

EN.600.475 Machine Learning

Introduction to ML

Raman Arora
Lecture 1
January 30, 2017

- Introduction
- Course overview

Slides courtesy: Mark Dredze, Aarti Singh



WHAT IS LEARNING?



Bait Shyness vs Pigeon Superstition

Bait Shyness

- rats learning to avoid poisonous baits
- avoid new unseen food with similar taste or smell
- generalization: inductive reasoning or inductive inference



Bait Shyness vs Pigeon Superstition

Bait Shyness

- rats learning to avoid poisonous baits
- avoid new unseen food with similar taste or smell
- generalization: inductive reasoning or inductive inference

Pigeon superstition

- inductive reasoning may lead to false conclusions
- associate delivery of food with whatever chance actions they had been performing when it was first delivered
- <https://www.youtube.com/watch?v=NCtF4aVlxgU>

No Free Lunch

Inductive bias (prior knowledge)

- rats fail to acquire conditioning between food and electric shock
- rats learning biased towards detecting patterns while ignoring temporal correlations
- pigeons happy with any explanation for occurrence of food
- distinguishes mechanisms resulting in superstition from useful learning

Incorporating prior knowledge or biasing learning is inevitable for the success of learning algorithms

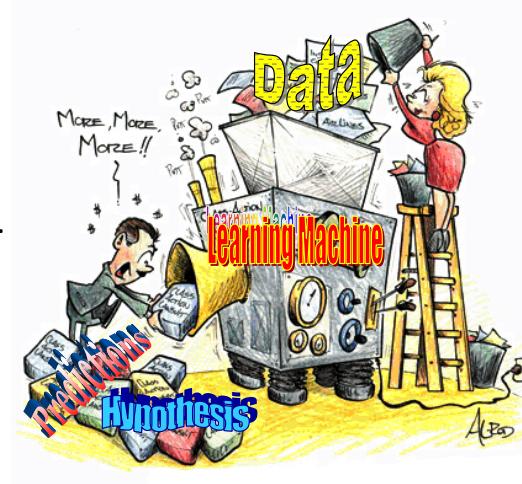
7

Using Experience to gain Expertise

Machine Learning

Design of programs that can

- learn “rules” from data,
- adapt to changes, and
- improve performance with experience.

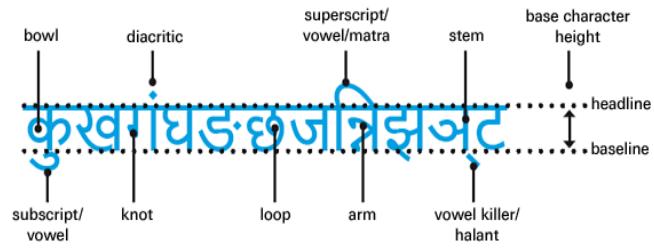


9

What is Machine Learning?

- Let's say you want to solve Character Recognition
- Hard way: Understand handwriting/characters

Aa Bb Cc Dd
Ee Ff Gg Hh
Ii Jj Kk Ll
Mm Nn Oo Pp
Qq Rr Ss Tt
Uu Vv Ww Xx
Yy Zz



What is Machine Learning?

- Let's say you want to solve Character Recognition
- Hard way: Understand handwriting/characters
 - Latin
 - Devanagri
 - Symbols: <http://detexify.kirelabs.org/classify.html>



Slide Credit: Dhruv Batra

11

What is Machine Learning?

- Let's say you want to solve Character Recognition
- Hard way: Understand handwriting/characters
- Lazy way: Throw data!

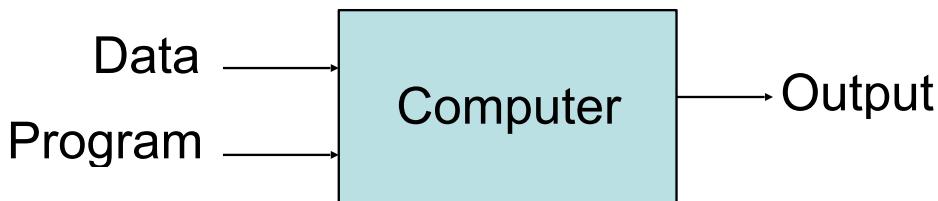


Slide Credit: Dhruv Batra

12

What is Machine Learning?

- Traditional Programming



- Machine Learning



(C) Dhruv Batra

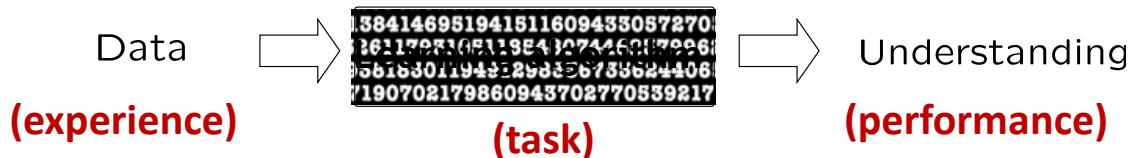
Slide Credit: Pedro Domingos, Tom Mitchel, Tom Dietterich

13

What is Machine Learning?

Study of algorithms that

- improve their performance
- at some task
- with experience



14

When/Why do we need machine learning?

Tasks that are too complex to program

- Tasks performed by humans (e.g. driving, understanding speech/language, understanding images/visual scenes)
- Tasks beyond human capabilities (analyzing large complex datasets e.g. astronomical data, turning medical archives into medical knowledge, weather prediction, analysis of genomic data, Web search engines, electronic commerce).

Adaptivity

- Programmed systems are limited by their rigidity. Many tasks change from time to time or from user to user.

15

When/Why do we need machine learning?

Types of learners

- A sponge, which absorbs everything
- A funnel, that lets in at one end and discharges at the other
- A sieve, that forgets the essential but retains the unimportant!
- A strainer, that memorizes the good and rejects the worthless



A computer can be an excellent ``sponge'' learner.

16

Why Machine Learning? System Design

- **Objective:** Automatically set up a complex system rather than design it manually
- **Benefits** of this approach:
 - save human (often expert) labor
 - may do better than human (even expert?) designer
- Some **examples** from computer science:
 - Branch prediction (binary classification)
 - Choice of optimization level (multi-class)
 - Predicting page faults and cache misses (detection)
 - Expected CPU or user I/O time (regression)

17

Slide Credit: Ray Mooney

Why Machine Learning? AI

- **Objective:** build intelligent systems
- What is intelligence?
- Many definitions almost equate intelligence with learning
- “Strong AI” will likely require learning
- Learning does not require AI
focus of most modern/recent ML research is pragmatic: emulate components of an intelligent system (perception, planning, natural language interfaces, . . .), not general intelligent behavior

18

Slide Credit: Ray Mooney

Why Machine Learning? Cognitive Science

- Computational studies of learning may help us understand learning in humans
 - and other biological organisms.
 - Hebbian neural learning: “Neurons that fire together, wire together.”

19

Slide Credit: Ray Mooney

Why Machine Learning? The Time is Ripe

- Algorithms
 - Many basic effective and efficient algorithms available.
- Data
 - Large amounts of on-line data available.
- Computing
 - Large amounts of computational resources available.

20

Slide Credit: Ray Mooney

Why Machine Learning?

The Time is Ripe

- “If you were a current computer science student what area would you start studying heavily?”
 - Answer: Machine Learning.
 - “The ultimate is computers that learn”
 - Bill Gates, Reddit AMA
- “Machine learning is the next Internet”
 - Tony Tether, Director, DARPA
- “Machine learning is today’s discontinuity”
 - Jerry Yang, CEO, Yahoo

21

Slide Credit: Ray Mooney

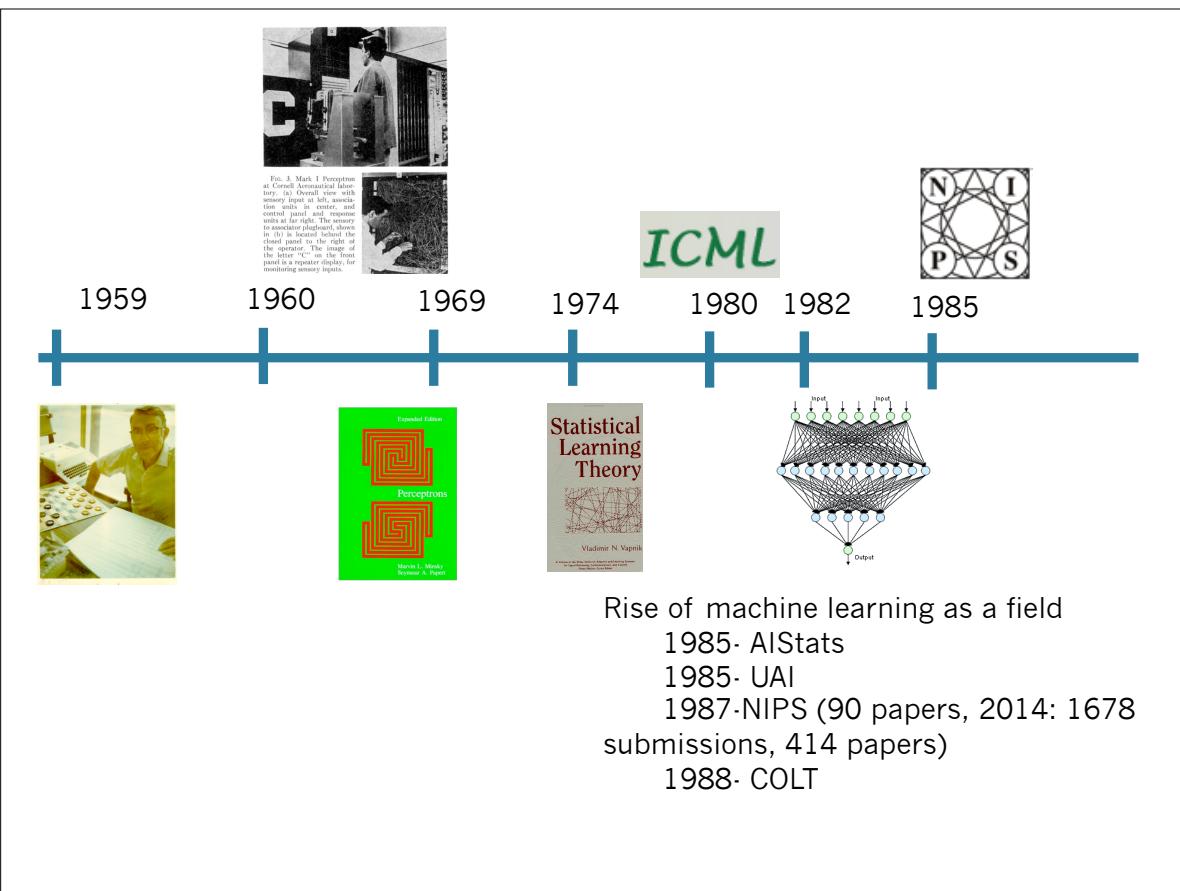
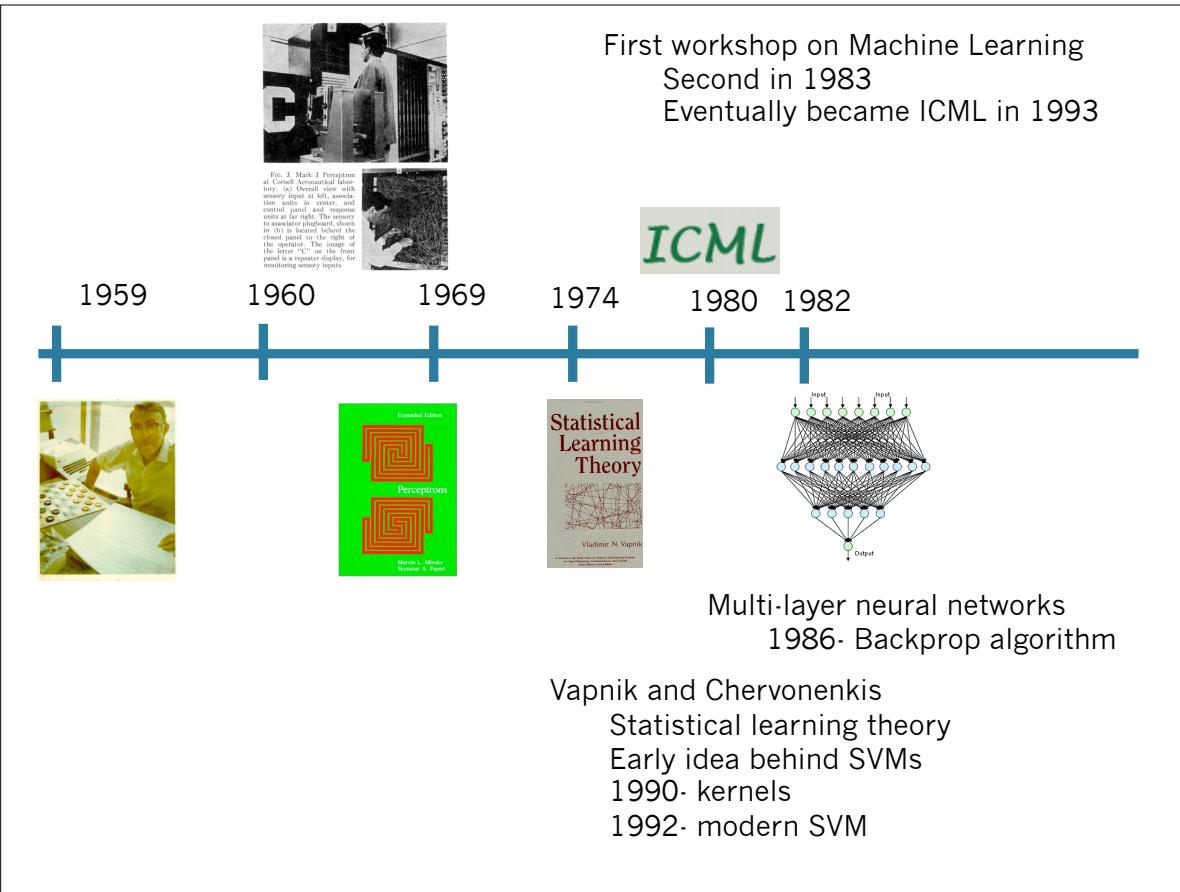
A horizontal timeline with three vertical tick marks. The first tick mark is labeled "1959" below it. The second tick mark is labeled "1960" below it. The third tick mark is labeled "1969" below it. To the left of the first tick mark is a small portrait of Arthur Samuel (IBM). Between the first and second tick marks is a photograph of Frank Rosenblatt at Cornell. Between the second and third tick marks is a book cover for Marvin Minsky and Seymour Papert's "Perceptrons". To the right of the timeline is a block of text about Frank Rosenblatt's work on the Perceptron.

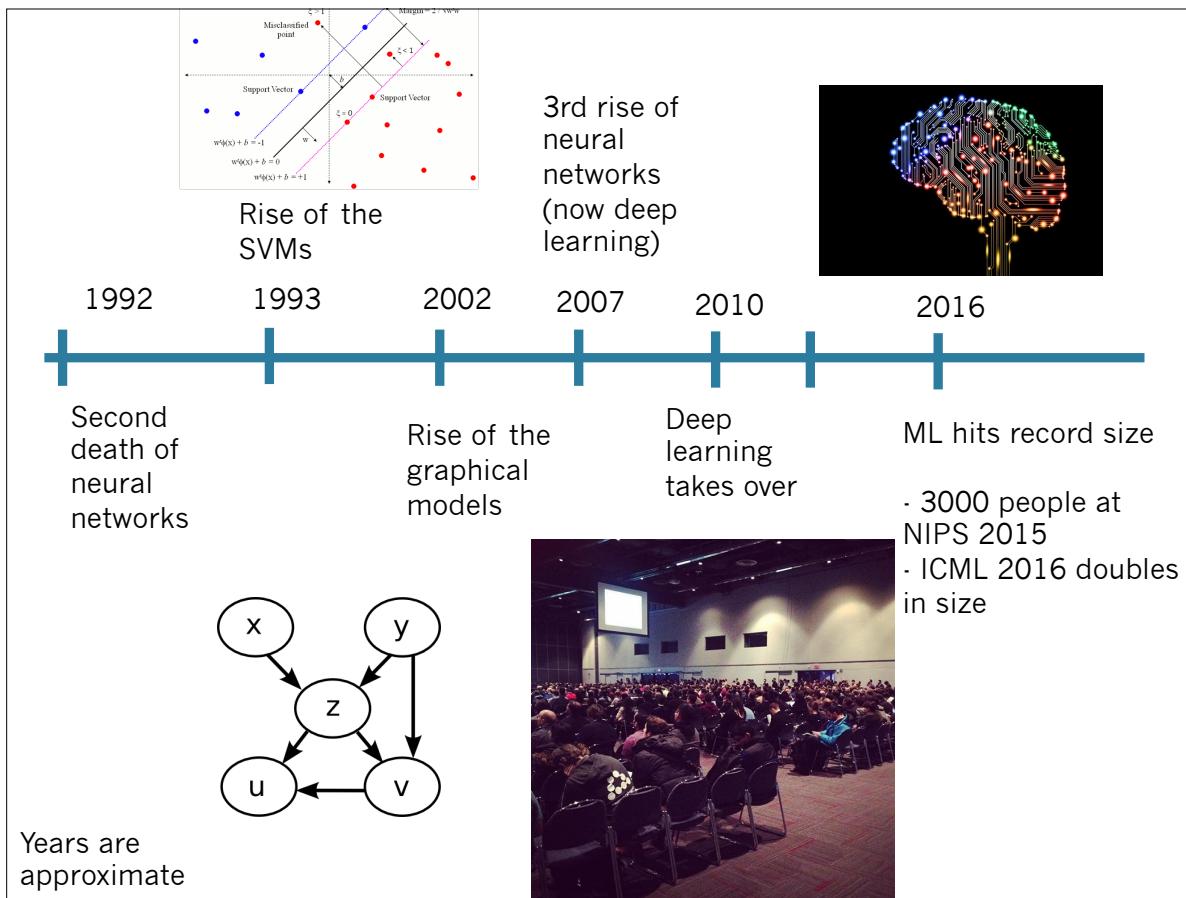
Frank Rosenblatt (Cornell)
Linear classifier for character recognition
Based on neurons in the brain
By studying the Perceptron, “the fundamental laws of organization which are common to all information handling systems, machines and men included, may eventually be understood.”

Arthur Samuel (IBM)
Checkers playing program
Rote learning- simulate many games to learn strategy
Early reinforcement learning

Marvin Minsky, Seymour Papert (MIT)
“Perceptrons: An Introduction to Computational Geometry”
Criticism of Perceptron demonstrating limitations
Research in Perceptrons stops for almost 20 years

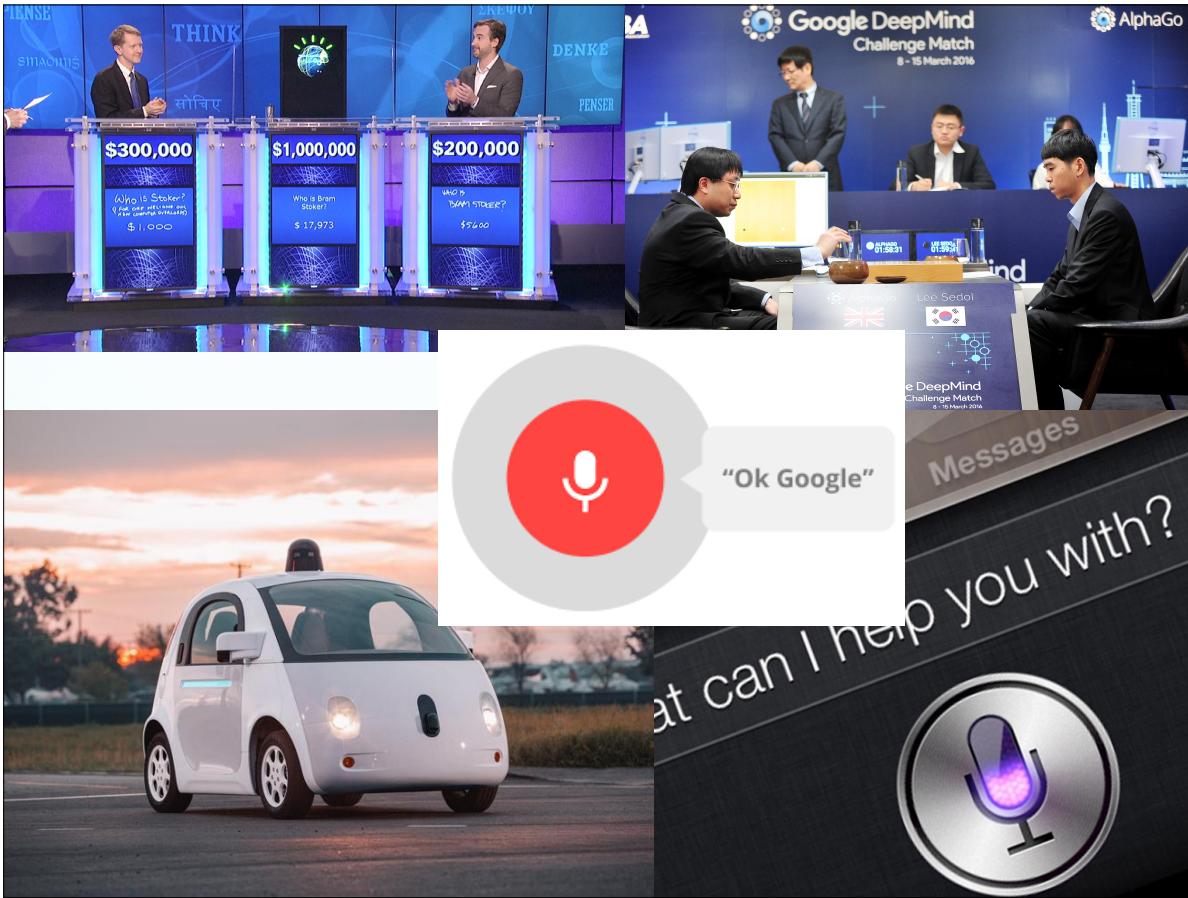
22





Machine Learning Today

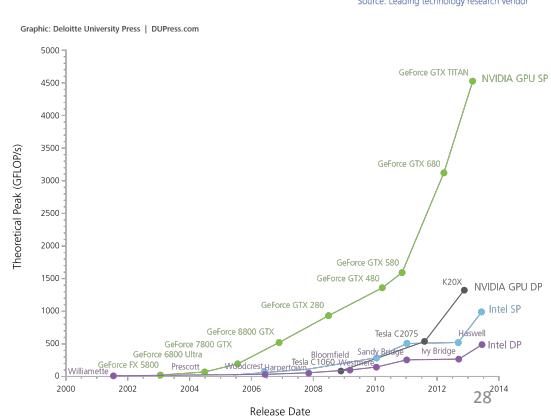
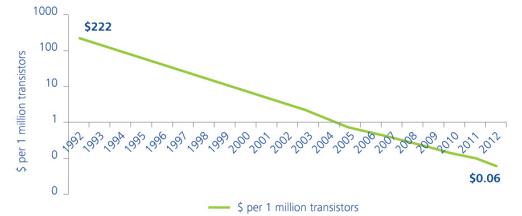
- Machine Learning
- Artificial intelligence
 - ML split from AI ~20 years ago, coming back
- Why?



Why success?

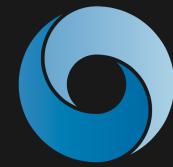
- (Somewhat) Better algorithms
 - Deep learning
- Better computers
- More data
- The rise of data science
 - More companies hire ML, more success (cycles)

Figure 1. Computing cost-performance (1992–2012)





ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE



Google DeepMind



OpenAI

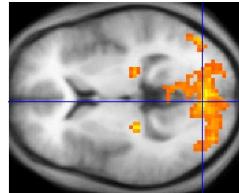
Stanford | One Hundred Year Study on Artificial Intelligence (AI100)

From Data to Understanding ...

Machine Learning in Action

Machine Learning in Action

- Decoding thoughts from brain scans



Rob a bank ...

Home » Health & Wellness

Brain Scans: Are You a Criminal?

Published February 07, 2007 by:
Andrea Okrentowich
[View Profile](#) | [Follow](#) | [Add to Favorites](#)

More: [Brain Scans](#) | [Brain Scan](#) | [Disposition](#) | [Defendant](#) | [Criminal Behavior](#)

MRI Scans as Courtroom Evidence

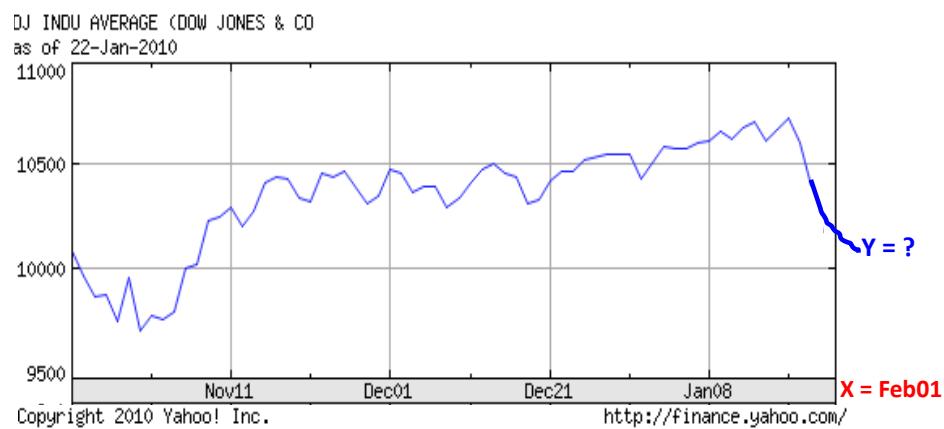
The average Joe's MRI scan can show a brain abnormality, do we proceed to check him into the nearest mental institution or prison? That would make about as much sense as trying to prove a defendant innocent of a violent



31

Machine Learning in Action

- Stock Market Prediction



Machine Learning in Action

- Document classification



33

Machine Learning in Action

- Spam filtering

Welcome to New Media Installation: Art that Learns

Hi everyone,
Welcome to New Media Installation:Art that Learns
The class will start tomorrow.
Make sure you attend the first class, even if you are on the Wait List.
The classes are held in Doherty Hall C316, and will be Tue, Thu 01:30-4:20 PM.
By now, you should be subscribed to our course mailing list: 10615-announce@cs.cmu.edu.

Natural _LoseWeight SuperFood Endorsed by Oprah Winfrey, Free Trial 1 bottle,
pay only \$5.95 for shipping mfw rlk [Spam](#) | [X](#)

Spam/
Not spam

==== Natural WeightLOSS Solution ====

Vital Acai is a natural WeightLOSS product that Enables people to lose weight and cleanse their bodies faster than most other products on the market.

Here are some of the benefits of Vital Acai that You might not be aware of. These benefits have helped people who have been using Vital Acai daily to Achieve goals and reach new heights in their dieting that they never thought they could.

- * Rapid WeightLOSS
- * Increased metabolism - BurnFat & calories easily!
- * Better Mood and Attitude

34

Machine Learning in Action

- Cars navigating on their own



Boss, the self-driving SUV
1st place in the DARPA Urban
Challenge.

Photo courtesy of Tartan Racing.



35

Machine Learning in Action

- The **best** helicopter pilot is now a computer!
 - it runs a program that learns how to fly and make acrobatic maneuvers by itself!
 - no taped instructions, joysticks, or things like that ...



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

36

Machine Learning in Action

- Robot assistant?
[<http://stair.stanford.edu/>]



37

Machine Learning in Action

- Many, many more...
Speech recognition, Natural language processing
Computer vision
Web forensics
Medical outcomes analysis
Computational biology
Sensor networks
Social networks
...

38

ML is trending!

- Wide applicability
- Very large-scale complex systems
 - Internet (billions of nodes), sensor network (new multi-modal sensing devices), genetics (human genome)
- Huge multi-dimensional data sets
 - 30,000 genes x 10,000 drugs x 100 species x ...
- Software too complex to write by hand
- Improved machine learning algorithms
- Improved data capture (Terabytes, Petabytes of data), networking, faster computers
- Demand for self-customization to user, environment

39

Glasses for the Visually Impaired

http://www.youtube.com/watch?v=k_C1iKIqi_o

Under the hood: Learning Problems

- Classification (assign a category: pedestrian, traffic, etc)
- Regression (predict a real value: depth in a scene)
- Ranking (order items: information extraction)
- Clustering (partitioning items into homogeneous regions: image segmentation)
- Dimensionality reduction/manifold learning (learn a lower dimensional representation: computer vision)

Machine Learning Paradigms

Broad categories -

- **Supervised learning**

Classification, Regression

- **Unsupervised learning**

Density estimation, Clustering, Dimensionality reduction

- Semi-supervised learning
- Active learning
- Reinforcement learning
- Many more ...

Supervised Learning

Feature Space \mathcal{X}



Words in a document

Label Space \mathcal{Y}

“Sports”
“News”
“Science”

...



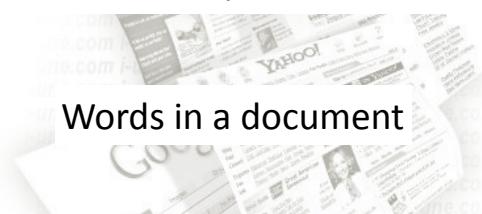
Share Price
“\$ 24.50”

Task: Given $X \in \mathcal{X}$, predict $Y \in \mathcal{Y}$.

43

Supervised Learning - Classification

Feature Space \mathcal{X}

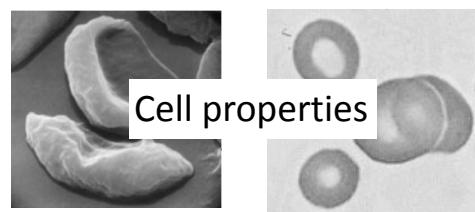


Words in a document

Label Space \mathcal{Y}

“Sports”
“News”
“Science”

...



Discrete Labels



“Anemic cell”
“Healthy cell”

44

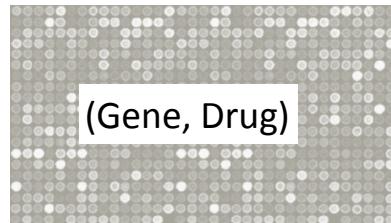
Supervised Learning - Regression

Feature Space \mathcal{X}



Label Space \mathcal{Y}

Share Price
"\$ 24.50"



Expression level
"0.01"

Continuous Labels

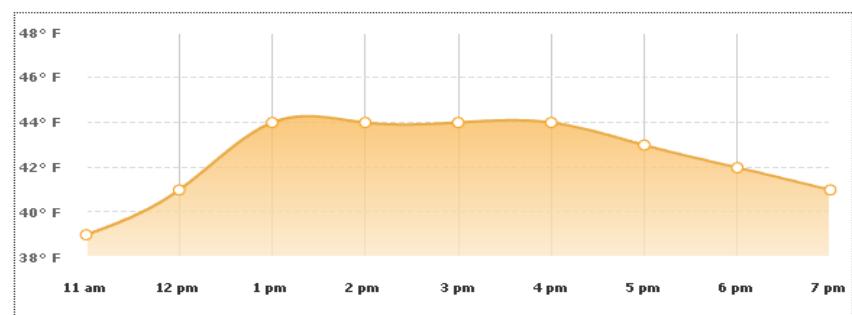
45

Supervised Learning problems

Features?

Labels?

Classification/Regression?



Temperature/Weather prediction

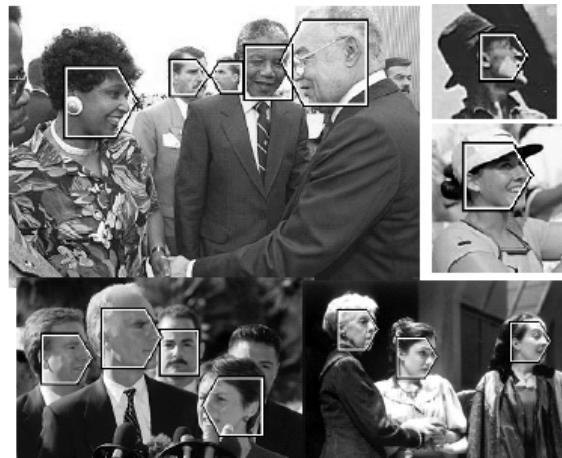
46

Supervised Learning problems

Features?

Labels?

Classification/Regression?



Face Detection

47

Supervised Learning problems

Features?

Labels?

Classification/Regression?



Environmental Mapping

48

Supervised Learning problems

Features?

Labels?

Classification/Regression?



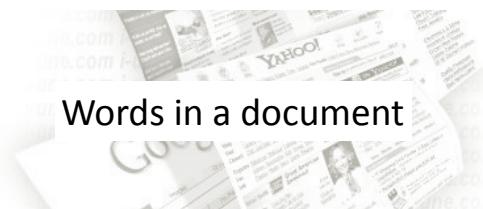
Robotic Control

49

Unsupervised Learning

Aka “learning without a teacher”

Feature Space \mathcal{X}



Words in a document



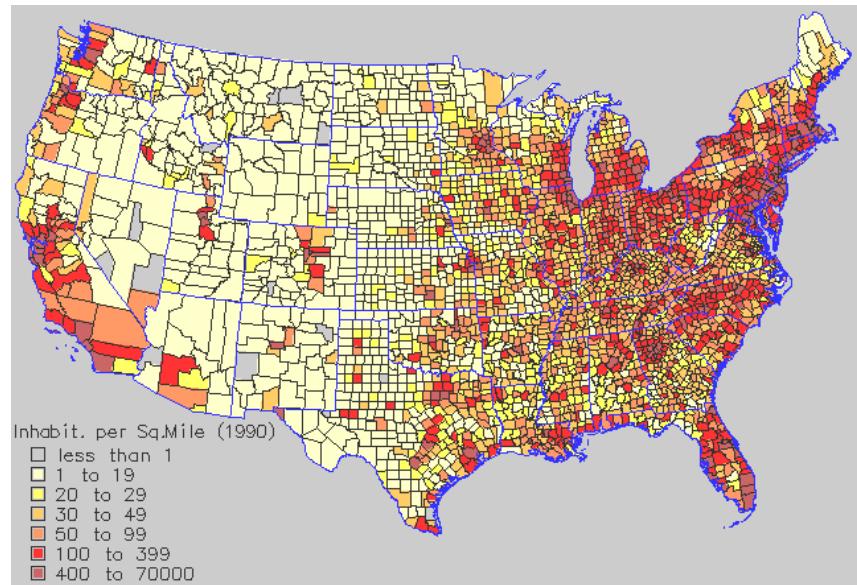
Word distribution
(Probability of a word)

Task: Given $X \in \mathcal{X}$, learn $f(X)$.

50

Unsupervised Learning – Density Estimation

Population density



51

Unsupervised Learning – clustering

Group similar things e.g. images

[Goldberger et al.]



52

Unsupervised Learning – clustering web search results

Clusty

web news images wikipedia blogs jobs more » race Search advanced preferences

Cluster Human contains 8 documents.

All Results (238)

- Car (28)
 - Race cars (1)
 - Photos, Races Scheduled (5)
 - Game (4)
 - Track (3)
 - Nascar (2)
 - Equipment And Safety (2)
 - Other Topics (7)
 - Photos (22)
 - Game (14)
 - Definition (13)
 - Team (18)
 - Human (8)
 - Classification Of Human (2)
 - Statement, Evolved (2)
 - Other Topics (4)
 - Weekend (8)
 - Ethnicity And Race (1)
 - Race for the Cure (8)

1. [Race \(classification of human beings\) - Wikipedia, the free ...](#) ⓘ ⓘ ⓘ
The term **race** or racial group usually refers to the concept of dividing **humans** into populations or groups on the basis of visible traits (especially skin color, cranial or facial features and hair texture), and self-identity by culture and over time, and are often controversial for scientific as well as social and political reasons. History · Mc en.wikipedia.org/wiki/Race_(classification_of_human_beings) - [cache] - Live, Ask

2. [Race - Wikipedia, the free encyclopedia](#) ⓘ ⓘ ⓘ
General. Racing competitions The **Race** (yachting **race**), or La course du millénaire, a no-rules round-the-world sail of **human** beings) **Race** and ethnicity in the United States Census, official definitions of "race" used by the US Census. Historical definitions of **race**; **Race** (bearing), the inner and outer rings of a rolling-element bearing. **RACE** Literature · Video games en.wikipedia.org/wiki/Race - [cache] - Live, Ask

3. [Publications | Human Rights Watch](#) ⓘ ⓘ ⓘ
The use of torture, unlawful rendition, secret prisons, unfair trials, ... Risks to Migrants, Refugees, and Asylum Seekers ... www.hrw.org/backgrounder/usa/race - [cache] - Ask

4. [Amazon.com: Race: The Reality Of Human Differences: Vincent Sarich ...](#) ⓘ ⓘ ⓘ
Amazon.com: **Race**: The Reality Of Human Differences: Vincent Sarich, Frank Miele: Books ... From Publishers Weekly www.amazon.com/Race-Reality-Differences-Vincent-Sarich/dp/0813340861 - [cache] - Live

5. [AAPA Statement on Biological Aspects of Race](#) ⓘ ⓘ ⓘ
AAPA Statement on Biological Aspects of **Race** ... Published in the American Journal of Physical Anthropology, vol. evolution and variation, ... www.physanth.org/positions/race.html - [cache] - Ask

6. [race: Definition from Answers.com](#) ⓘ ⓘ ⓘ
race n. A local geographic or global **human** population distinguished as a more or less distinct group by genetically www.answers.com/topic/race-1 - [cache] - Live

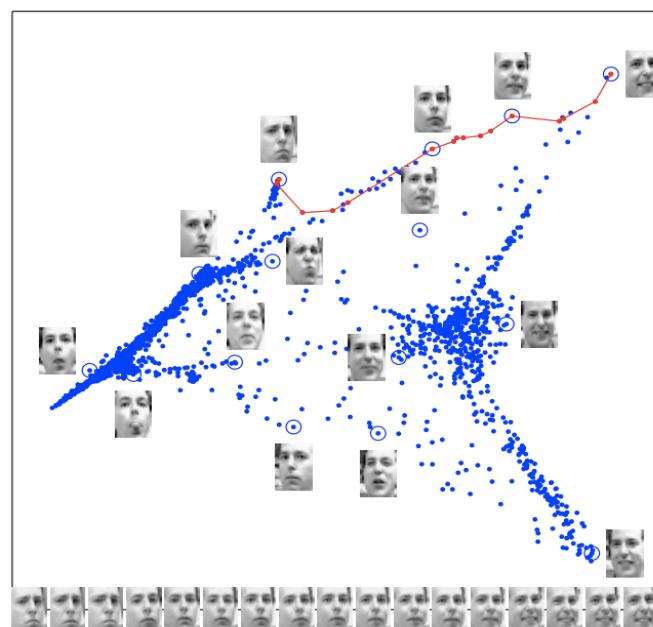
Unsupervised Learning - Embedding

Dimensionality Reduction

[Saul & Roweis '03]

Images have thousands or millions of pixels.

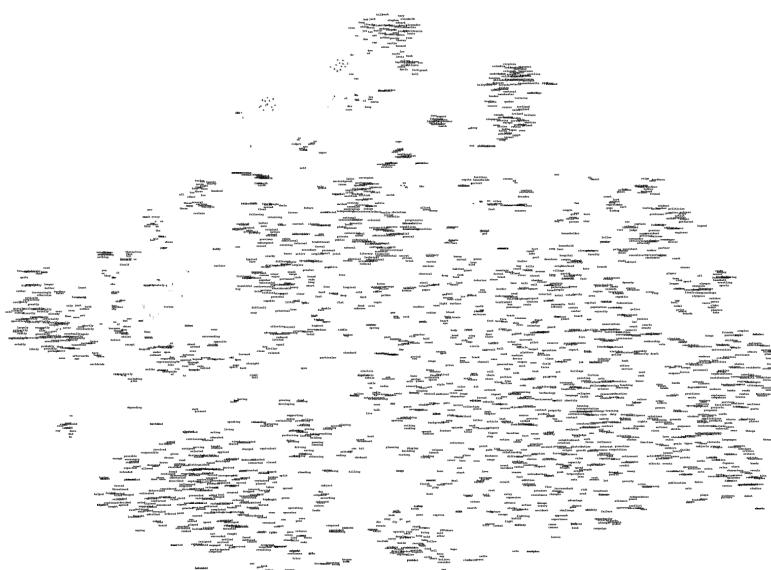
Can we give each image a coordinate, such that similar images are near each other?



Unsupervised Learning - Embedding

Dimensionality Reduction - words

[Joseph Turian]



55

Unsupervised Learning - Embedding

Machine Learning Paradigms

Broad categories -

- **Supervised learning**

Classification, Regression

- **Unsupervised learning**

Density estimation, Clustering, Dimensionality reduction

- Semi-supervised learning

- Active learning

- Reinforcement learning

- Many more ...

57

About the course

- Machine Learning Algorithms and Principles

- Classification: kNN, naive Bayes, logistic regression, perceptrons, large margin, kernels, support vector machines

- Regression: Linear regression, Gaussian Process regression, Kernel regression

- Kernel density estimation, Hidden Markov Models, Graphical Models, k-means clustering, dimensionality reduction, neural networks, deep belief networks, Boosting, Decision Trees, etc.

- Optimization, Theory, Model selection, overfitting, bias-variance tradeoffs ...

- **It's going to be fun and hard work 😊**

- See **tentative** lecture outline on Piazza – MAY CHANGE

- Material: Class slides + Reading material

58