CSE 4404A/5327A 3.0 (F) 2011-12
Introduction to Machine Learning and Pattern Recognition

Chemistry Building 120 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: Monday 10:30-11:30

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 19 Assignment 1 Due
- Mon, Nov 7 Midterm
- Fri, Nov 11 Drop Date
- Mon, Dec 5 Assignment 2 Due
- Fri, Dec 16 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.

Main Texts:

This 2-book package will be available from the York bookstore for $120 (by ~Sept 13). Also available from Amazon & Chapters-Indigo.

- S. Theodoridis & K. Koutroumbas (2009) *Pattern Recognition*. Amsterdam: Academic Press.
  - One copy is on the 1-day reserve list at Steacie Library
  - Also available as an e-resource from the York Library
- S. Theodoridis, A. Pikrakis, K. Koutroumbas & D. Cavouras (2010) *Introduction to Pattern Recognition: A MATLAB Approach.* Amsterdam: Academic Press.
  - One copy is on the 1-day reserve list at Steacie Library

(For the module on linear regression, I will provide a reading from Bishop – see below.)

Optional Text:

- S. Marsland (2009) *Machine Learning: An Algorithmic Perspective*. Boca Raton, FL: CRC Press.
  - One copy is on the 1-day reserve list at Steacie Library

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
- S.J.D. Prince (2010) *Computer Vision Models*. Available in draft form at
- http://computervisionmodels.blogspot.com/
- 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	Required Reading (Theodoridis & Koutroumbas, except where noted)	Optional Reading (Marsland)
W Sept 7	Introduction		1	1
M Sept 12 W Sept 14	Bayesian Decision Theory Bayesian Decision Theory	Bayesian Classifiers Parametric Density Estimation	2.1 – 2.4, A.1 – A.4 2.5.1 – 2.5.4	8.1-8.2
M Sept 19 W Sept 21	Bayesian Decision Theory Bayesian Decision Theory	Non-Parametric Density Estimation Training & Evaluating Classifiers Bayes Nets	2.5.5 – 2.5.7 2.6 – 2.7 10.1 – 10.4	8.3 8.4, 15.1 3.3.4 – 3.3.5
M Sept 26 W Sept 28	Linear Regression Linear Regression	Linear Basis Function Models Bayesian Linear Regression	Bishop 3.1 Bishop 3.2 – 3.3	
M Oct 3 W Oct 5	Linear Classifiers Linear Classifiers	Perceptrons Least-Squares Logistic Classifiers	3.1 – 3.4 3.5 – 3.6	2
M Oct 10 W Oct 12	Reading Week			
M Oct 17 W Oct 19	Linear Classifiers Feature Selection	SVMs Preprocessing Hypothesis Testing	3.7 5.1 – 5.5	5
M Oct 24 W Oct 26	Feature Selection Dimensionality Reduction	Optimal Feature Sets PCA	5.6 – 5.8 6.1 – 6.3	 10.2
M Oct 31 W Nov 2	Dimensionality Reduction Dimensionality Reduction	SVD & ICA Kernel PCA Manifold Learning	6.4 – 6.5 6.6 – 6.7	10.4 10.5 – 10.6
M Nov 7 W Nov 9	Midterm Nonlinear Classifiers	 Multi-Layer Perceptrons Backpropagation	 4.1 – 4.4, 4.6	 3
M Nov 14 W Nov 16	Nonlinear Classifiers Nonlinear Classifiers	Backpropagation RBFs Nonlinear SVMs	4.7 – 4.12 4.13 – 4.18	3 4
M Nov 21 W Nov 23	Nonlinear Classifiers Nonlinear Classifiers	Kernel Methods Boosting	4.19 4.21 – 4.24	6 7
M Nov 28 W Nov 30	Markov Models Markov Models	Markov Chains HMMs	8.1 – 8.2.1, 9.1 – 9.4 9.5 – 9.6	15.2 – 15.3
M Dec 5	Markov Models	Training HMMs and MRFs	9.7 – 9.9	

Application Paper Presentations

Date	Application Area	Methods	Paper	Presenter
W Oct 26	Text Classification	EM, Naive Bayes, Semi-Supervised Learning	Nigam et al (2000)	**Elnaz**
W Nov 16	Face Recognition	Linear Regression, Lasso	Wright et al (2009)	**Alex**
W Nov 23	Face Recognition	FLD, PCA, RBFs	Er (2002)	**Ying**
W Nov 30	Visual Tracking	SVMs, RVMs,	Williams et al (2005)	**Larry**
M Dec 5	Face Detection	Boosting	Viola & Jones (2004)	**Ron**

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!