

CSC 411

MACHINE LEARNING and DATA MINING

Lectures: Monday, Wednesday 3-4
Lecture Room: Bahen 1230
Instructor: Richard Zemel <zemel@cs.toronto.edu>
Office hours: Thursday 1-3 Pratt 290D, and by appointment
Tutors: David Ross, Xuming He, Xin Gu, & Xiaodan Zhu
Tutorials: Friday 3-4
Tutorial Room: Bahen 1230
Class URL: www.cs.toronto.edu/~zemel/Courses/csc411.html

Overview

Machine learning research aims to build computer systems that learn from experience. Learning systems are not directly programmed by a person to solve a problem, but instead they develop their own program based on examples of how they should behave, or from trial-and-error experience trying to solve the problem. These systems require learning algorithms that specify how the system should change its behavior as a result of experience. Researchers in machine learning develop new algorithms, and try to understand which algorithms should be applied in which circumstances.

Machine learning is an exciting interdisciplinary field, with historical roots in computer science, statistics, pattern recognition, and even neuroscience and physics. In the past 10 years, many of these approaches have converged and led to rapid theoretical advances and real-world applications.

This course will focus on the machine learning methods that have proven valuable and successful in practical applications. This course will contrast the various methods, with the aim of explaining the circumstances under which each is most appropriate. We will also discuss basic issues that confront any machine learning method.

Pre-requisites

You should understand basic probability and statistics, and college-level algebra and calculus. Knowledge of linear algebra will be a big help, as will knowledge of mathematics underlying probability models. For the programming assignments, you should have some background in programming, and it would be helpful if you know Matlab.

Readings

Most of the readings in the course will come from the following book, which is available in the bookstore: Mitchell, T. M. (1997). *Machine Learning*. Boston, MA: McGraw-Hill.

This is currently the best and most thorough text in the field. However, some of the topics I will cover are not included in the book. I will hand out supplementary readings for some of these topics. Reading assignments from the text are listed in this syllabus; other reading assignments will be given out in class.

Two optional textbooks for the course are: *Pattern Classification (2nd edition)* by Duda, Hart, and Stork; and *Neural Networks for Pattern Recognition*, by Bishop. Some of the supplementary readings will be taken from these books.

Course requirements and grading

The format of the class will be lecture, with some discussion. I strongly encourage interaction and questions. There are assigned readings for each lecture that are intended to prepare you to participate in the class discussion for that day.

Course handouts will be brought to class once. I will make postscript copies of the handouts available on the class web site (see URL above).

The grading in the class will be divided up as follows:

Assignments	50%
Mid-Term Exam	20%
Final Exam	25%
Attendance and participation	5%

There will be four assignments, each of which will be worth 1/8 of your grade.

Final grades will be based on a class curve. I will give each assignment and test a numerical score, and will try to give you an idea of where you lie along the curve as the course progresses.

Homework assignments

The best way to learn about a machine learning method is to program it yourself and experiment with it. So the assignments will generally involve implementing machine learning algorithms, and experimentation to test your algorithms on some data. You will be asked to summarize your work in brief (3-4 page) write ups. The implementations must be done in Matlab, but prior knowledge of Matlab is not required. A brief tutorial on Matlab is available from the course web-site.

Collaboration on the assignments is not allowed. Each student is responsible for his or her own work. Discussion of assignments and programs should be limited to clarification of the handout itself, and should not involve any sharing of pseudocode or code or simulation results. Violation of this policy is grounds for a semester grade of F, in accordance with university regulations.

The schedule of assignments is included in the syllabus. Assignments are due at the beginning of class on the due date. Because they may be discussed in class that day, it is important that you have completed them by that day. Assignments handed in late but before 5 pm of that day will be penalized by 5% (i.e., total points multiplied by 0.95); a late penalty of 20% per day will be assessed thereafter. Extensions will be granted only in special situations, and you will need a Student Medical Certificate or a written request approved by the instructor at least one week before the due date.

Exams

There will be a mid-term in class on October 27th, which will be a closed book exam on all material covered up to that point in the course and assigned readings.

The final will not be cumulative, except insofar as concepts from the first half of the semester are essential for understanding the later material.

Attendance

I expect students to attend all classes, and all tutorials. This is especially important because I will cover material in class that is not included in the textbook. Also, the tutorials will not only be for review and answering questions, but new material will be covered. So 5% of your grade will be based on attendance and participation.

CLASS SCHEDULE

(subject to change)

Lecture	Topic	Reading	Assignments
Sep 13, 15	Introduction <ul style="list-style-type: none"> · basic problems · examples of learning systems · generalization 	Chapter 1, 5	
Sep 20, 22	Simple classifiers <ul style="list-style-type: none"> · naive Bayes · k-nearest neighbors 	Sections 6.1-6.2; 6.9; 8.1-8.2	Asst 1 Out (9/20)
Sep 27, 29	Decision Trees <ul style="list-style-type: none"> · attribute selection · pruning methods · missing data 	Chapter 3	
Oct 4, 6, 13	Neural Networks <ul style="list-style-type: none"> · perceptrons · LMS · multi-layer networks 	Chapter 4	Asst 1 In (10/4) Asst 2 Out (10/6)
Oct 18, 20, 25	Algorithm independent topics <ul style="list-style-type: none"> · bias-variance · bagging & boosting 	Supplementary reading	Asst 2 In (10/20)
Oct 27	MIDTERM		

Lecture	Topic	Reading	Assignments
Nov 1, 3, 8, 10	Unsupervised Learning <ul style="list-style-type: none"> · density estimation · mixture models · <i>k</i>-means · components analysis · multidimensional scaling 	Sections 6.11-6.12 Supplementary reading	Asst 3 Out (11/3)
Nov 15, 17, 22	Reinforcement Learning <ul style="list-style-type: none"> · Markov models · Q learning · TD learning 	Chapter 13	Asst 3 In (11/15) Asst 4 Out (11/17)
Nov 24, 29	Support Vector Machines <ul style="list-style-type: none"> · kernel methods · the kernel trick 	Supplementary reading	
Dec 1, 6, 8	HMMs <ul style="list-style-type: none"> · forward-backward algorithm · application: speech recognition 	Supplementary reading	Asst 4 In (12/1)