CSE 4404A/5327A 3.0 (F) 2012-13
Introduction to Machine Learning and Pattern Recognition

Chemistry Building 122 MW 17:30-19:00
Course Website: www.cse.yorku.ca/course/4404-5327

Instructor Information:
James Elder
0003G Computer Science and Engineering Building
tel: (416) 736-2100 ext. 66475 fax: (416) 736-5857
email: jelder@yorku.ca website: www.yorku.ca/jelder
Office Hour: Wednesday 12:30 – 13:30

Purpose:
Machine learning is the study of algorithms that learn how to perform a task from prior experience. Machine learning algorithms find widespread application in diverse problem areas, including machine perception, natural language processing, search engines, medical diagnosis, bioinformatics, brain-machine interfaces, financial analysis, gaming and robot navigation. This course will thus provide students with marketable skills and also with a foundation for further, more in-depth study of machine learning topics.

This course introduces the student to machine learning concepts and techniques applied to pattern recognition problems in a diversity of application areas. The course takes a probabilistic perspective, but also incorporates a number of non-probabilistic techniques.

Learning Objectives:
Upon completing this course the student will, through the assignments, test, and final project, have demonstrated an ability to:

- Use probabilistic modeling and statistical analysis of data to develop powerful pattern recognition algorithms.
- Identify machine learning models and algorithms appropriate for solving specific problems.
- Explain the essential ideas behind core machine learning models and algorithms
- Identify the main limitations and failure modes of core machine learning models and algorithms
- Program moderately complex machine learning algorithms
- Manage data and evaluate and compare algorithms in a supervised learning setting
- Access and correctly employ a variety of machine learning toolboxes currently available.
- Identify a diversity of pattern recognition applications in which machine learning techniques are currently in use.

Prerequisites:
One of MATH2030 3.0 or MATH1131 3.0. MATH1025 3.0 or a similar introductory course in linear algebra is strongly recommended.

Course Format:
Through lectures and short presentations of application papers by graduate students, students will learn the fundamental concepts and principles of machine learning techniques, and how they can be applied. This knowledge will be evaluated in part through the midterm.

Through the two assignments, students will have a chance to solve specific, relatively short programming problems in which machine learning algorithms are coded and applied to specific datasets. **Please note that the two marked assignments are to be done individually: collaboration is not allowed.**

Through the course project, students will have an opportunity to design and implement a more complete machine learning solution to a problem of their choosing: a selection of candidate problems and datasets will be made available by the instructor.

Basis of Evaluation:

Undergraduate Student Evaluation:

- 30% Assignments (2)
- 35% Midterm (Closed book)
- 35% Final Project

Graduate Student Evaluation:

- 15% Presentation(s) to class on one or more pattern recognition applications of machine learning
- 25% Assignments (2) - each will include one additional, more in-depth exercise
- 30% Midterm (Closed book) - will include one additional, in-depth question
- 30% Final Project – expected to be more in-depth and to incorporate more novel elements.

Important Dates:

- Wed, Oct 17 Assignment 1 Due
- Mon, Oct 29 Midterm
- Fri, Nov 9 Drop Date
- Mon, Dec 3 Assignment 2 Due
- Fri, Dec 14 Final Project Due

Software:

Problems and assignments will be based upon MATLAB, which is available:

- In the PRISM lab (including some toolboxes). To use MATLAB remotely, ssh to red.cse.yorku.ca with the –X option and enter the **matlab** command at the shell prompt.
- In the Computing Commons and through WebFAS (no toolboxes)
- Individual student licenses available from www.mathworks.com for $99.

MATLAB is a vector-based language with a *c*-like syntax. Tutorials are available at http://www.mathworks.com/academia/student_center/tutorials, and webinars are available at www.mathworks.com/company/events/webinars/.

Projects may be completed using any programming language, but use of MATLAB is encouraged.

Texts:

- **Primary:** K.P. Murphy (2012) *Machine Learning: A Probabilistic Perspective.* Cambridge, MA: MIT Press.
    - Available from amazon.ca ($73) – expected delivery Oct 3, 2012
    - Available from amazon.com – expected Sept 11 with Priority Shipping ($118 all-in)
    - Now appears as available in the York bookstore ($95)
- **Secondary:** S.J.D. Prince (2012) *Computer Vision: Models, Learning and Inference.* Cambridge, UK: Cambridge University Press.
    - $87 at amazon.ca.
    - Electronic version also available for free at http://www.computervisionmodels.com/.

Additional References:

- C.M. Bishop (2006) *Pattern Recognition and Machine Learning*. New York: Springer.
    - One copy is on the 1-day reserve list at Steacie Library
- Pearl J. (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* San Mateo, CA: Morgan Kauffman, 1988.
- Li S.Z. (2009) Markov Random Field Modeling in Image Analysis, 3rd ed. London: Springer-Verlag.
- Duda R.O., Hart, P.E. & Stork D.G. (2001) Pattern Classification, 2nd ed. New York: Wiley.

Additional readings will be made available through Moodle (http://moodle.yorku.ca). Please check the Moodle site frequently.

Lecture Schedule (Approximate – subject to revision as the course unfolds):

Date	Topic	Selected Sub-Topics	Murphy	Prince
W Sept 5	Introduction		1.1 – 1.3	
M Sept 10 W Sept 12	Probability Theory	Probability Univariate Distributions Multivariate Distributions Transformations of Variables	2.1 – 2.2 2.4.1 – 2.4.4 2.5.1 – 2.5.3 2.6	2 3.5 3.7 5
M Sept 17 W Sept 19	Multivariate Normal Distribution	Naïve Bayes MVN Distribution MVN Discriminant Analysis MVN Conditionals & Marginals	3.5 – up to and incl 3.5.3 4.1.1 – 4.1.3 4.2.1 – 4.2.5 4.3.1 – 4.3.2	4.1-4.4 7.1
M Sept 24 W Sept 26	Linear Regression Linear Regression	Least Squares, Ridge Regression, Robust Regression	7.1 – 7.6 (omit 7.5.3)	8.1-8.2
M Oct 1 W Oct 3	Linear Classifiers Linear Classifiers	Logistic Regression The Perceptron Algorithm Generative vs Discriminative Fisher's LDA	8.1 – 8.3 8.5.4 8.6.1 – 8.6.2 8.6.3.1 – 8.6.3.2	9.1 6.2
M Oct 8 W Oct 10	Mixture Models and EM	Mixture of Gaussians Mixture of Experts	11.1 – 11.4.3 (21 pgs)	7.2-7.4 7.5
M Oct 15 W Oct 17	Latent Linear Models & Dimensionality Reduction	Factor Analysis Principal Components Analysis	12.1 12.2	7.6
M Oct 22 W Oct 24	Latent Linear Models & Dimensionality Reduction	Independent Components Analysis	12.6 up to and incl. 12.6.2	
M Oct 29 W Oct 31	Midterm Co-Curricular Day			
M Nov 5 W Nov 7	Sparse Linear Models	Bayesian Variable Selection L^1 Minimization	13.1 – 13.3.1, 13.3.3 13.4.1	8.6
M Nov 12 W Nov 14	Kernel Methods	Types of Kernels Using Kernels Inside GLMs The Kernel Trick	14.1, 14.2.1, 14.2.3, 14.2.4 14.3 14.4	8.4
M Nov 19 W Nov 21	Kernel Methods	Support Vector Machines Comparison of Kernel Methods Kernels for Generative Models	14.5 14.6 14.7.1 – 14.7.5	
M Nov 26 W Nov 28	Adaptive Basis Function Models	Boosting Feedforward Neural Nets (Multilayer Perceptrons)	16.4.1 – 16.4.4 16.5 - up to and incl. 16.5.5	9.7
M Dec 3	Review			

Application Paper Presentations

Presentation Guidelines
For Presenters:

- 10-15 minutes, with a maximum of 15 slides. **Please practice your talk to ensure you can complete it in the allotted time!**
- Structure your talk as follows:
  1. Define and motivate the problem.
  2. Briefly discuss the range of solutions that have been proposed
  3. Outline the solution proposed by this paper. For our benefit, please pay particular attention to the machine learning techniques employed.
  4. Are the authors applying these techniques in a straightforward way, or are they innovating in some fashion?
  5. Try to identify the key idea(s) behind the work
  6. Report on how well the methods worked for the problem as defined.
  7. Identify anything in particular that you personally learned from the paper, and anything you still do not fully understand.
  8. Identify open questions.
  9. Did you like the paper?

For the Audience:

1. Please read the paper prior to the meeting.
2. Please ask the speaker questions after his or her talk. Let them know that you were interested and were paying attention!
3. Think about how you might use some of these ideas for your own course project!