University of Toronto at Mississauga, Spring 2007

csc411—Machine Learning and Data Mining

Machine learning aims to build computer systems that learn from experience, instead of being directly programmed. It is an exciting interdisciplinary field, with historical roots in computer science, statistics, pattern recognition, and even neuroscience and physics. In the past ten years, many of these approaches have converged and led to rapid theoretical advances and real-world applications. This course should be of interest to students in computer science, statistics, mathematics, engineering and bioinformatics.

Instructor: Anthony Bonner, email: ajbonner [at] sympatico [dot] ca, office hours: TBA, phone: `905-828-3813` (UTM), `416-978-7441` (St. George).

Classes: Wed 11am–1pm, CC 2150.

Tutorials: Fri 2–3pm, SE 3007.

Web Page: `http://www.cs.toronto.edu/~bonner/courses/2007s/csc411/`

Grading Scheme: Four assignments, 13% each; Midterm test, 13%; Final exam, 35%. The final exam will be based on the assignments and will assume that you have completed them by yourself. On all work, 20% of the mark will be for quality of presentation, including the use of *good English*. Late assignments will not be accepted. Final marks may be adjusted up or down to conform with University of Toronto grading policies.

Text: Ethem Alpaydin (2004), *Introduction to Machine Learning*, The MIT Press.

Additional References:

1. Christopher Bishop (2006), *Pattern Recognition for Machine Learning*, Springer.
2. Duda, Hart and Stork (2001), *Pattern Classification (2nd edition)*, Wiley–Interscience.
3. Hastie, Tibshirani and Friedman (2001), *The Elements of Statistical Learning: Data Mining, Inference and Prediction,* Springer.
4. David Mackay (2003), *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press. Freely available on the web at `http://www.inference.phy.cam.ac.uk/mackay/itprnn/book.html`.
5. Michell (1997), *Machine Learning*, McGraw–Hill.
6. Hand, Mannila and Smyth (2001), *Principles of Data Mining,* The MIT Press.

Topics Covered: The course will start with basic methods of clustering, regression and classification, and then move on to more sophisticated methods such as neural networks, hidden Markhov models, and reinforcement learning, as well as newer methods such support vector machines, multidimensional scaling and Bayesian learning. Both supervised and unsupervised learning will be covered, as well as the evaluation of learning algorithms.

Prerequisites: The official prerequisites are CSC263/270, STA257, STA248/258/261. However, in most cases, intellectual maturity can substitute for this as long as the student has a basic knowledge of computer programming, calculus (including partial derivatives), linear algebra and probability. Experience with Matlab, Octave or R would be helpful, but is not essential.

For senior Computer Science students, the Math and Statistics requirements in the Computer Science specialist program provides sufficient preparation. For senior Math and Statistics students, CSC148 should provide sufficient Computer Science preparation (or, if the student is experienced in R or Matlab, CSC108 could be sufficient). Students lacking the official prerequisites should see the instructor.

Plagiarism and Cheating: The academic regulations of the University are outlined in the *Code of Behaviour on Academic Matters* which can be found in the UTM Calendar or on the web at http://www.utm.utoronto.ca/regcal/WEBGEN117.html.