Stanford Autonomous Helicopter

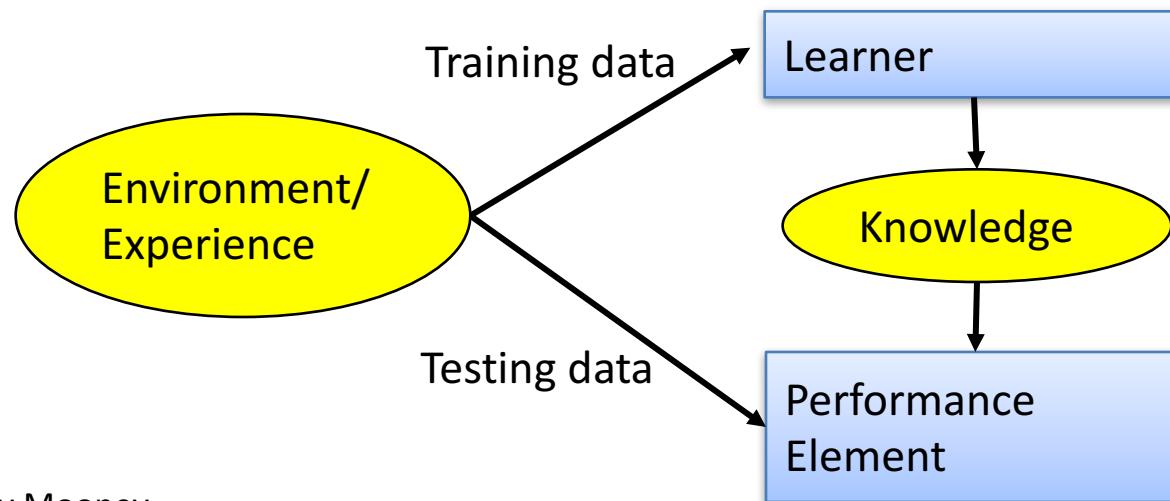
<http://heli.stanford.edu/>

<https://www.youtube.com/watch?v=VCdxqn0fcnE>

Framing a Learning Problem

Designing a Learning System

- Choose the training experience
- Choose exactly what is to be learned
 - i.e. the **target function**
- Choose how to represent the target function
- Choose a learning algorithm to infer the target function from the experience



Training vs. Test Distribution

- We generally assume that the training and test examples are independently drawn from the same overall distribution of data
 - We call this “i.i.d” which stands for “independent and identically distributed”
- If examples are not independent, requires ***collective classification***
- If test distribution is different, requires ***transfer learning***

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

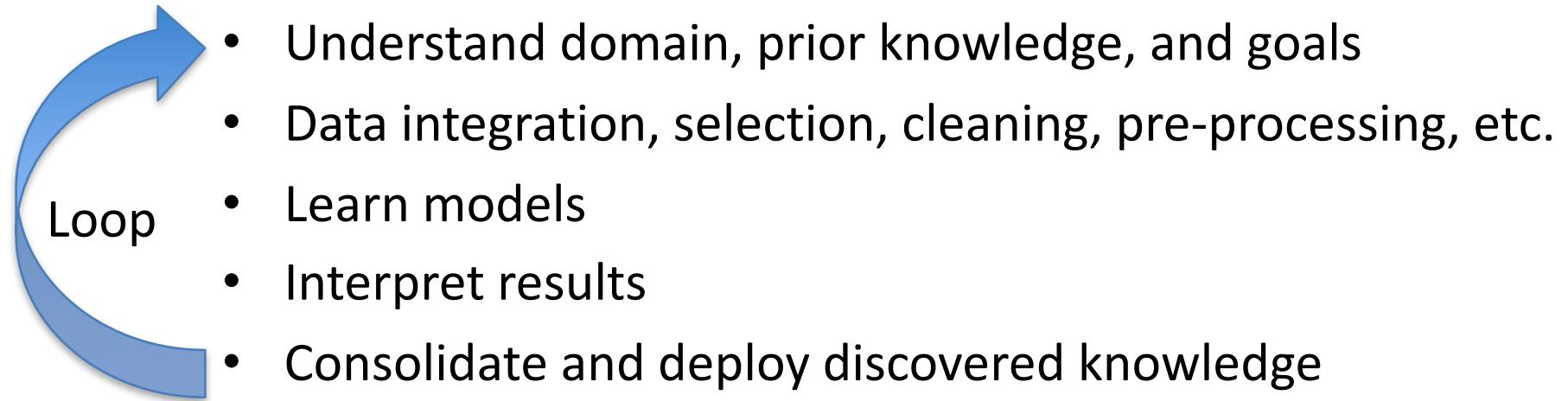
Various Search/Optimization Algorithms

- 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



Lessons Learned about Learning

- Learning can be viewed as using direct or indirect experience to approximate a chosen target function.
- Function approximation can be viewed as a search through a space of hypotheses (representations of functions) for one that best fits a set of training data.
- Different learning methods assume different hypothesis spaces (representation languages) and/or employ different search techniques.

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.
 - ???

What We'll Cover in this Course

- **Supervised learning**
 - Decision tree induction
 - Linear regression
 - Logistic regression
 - Support vector machines & kernel methods
 - Model ensembles
 - Bayesian learning
 - Neural networks & deep learning
 - Learning theory
- **Unsupervised learning**
 - Clustering
 - Dimensionality reduction
- **Reinforcement learning**
 - Temporal difference learning
 - Q learning
- **Evaluation**
- **Applications**

Our focus will be on applying machine learning to real applications