CMSC 726
Lecture 1: Introduction

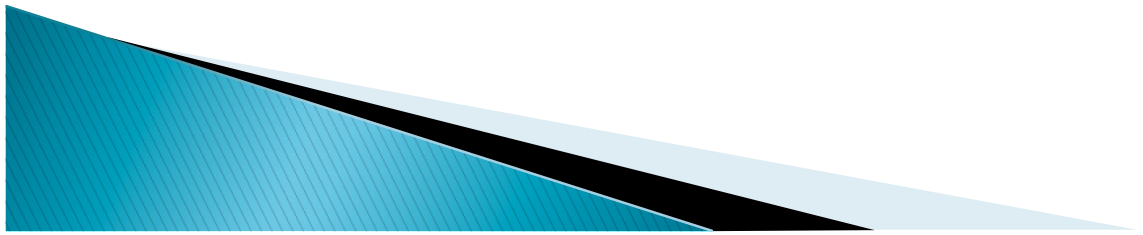
Lise Getoor

August 31, 2010

ACKNOWLEDGEMENTS: The material in this course is a synthesis of materials from many sources, including: Hal Daume III, Mark Drezde, Carlos Guestrin, Andrew Ng, Ben Taskar, Eric Xing, and others. I am very grateful for their generous sharing of insights and materials.

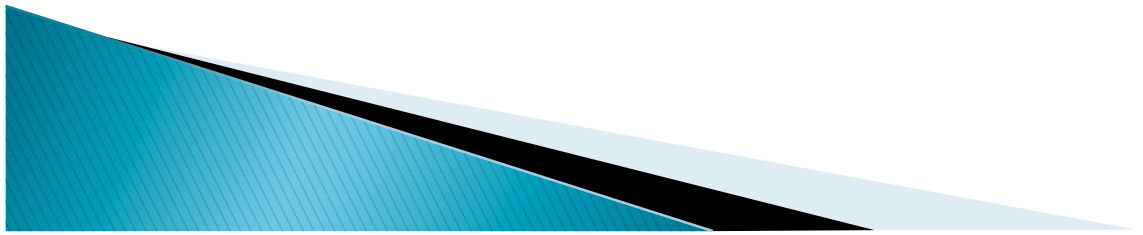
Today's Topics

- Course Information
  - Goals, Logistics, & Other practical information
- Overview of ML
  - History
  - ML Basic Paradigms
  - Why is ML Cool?



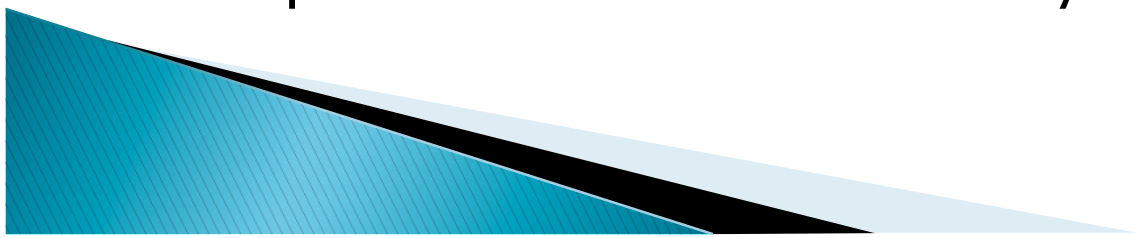
Objectives for this Class

- Learn basic ML algorithms
- Be able to recognize what problems may be amenable to ML
- Be able to critically evaluate applications of ML (aka recognize *baaaad* ML)
- Have ideas about how to apply ML in your PhD research



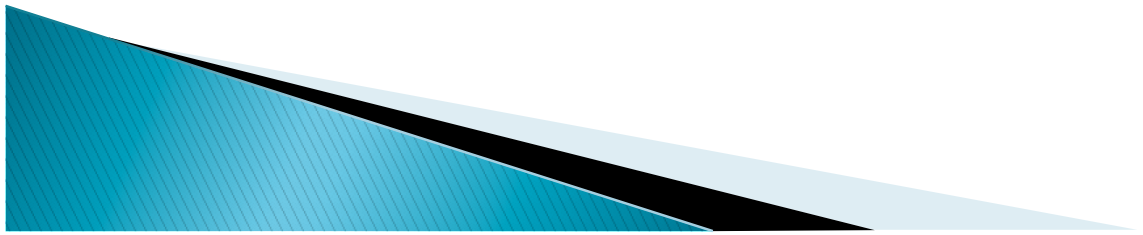
This class…

- This class will:
  - Give you the foundations for applying basic ML techniques
  - Give you enough knowledge so that you know where to start for identifying current research trends in ML
  - Be challenging, a lot of work, and fun, ☺!
- This class will not:
  - Cover all topics in ML (there's not enough time!)
  - Spoon feed the material to you…



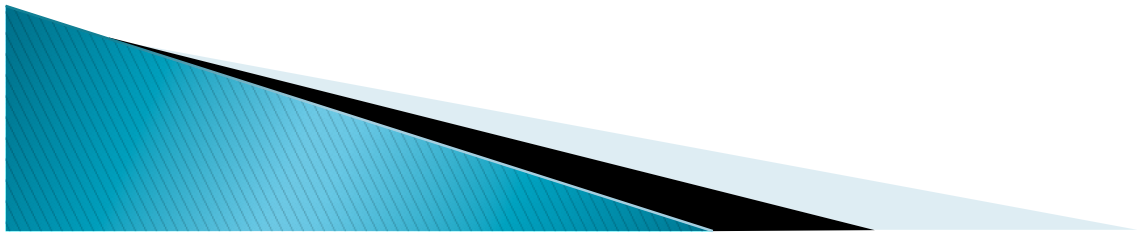
Course Logistics #1

- Course is Oversubscribed…
  - Options:
    - Course will be offered again next year…
    - Continue attending course and see if some folks drop



Course logistics #2

- Webpage:
  - http://www.cs.umd.edu/class/fall2010/cmsc726/

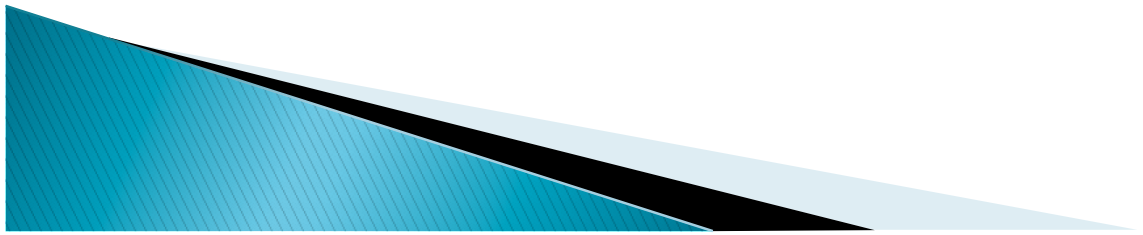


Logistics

- 5 homework assignments: 30% of grade
  - Theory exercises
  - Implementation exercises
- Final project: 20% of grade
  - Applying machine learning to your research area
    - NLP, IR, Computational biology, vision, robotics …
  - Theoretical and/or algorithmic work
    - a more efficient approximate inference algorithm
    - a new sampling scheme for a non-trivial model …
  - 3-stage reports
- Two exams: 20% Midterm, 25% final
  - Theory exercises and/or analysis
- Class participation: %5
- Policies …

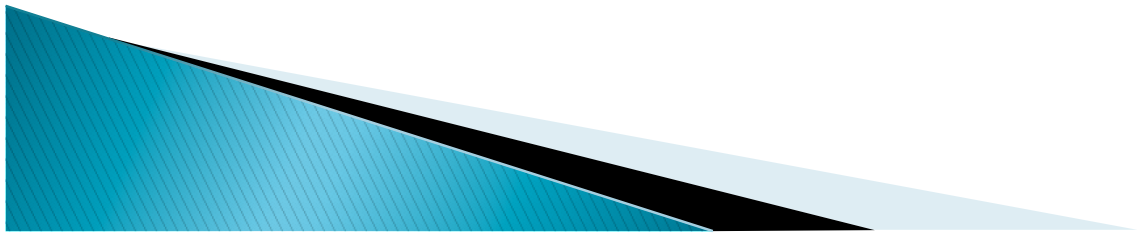
Policies: Reading

- You are REQUIRED to read the material before class
  - To facilitate this, I will try to keep the readings short and focused
- I will ASSUME that you have read the material before class
  - To facilitate this, I will try not to regurgitate material
  - I will expect you to be able to answer simple questions about the reading materials



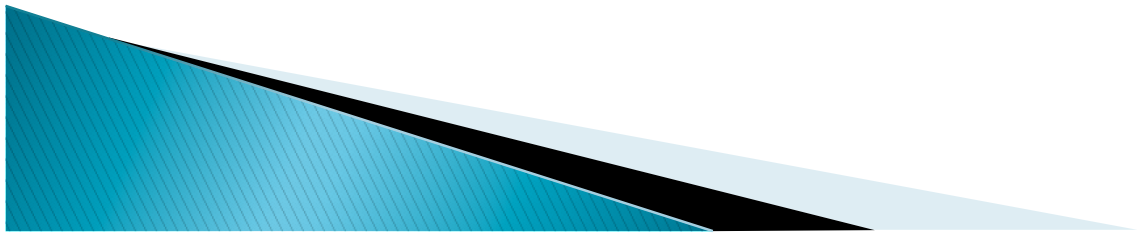
Policies: In Class

- **NO** laptop or cell phone usage during class
  - It is distracting to other people in the class
- Class participation is encouraged
  - It is a portion of the grade
  - There will be in class exercises, and you are expected to participate



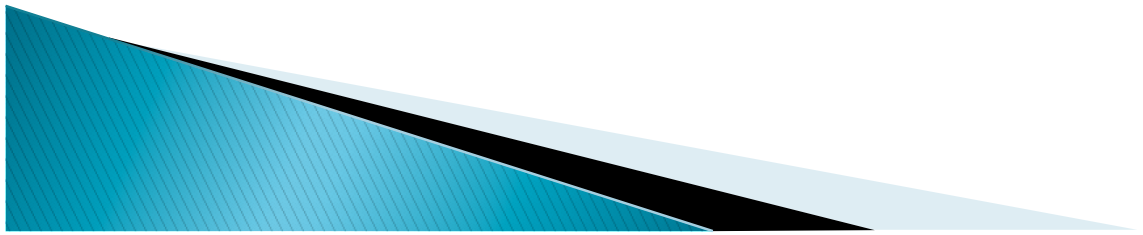
Policies: Homework

- Homeworks are due at a fixed time and will be submitted electronically.
- Late policy
  - You are given a quota of 5 FREE Late days
    - A late day is any portion of a 24 hour period
    - You may use these however you see fit
    - No homework will be accepted > 5 days late
    - Once you have used up your late days, each late day deducts 25% of the grade
- Regrade policy
  - All regrade request must be made in writing within 1 week of the homework return. When you request a regrade, your *entire* assignment will be regraded.



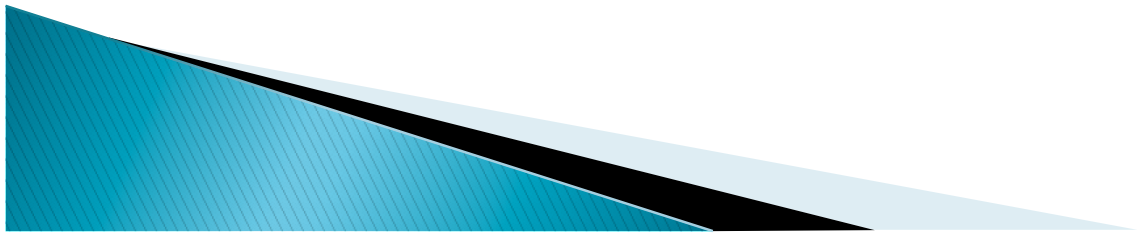
Policy: Collaboration vs. Cheating

- I *encourage* you to work together on homeworks, programming assignments and projects
  - This means you can discuss the problems at a high-level, work things out on the board, etc.
- However, I treat **CHEATING** very seriously
  - You will get a 0 grade at minimum; more likely you will be reported to the honor board; this will result in an XF or expulsion, this is not something that you want to risk!!!
- What is cheating?
  - Copying someone's solution/code
  - **Letting someone else copy your solution/code**
  - Copying ANYTHING from the web without proper attribution



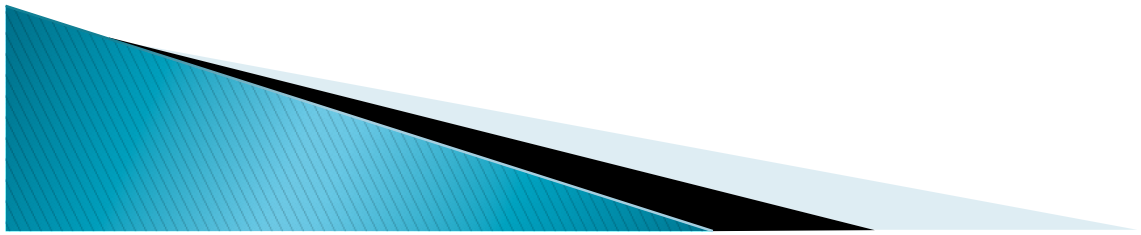
Class Exercise #1

- On the white index card that is being handed out
- You have 2.5 minutes to:
  - Collect the names and emails addresses of as many neighboring students as you can!



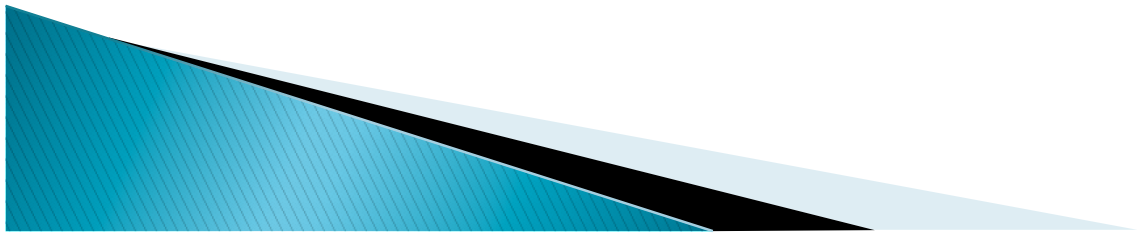
Homework #0

- Do course survey here:
  - http://www.surveymonkey.com/s/CXDZD9
- Sign up for class mailing list here:
  - http://mailman.cs.umd.edu/mailman/listinfo/cmsc726-f10
- Due by noon, Thu Sept 2, before next class meeting!



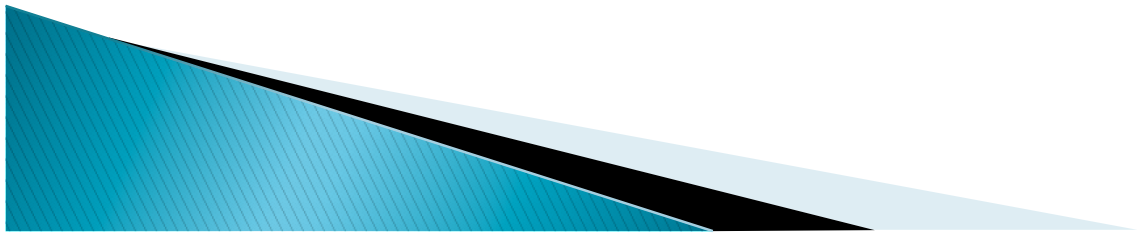
End administrivia ---

- Questions????



Intro to ML

- History
- Basic ML Paradigms
- Why ML is cool



Why a History is Hard?

- Intersection of fields:
  - Artificial Intelligence
  - Pattern recognition
  - Statistics
  - Linear algebra
  - Information theory
  - Probability
  - Optimization
  - Decision theory
  - Theoretical physics
  - …
- And applications:
  - Computer security, networks, systems, architecture, databases
  - Computational biology
  - Computational chemistry
  - Computational social science
  - Physics
  - Language processing
  - Speech
  - Vision
  - Robotics
  - Information retrieval
  - …

Origins… Artificial Intelligence

- As a broad subfield of Artificial Intelligence, "**Machine Learning**" is concerned with the development of algorithms and techniques that allow computers to "learn".

  – Wikipedia

What does it mean to learn?

ML Definitions

- **Definition 1.** From Tom Mitchell, 1997: A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

- **Definition 2.** Statistics + Algorithms

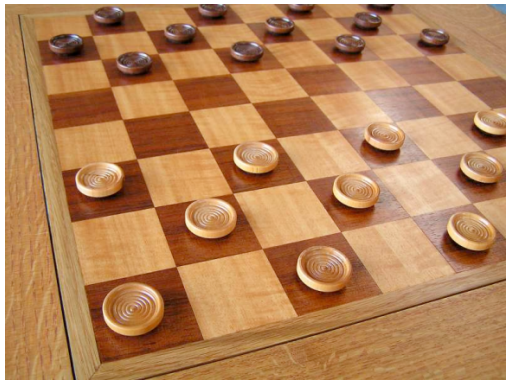
- **Definition 3.** Fancy Function Fitting

Checkers Playing Program

- Arthur Samuel
  - Masters in electrical engineering, MIT 1928
- 1959: To demonstrate IBM 701 created
  checkers program
- Program taught computer to play checkers
  - Strategy for exploring moves
  - Scores board positions based on chance of winning
  - Features from position of pieces, proximity of pieces to being kinged, etc.

Checkers Playing Program

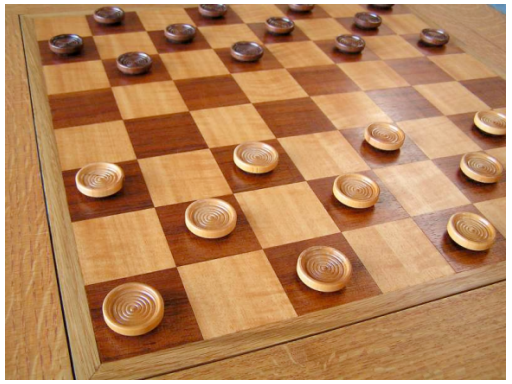
- Rote learning
  - Remember every observed configuration
  - Save the end result for the configuration
  - Score positions by tuning parameters
  - Play against self to learn best strategy



- An early form of reinforcement learning

AI Success?

- "We'll have AI solved in 10 years!"
- If a computer could play checkers, it could do X.
- What went wrong?
  - Learning done by human (picked features/parameters)
  - Parameter tuning done by computer (picked values)



Perceptron

- Frank Rosenblatt
  - Cornell University, 1957-1960
  - MARK 1 (Physical machine, 1960)
- Motivation
  - Output is a combination of input
  - Similar to neurons in the brain
- Application
  - Character recognition

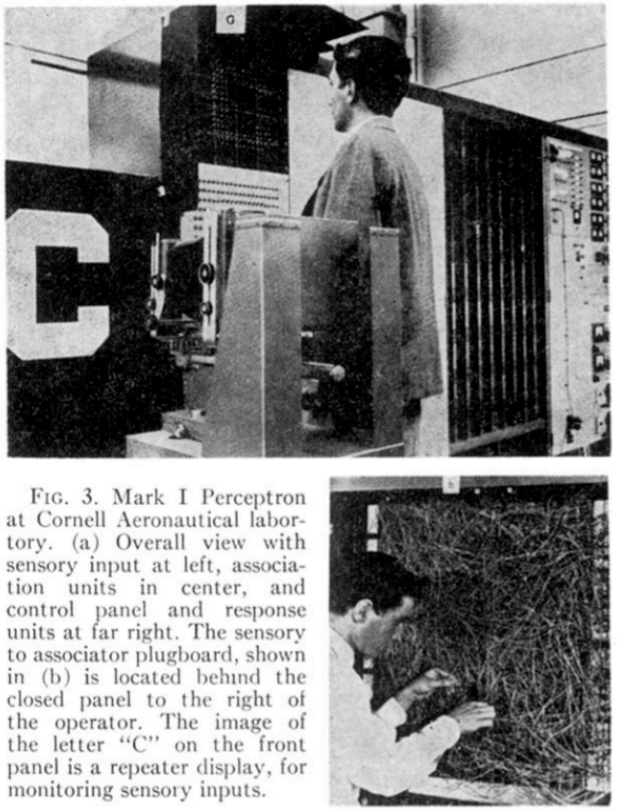


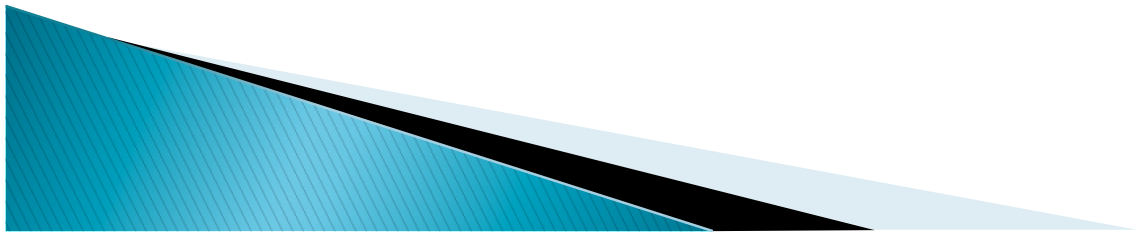
FIG. 3. Mark I Perceptron at Cornell Aeronautical laboratory. (a) Overall view with sensory input at left, association units in center, and control panel and response units at far right. The sensory to associator plugboard, shown in (b) is located behind the closed panel to the right of the operator. The image of the letter "C" on the front panel is a repeater display, for monitoring sensory inputs.

Rev. Mod. Phys. 34, 123 – 135 (1962)

Perceptron

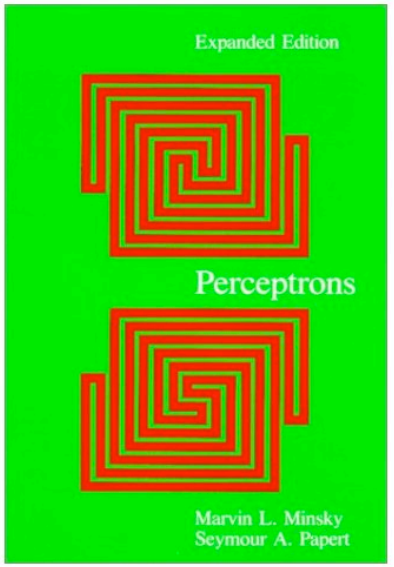
- General purpose machine to learn by trial and error
- Now known as a linear classifier
- By studying the Perceptron, “the fundamental laws of organization which are common to all information handling systems, machines and men included, may eventually be understood.”

  – Frank Rosenblatt



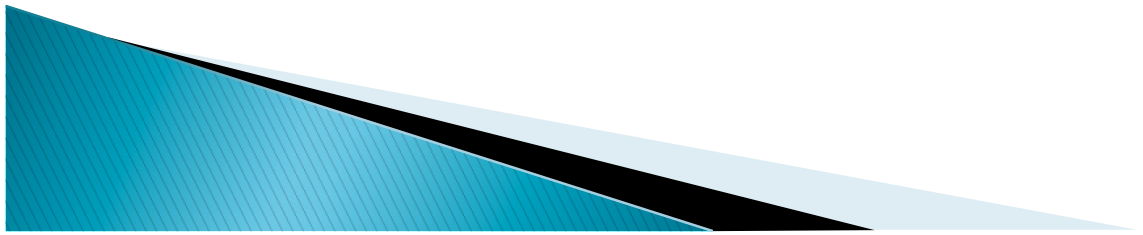
Perceptron

- Perceptrons: An Introduction to Computational Geometry
  - Marvin Minsky, Seymour Papert
  - 1969
- Criticism
  - Demonstrated some limitations of Perceptrons
- Result
  - Research in Perceptrons (aka neural networks) stops for almost 20 years



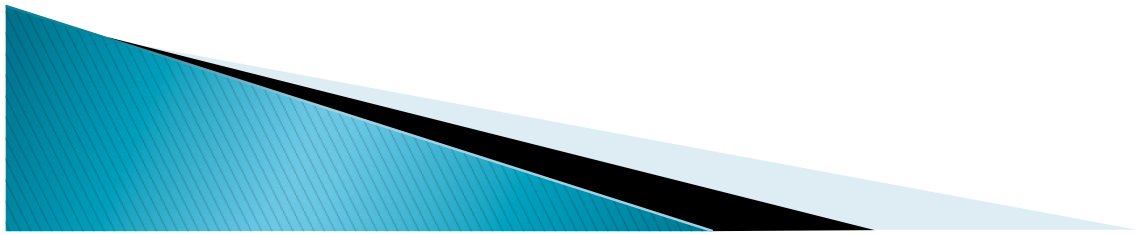
Machine Learning as a Field

- July 1980- first workshop on machine learning
  - Second workshop in 1983
  - 1983 Workshop Report
  - Focus on knowledge induction, rule induction
  - Continued every year (now ICML)



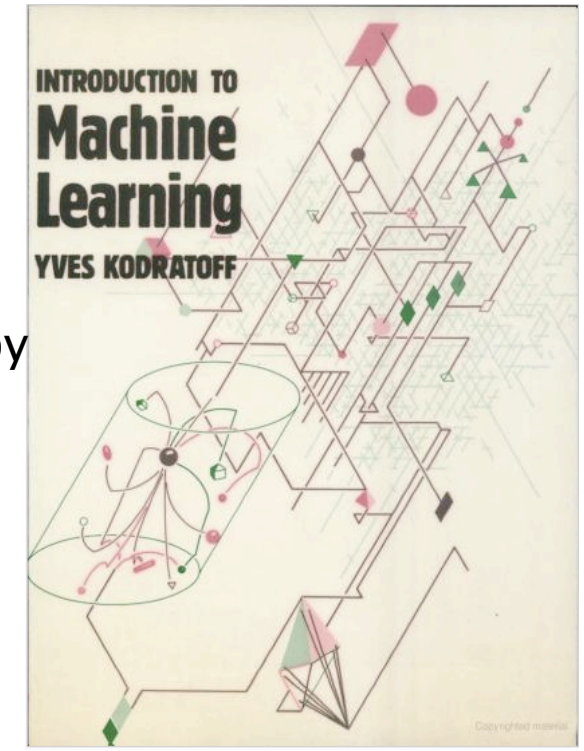
Machine Learning as a Field

- Focus on solving problems
- "What does it mean to learn?"
  - The answer keeps changing
- Initially grew from AI
- Statistical methods revolutionized the field
  - Some methods in late 1980s
  - Took off in late 1990s



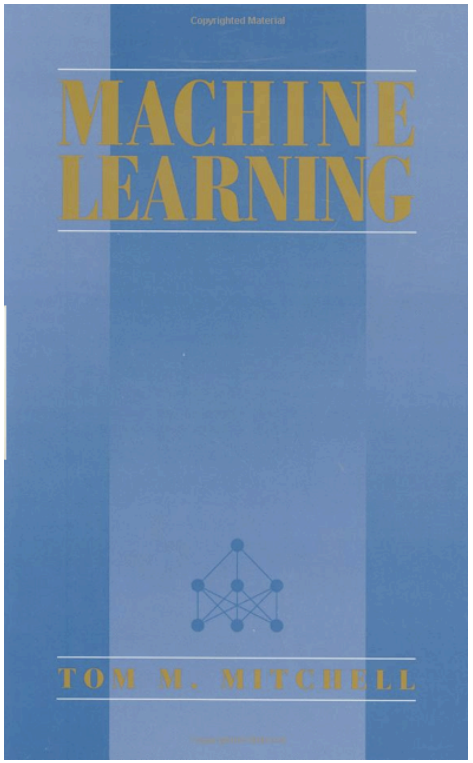
Textbooks: 1988

- Introduction to Machine Learning
  - Yves Kodratoff
- Main Topics
  - Theory: clauses, unification
  - Representation of complex knowledge by clauses
  - Representation of knowledge about actions
  - Learning by doing- rule acquisition
  - Version spaces
  - Explanation based learning
  - Learning by similarity detection
  - Construction of taxonomies
  - Learning by analogy



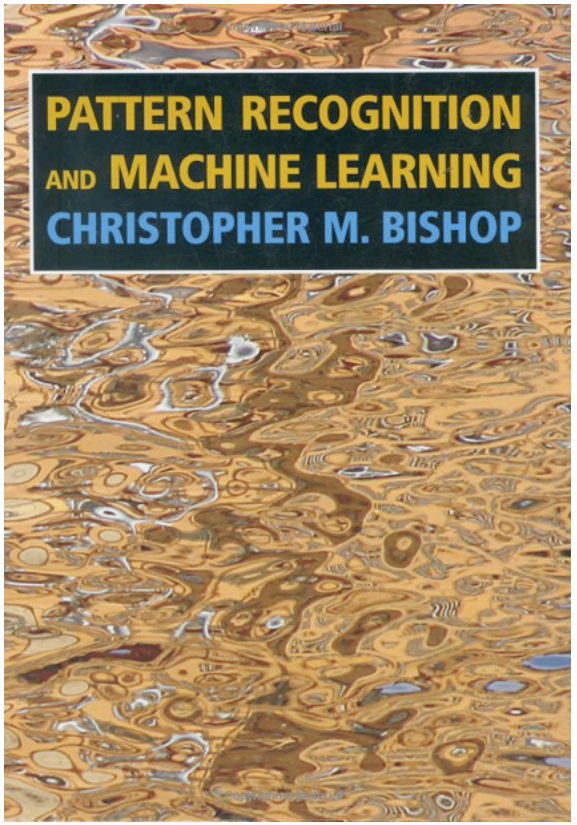
Textbooks: 1997

- Machine Learning
  - Tom Mitchell
- Main Topics
  - Concept Learning and the General-to-Specific Ordering
  - Decision Tree Learning
  - Artificial Neural Networks
  - Evaluating Hypotheses
  - Bayesian Learning
  - Computational Learning Theory
  - Instance-Based Learning
  - Genetic Algorithms
  - Learning Sets of Rules
  - Analytical Learning
  - Combining Inductive and Analytical Learning
  - Reinforcement Learning

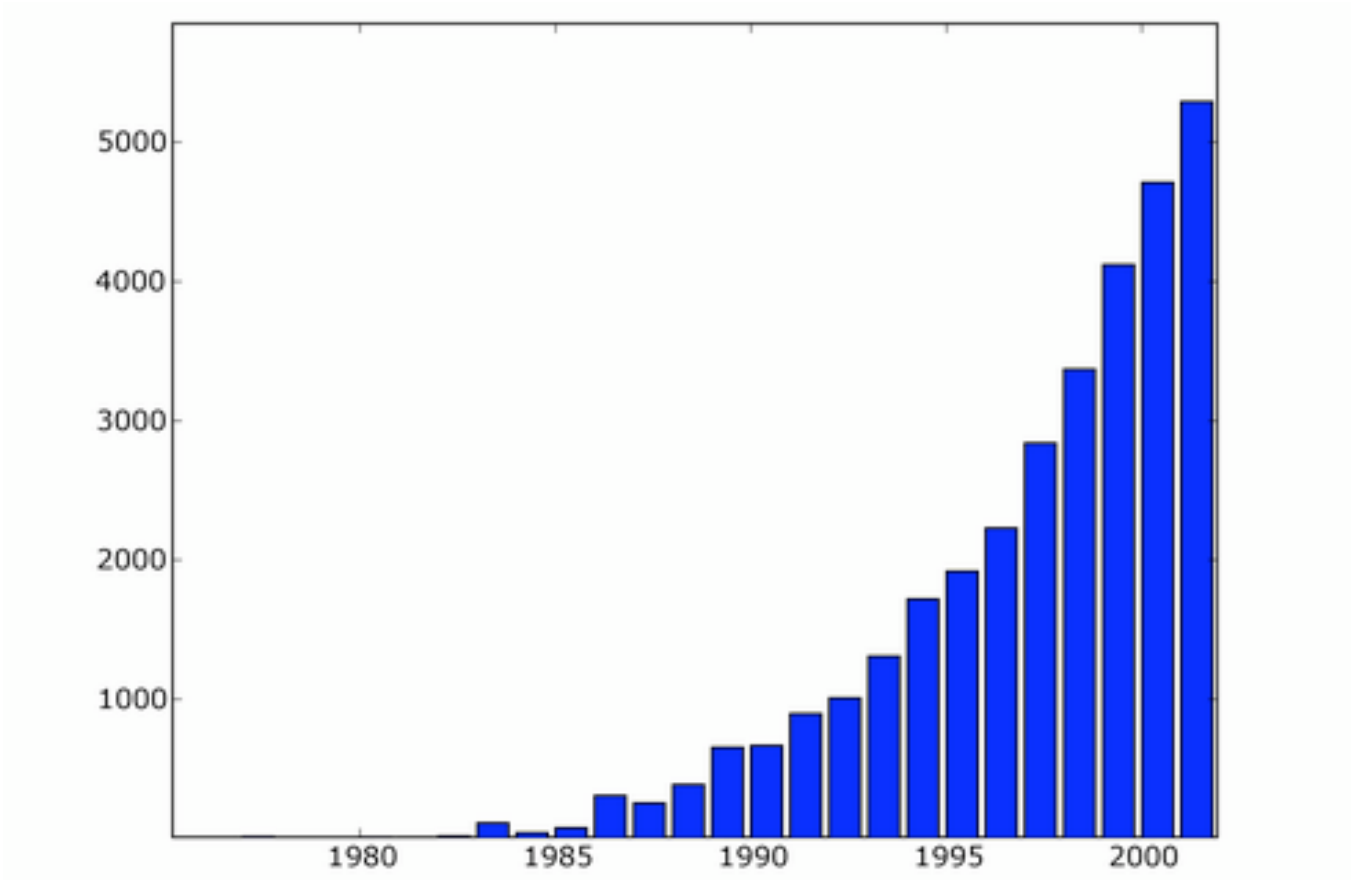


Textbooks: 2008

- Pattern Recognition and Machine Learning
  - Chris Bishop
- Main Topics
  - Probability distributions
  - Linear models
  - Neural networks
  - Kernel methods/SVMs
  - Graphical models
  - Mixture Models
  - Sampling methods and inference
  - Markov models

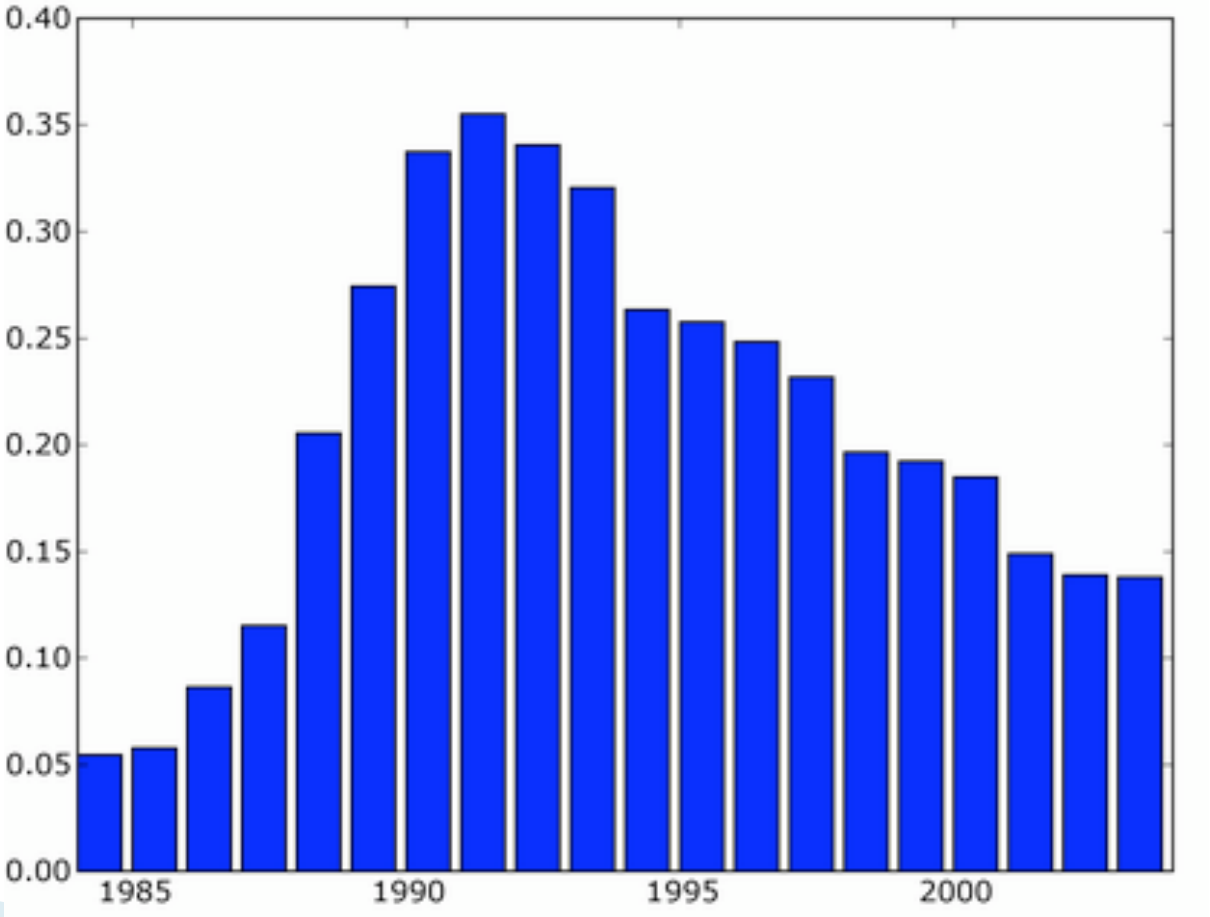


Machine Learning Articles by Year

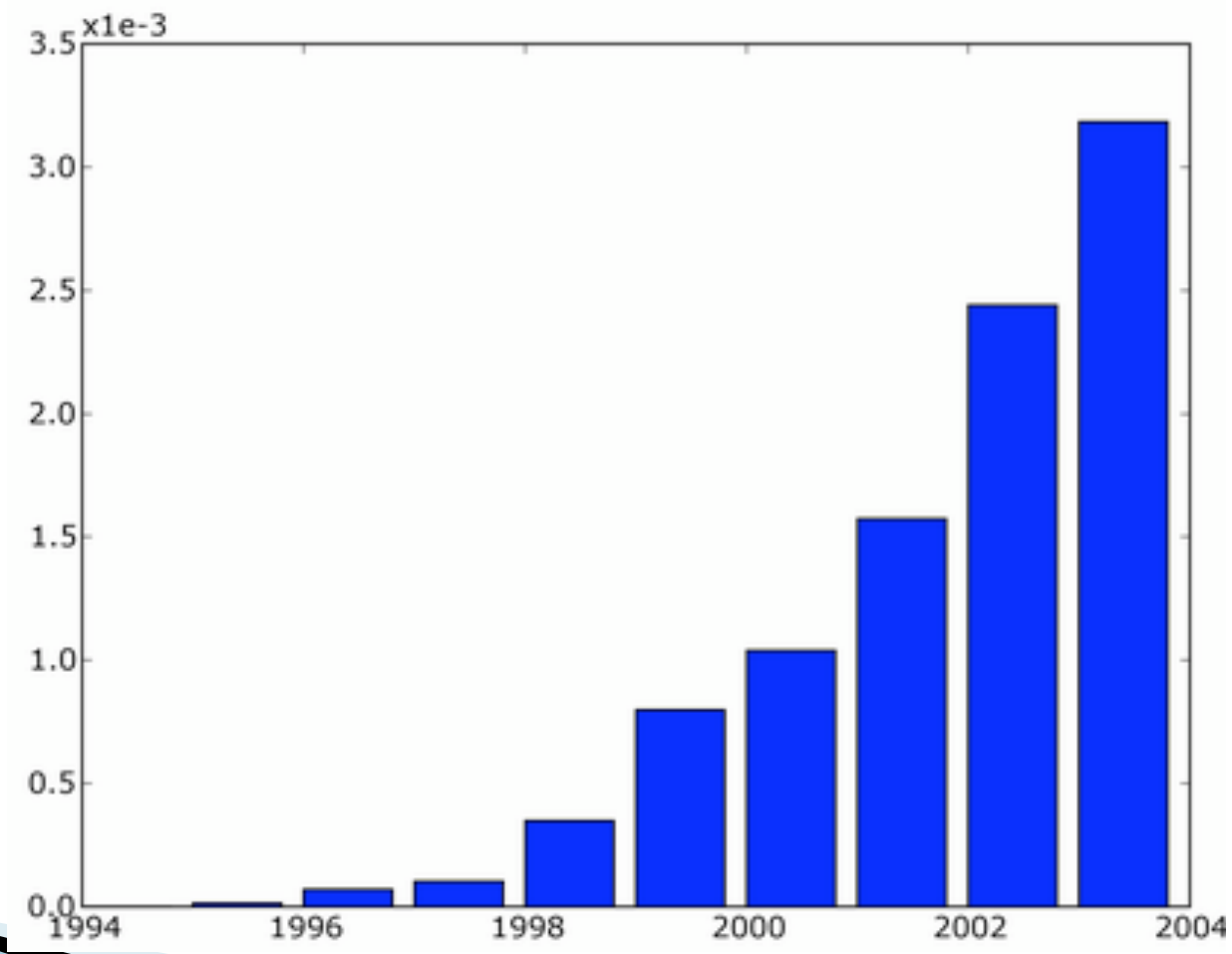


http://yaroslavvb.blogspot.com/2005_12_01_archive.html

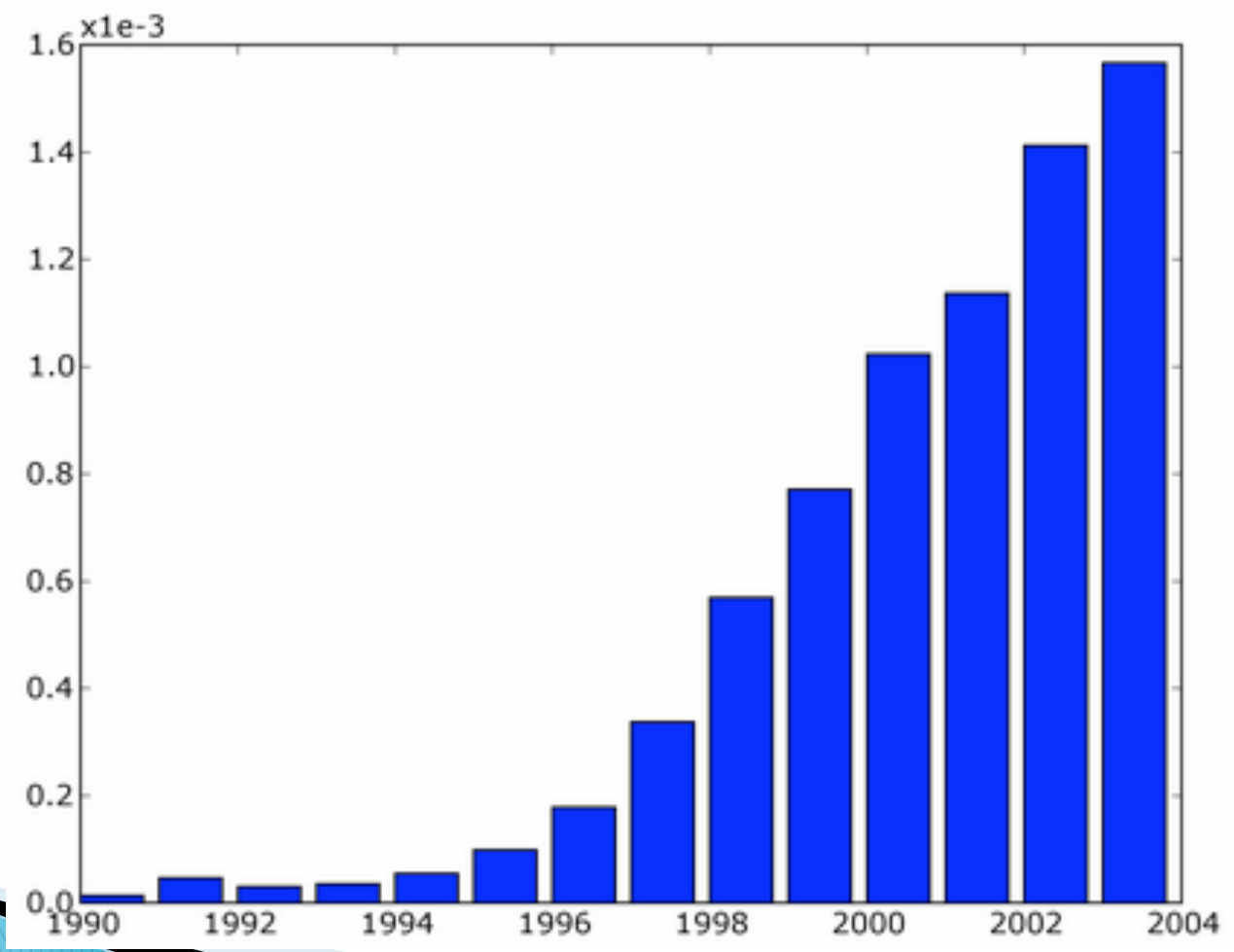
Neural Networks



SVMs

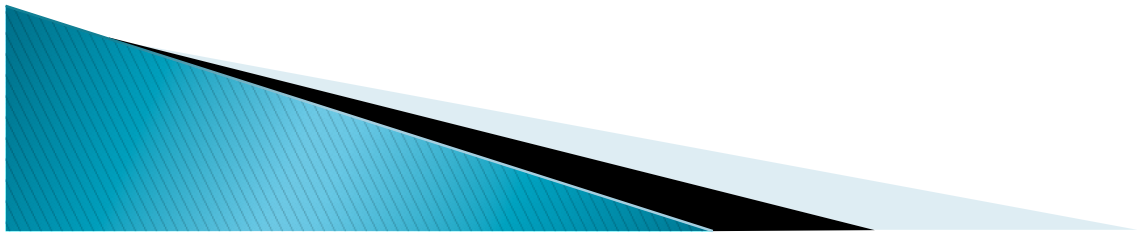


Naïve Bayes


x1e-3
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1990
1992
1994
1996
1998
2000
2002
2004

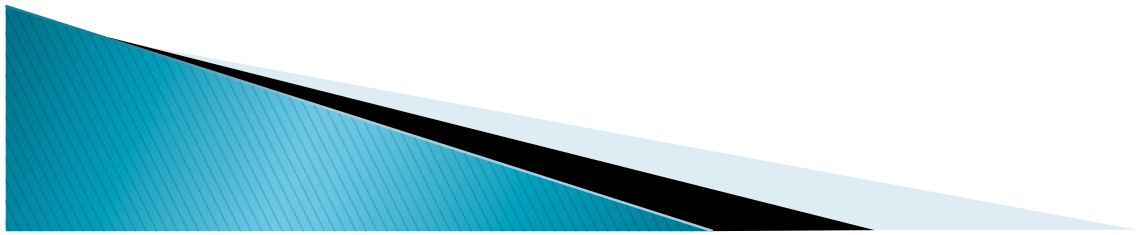
Basic ML Paradigms

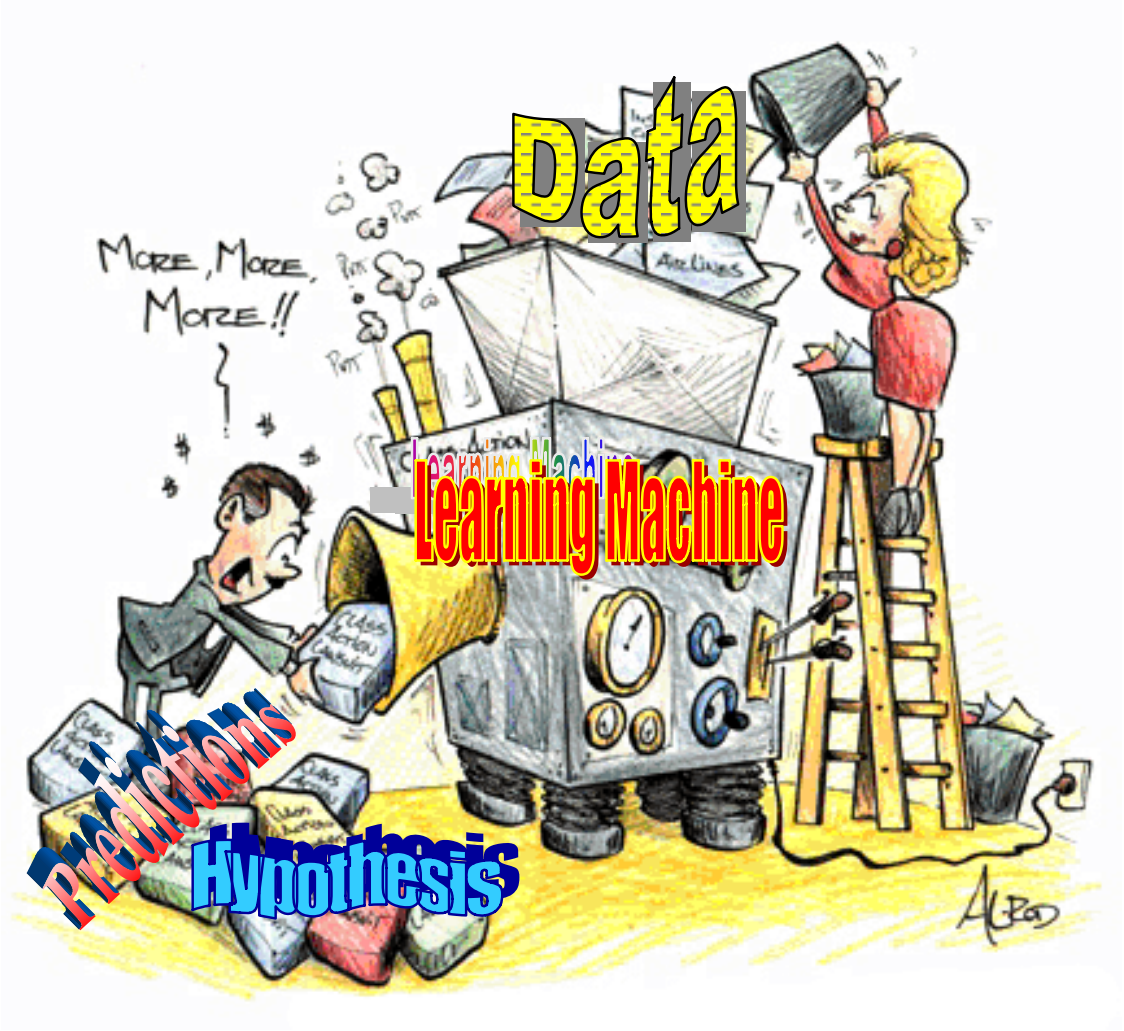
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- Learning with Complex Output Spaces



Why is ML Hot?

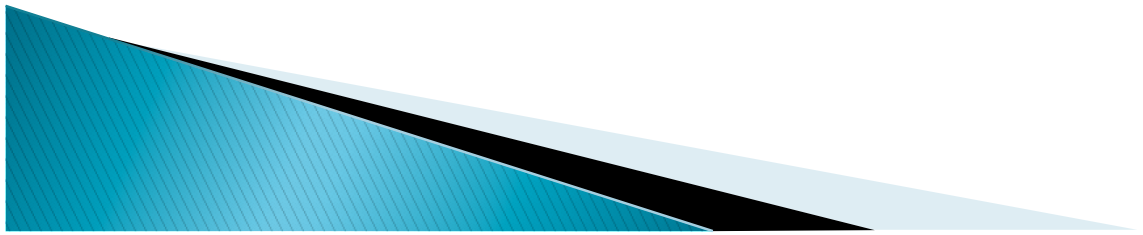
- ???




Data
MORE, MORE, MORE!!
Learning Machine
Predictions
Hypothesis

Why is ML Cool?

- Wonderful mix of theory and practice
  - Rich, elegant mathematics behind many of the methods
  - Many cool and relevant applications



Next Time….

- Reading: Bishop 1–1.1, 1.3–1.4