Outline of Today

- Who we are?
- What is learning?
- Why should a computer be able to learn?
- Examples of machine learning
- What it takes to build a learning system?
- Syllabus
- Administrivia
  - Pre-Requisites
  - Assignments
  - Grading
  - Textbook and course material
  - Office Hours
- Machine learning case study

(One) Definition of Learning

Definition [Mitchell]:
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.

Syllabus

- Concept Learning: Hypothesis space, version space, target concept
- Instance-Based Learning: K-nearest neighbor, collaborative filtering
- Decision Trees: TDIDT, Representation bias vs. search bias
- Hypothesis Tests: Confidence intervals, resampling estimates
- Linear Rules: Perceptron, Winnow
- Support Vector Machines: Optimal hyperplane, Kernels
- Generative Models: Naïve Bayes, MAP and Bayesian learning
- Hidden Markov Models: Viterbi, Expectation-Maximization
- Complex Output Prediction: natural language parsing
- Learning Theory: PAC learning, Mistake Bounds, No-Free-Lunch
- Clustering: HAC, k-means, latent semantic indexing
- Reinforcement Learning: Q-Learning, Temporal difference learning

Textbook and Course Material

- Main Textbook
- Additional Reference (optional)
- Handouts for topics not covered in textbook
- Course Notes
  - Slides available on course homepage
  - Material on blackboard

Pre-Requisites and Related Courses

- Pre-Requisites
  - Programming skills (e.g. COM S 211)
  - Basic linear algebra
  - Basic probability theory (e.g. COM S 280).
- Related Courses
  - CS472: Foundations of Artificial Intelligence
  - CS578: Empirical Methods in Machine Learning
  - CS678: Advanced Topics in Machine Learning
  - CS778: Seminar in Machine Learning
  - CS630: Language Technologies
Assignments and Grading

- **Deliverables**
  - 2 Prelim Exams (40% of Grade)
  - Final Project (15% of Grade)
  - Homeworks (~5 assignments) (40% of Grade)
  - Class Participation (5% of Grade)

- **Policies**
  - Assignments are due at the beginning of class on the due date.
  - Assignments turned in late will drop 5 points for each period of 24 hours for which the assignment is late.
  - No assignments will be accepted after the solutions have been made available.
  - Collaborations are not allowed (except when explicitly permitted).
  - Must state all sources of material used in assignments or project.
  - Academic Integrity

Final Project

- **Organization**
  - Self-defined topic related to your interests and research
  - Groups of 3-4 students
  - Each group has TA as advisor

- **Deliverables**
  - Project proposal (~ week after spring break)
  - Meetings with TA to discuss progress
  - Short presentation in class (last week of classes)
  - Project report (~ exam period)

How to Get in Touch

- **WWW Page**

- **Email Addresses**
  - Thorsten Joachims: tj@cs.cornell.edu
  - Rich Caruana: caruana@cs.cornell.edu
  - Chun-Nam Yu: cnyu@cs.cornell.edu
  - Chris Quinn, Bob Albright, Nick Gallo, Dave Golland
  - Mailing list to all course staff: cs478-l@lists.cs.cornell.edu

- **Office Hours**
  - Thorsten Joachims:
    - TBD, 4153 Upson Hall
  - Rich Caruana
    - TBD, 4157 Upson Hall
  - Other office hours:
    - TBD