Machine Learning: CS 6375

Introduction

Instructor: Vibhav Gogate
The University of Texas at Dallas
Logistics

• **Instructor:** Vibhav Gogate
  – Email: Vibhav.Gogate@utdallas.edu
  – Office: ECSS 3.406
  – Office hours: Mondays 4 p.m. to 6 p.m.

• **TA:** To be announced

• **Web:**

• **Discussion Board**
  – Discussion board on **Piazza**. This will be the main on-line forum for discussing assignments and course material, and interacting with other students, TA and me. We will also post course-wide announcements on **Piazza**.
Evaluation

• Five homeworks (25%)
  – 5% each
  – Due two weeks later
  – Some programming, some exercises
  – Assigned via elearning.

• One Project (25%)

• One Midterm (15%)
  – March 28, in class

• One Final (35%)
  – May 14, ECSS 2.303, 2:00 to 4:45 p.m.
  – Exams are closed book. You will be allowed a cheat sheet, a double-sided 8.5 x 11 page.

• A (90 or above), A- (85 to 89), B+ (80 to 84), B (75 to 79), B- (70 to 74), Fail (69 and below).
Source Materials


• C. Bishop, *Pattern Recognition and Machine Learning*, Springer (Required/Recommended)

• R. Duda, P. Hart & D. Stork, *Pattern Classification* (2nd ed.), Wiley (Recommended)

• Papers
Why Study Machine Learning: A Few Quotes

• “A breakthrough in machine learning would be worth ten Microsofts” (Bill Gates, Microsoft)
• “Machine learning is the next Internet” (Tony Tether, Former 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)
So What Is Machine Learning?

• Automating automation
• Getting computers to program themselves
• Writing software is the bottleneck
• Let the data do the work instead!
Traditional Programming

Data → Computer → Output

Program → Computer

Machine Learning

Data → Computer → Program

Output → Computer
Magic?

No, more like gardening

- **Seeds** = Algorithms
- **Nutrients** = Data
- **Gardener** = You
- **Plants** = Programs
Definition: Machine Learning!

- T. Mitchell: Well posed machine learning
  - Improving performance via experience
  - Formally, 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.

- 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 (Learning = Generalization)
Example 1: A Chess learning problem

- Task $T$: playing chess
- Performance measure $P$: percent of games won against opponents
- Training Experience $E$: playing practice games against itself
Example 2: Autonomous Vehicle Problem

- Task T: driving on a public highway/roads using vision sensors
- Performance Measure P: percentage of time the vehicle is involved in an accident
- Training Experience E: a sequence of images and steering commands recorded while observing a human driver
ML in a Nutshell

• Tens of thousands of machine learning algorithms
• Hundreds new every year
• Every machine learning algorithm has three components:
  – Representation
  – Evaluation
  – Optimization
Representation

• Decision trees
• Sets of rules / Logic programs
• Instances
• Graphical models (Bayes/Markov nets)
• Neural networks
• Support vector machines
• Model ensembles
• Etc.
Evaluation

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

• Combinatorial optimization
  – E.g.: Greedy search

• Convex optimization
  – E.g.: Gradient descent

• Constrained optimization
  – E.g.: Linear programming
Types of Learning

• **Supervised (inductive) learning**
  – Training data includes desired outputs

• **Unsupervised learning**
  – Training data does not include desired outputs
  – Find hidden structure in data

• **Semi-supervised learning**
  – Training data includes a few desired outputs

• **Reinforcement learning**
  – The learner interacts with the world via “actions” and tries to find an optimal policy of behavior with respect to “rewards” it receives from the environment
Types of Supervised Learning Problems

• Classification: learning to predict a discrete value from a predefined set of values
• Regression: learning to predict a continuous/real value
Machine Learning: Applications

Examples of what you will study in class in action!
Classification Example: Spam Filtering

Classify as “Spam” or “Not Spam”
Classification Example: Weather Prediction
Regression example: Predicting Gold/Stock prices

Given historical data on Gold prices, predict tomorrow’s price!

Good ML can make you rich (but there is still some risk involved).
Similarity Determination
Collaborative Filtering

- The problem of collaborative filtering is to predict how well a user will like an item that he has not rated given a set of historical preference judgments for a community of users.
Collaborative Filtering
Collaborative Filtering
Clustering: Discover Structure in data
Clustering images

Set of Images

[Goldberger et al.]
Machine learning has grown in leaps and bounds

• The main approach for
  – Speech Recognition
  – Robotics
  – Natural Language Processing
  – Computational Biology
  – Sensor networks
  – Computer Vision
  – Web
  – And so on

Alice/Bob says: I know machine learning very well!
Potential Employer: You are hired!!!
What We’ll Cover

- **Supervised learning:** Decision tree induction, Rule induction, Instance-based learning, Bayesian learning, Neural networks, Support vector machines, Linear Regression, Model ensembles, Graphical models, Learning theory, etc.

- **Unsupervised learning:** Clustering, Dimensionality reduction

- **Reinforcement learning:** Markov Decision Processes, Q-learning, etc.

- **General machine learning concepts and techniques:** Feature selection, cross-validation, maximum likelihood estimation, gradient descent, expectation-maximization

- **Your responsibility:**
  - Brush up on some important background
  - Linear algebra, Statistics 101, Vectors, Probability theory