COMP 551 - Applied Machine Learning

McGill's introductory course in machine learning

Home   Syllabus   Schedule

General Information

- Instructor: William L. Hamilton
  - Email: wlh@cs.mcgill.ca
  - Office hours: Mondays, 2-4pm, MC309
- Term: Fall 2019
- When: Mondays and Wednesdays, 4:05-5:25pm
- Where: Strathcona M-1

Course Description

The course will cover selected topics and new developments in data mining and applied machine learning, with a particular emphasis on good methods and practices for effective deployment of real systems. We will study commonly used algorithms and techniques, including linear and logistic regression, clustering, neural networks, support vector machines, decision trees and more. We will also discuss methods to address practical issues such as empirical validation, feature selection, dimensionality reduction, and error estimation.

Course content (subject to minor changes):

1. Linear regression.
2. Linear classification.
3. Performance evaluation, overfitting, cross-validation, bias-variance analysis, error estimation.

4. Regularization.
5. Naive Bayes.
6. Decision trees, regression trees and ensemble methods.
7. Support vector machines.
8. Artificial neural networks and deep learning (e.g., RNNs and CNNs).
9. Feature selection and dimensionality reduction.
10. Unsupervised learning and clustering.
11. Semi-supervised learning.
12. Generative models.
13. Applications.

Reference Materials

There is no required textbook. Lecture notes and references will be available from the course web page. The following texts can also be very useful:

1. Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press. 2014. Available free online.
2. Trevor Hastie, Robert Tibshirani and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd Edition*. Springer. 2009. Available free online.
3. Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer. 2007. Available here with some solutions here.
4. Ian Goodfellow, Yoshua Bengio and Aaron Courville. *Deep Learning*. The MIT Press. 2016. free online
5. Kevin Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press. 2012.
6. David MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press. 2003.
7. Richard Duda, Peter Hard and David Stork. *Pattern Classification. 2nd Edition*. Wiley & Sons. 2001.

Prerequisites / Anterequesites

Basic knowledge of a programming language is required. Basic knowledge of probabilities/statistics, calculus and linear algebra is required. Example courses at McGill providing sufficient background in probability are MATH-323 or ECSE-305. Some AI background

is recommended, as provided, for instance by COMP-424 or ECSE-526, but not required. Note that while the course does not have strict prerequesites, it is a graduate-level course in computer science.

Students who took COMP-652 in Winter 2013 or before CANNOT take COMP-551. Starting in Fall 2013, COMP-551 and COMP-652 were designed to avoid significant overlap; you can take either or both.

The courses is intended for hard-working, technically skilled, highly motivated students. Participants will be expected to display initiative, creativity, scientific rigour, critical thinking, and good communication skills.

Evaluation Criteria

The class grade will be based on the following components:
- Weekly online quizzes - 10%
- One written midterm examination - 35%
- Mini-project 1 - 9%
- Mini-project 2 - 12%
- Mini-project 3 - 17%
- Mini-project 4 - 17%

Quizzes

The weekly quizzes are designed to develop basic understanding of the course material and to provide some practice for the midterm. There will be two quizzes per week: one available on Tuesday-Wednesday and one Thursday-Friday. Each quiz will be available for a 48 hour period, and you will have 1 hour to complete the quiz once it is started. **We will only use the best grade from each week (i.e., if you write the quiz on Wednesday and on Friday, we will only take the best grade of the two).** There will be 10 weeks of graded quizzes (the first self-assessment quiz does not count towards you grade). Thus, in the end, each quiz that counts towards your grade is worth 1%.

Midterm

The midterm is designed to assess in-depth understanding of fundamental methods and algorithms. There is no final exam.

Projects

The projects will require reading, writing, programming and experiments to gain hands-on experience with the application of recent machine learning methods, including concepts covered in the lectures, and concepts drawn from the literature. Students will be responsible for characterizing the problem, developing methods of analysis, and presenting the results of their work. All projects will be done in groups of 3 students.

Evaluation Policy

All course work should be submitted online (details to be given in class), by 11:59pm, on the assigned due date. Late work will be automatically subject to a 20% penalty, and can be submitted up to 5 days after the deadline.

No make-up quizzes will be given.

Some of the course work will be individual, other components can be completed in groups. It is the responsibility of each student to understand the policy for each work, and ask questions of the instructor if this is not clear. It is also the responsibility of each student to carefully acknowledge all sources (papers, code, books, websites, individual communications) using appropriate referencing style when submitting work.

We will use automated systems to detect possible cases of text or software plagiarism. Cases that warrant further investigation will be referred to the university disciplinary officers. Students who have concerns about how to properly use and acknowledge third-party software should consult the course instructor or TAs.

McGill University values academic integrity. Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures (see www.mcgill.ca/students/srr/honest/) for more information).

In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded.

In the event of extraordinary circumstances beyond the University's control, the content and/or evaluation scheme in this course is subject to change.

As the instructor of this course I endeavor to provide an inclusive learning environment. However, if you experience barriers to learning in this course, do not hesitate to discuss them with me and the Office for Students with Disabilities, 514-398-6009.

COMP 551 - Applied Machine Learning is maintained by **William L. Hamilton.**