### New Course

**Program Affected?**
Y

**Program Change Form Submitted?**
N (Simple Change) - Please add this explicitly to the Complementary courses list for the following School of Computer Science programs: Major in Computer Science: Computer Games Option (in the block containing COMP 424 Artificial Intelligence), Major in Software Engineering & Honours in Software Engineering (for both, in the block: Applications Specialties), Joint Honours in Statistics and Computer Science & Joint Major in Statistics and Computer Science (in the general list of Complementaries). Note: the course is automatically offered as a Complementary for several other programs because it is at the 300-level or above.

**Subject/Course/Term**
COMP 551  
- one term

**Credit Weight or CEU's**
4 credits

**Course Activities**

<table>
<thead>
<tr>
<th>Schedule Type</th>
<th>Hours per week</th>
</tr>
</thead>
<tbody>
<tr>
<td>A - Lecture</td>
<td>3</td>
</tr>
</tbody>
</table>

Total Hours per Week: 3  
Total Number of Weeks: 13

**Course Title**

<table>
<thead>
<tr>
<th>Official Course Title</th>
<th>Course Title in Calendar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied Machine Learning</td>
<td>Applied Machine Learning</td>
</tr>
</tbody>
</table>

**Rationale**
The course has been offered three times as a Topics in Computer Science course (COMP 598). It is popular with students: 36 students in 2013, 63 students in 2014, 81 registered for Fall 2015. The topic is expected to remain of interest for several years to come. The current instructor is willing to commit to teaching the course annually for the foreseeable future, and other faculty members are able to teach the course if necessary. Other departments (eg. ECSE) have indicated a strong interest in co-listing the course, which they cannot do while it is a only a Topics in CS course. The course is related COMP-652 Machine Learning and indeed was created because the material in COMP-652 had become too rich for a single course. Since 2013, COMP-652 focuses on more advanced material with emphasis on theoretical concepts, whereas this new course covers more introductory material with an emphasis on practical aspects. The instructors consistently work together to ensure good coordination and minimal overlap. Students are able to take either or both of these courses. Other than COMP 652 (which is offered in ECSE under the course number ECSE 608), there are no other courses in the University that cover this material. The course is being proposed as 4 credits because the number of hours of
work is approximately 180, broken down as follows: lectures (~40), review of lecture material + reading (~50), study for "midterm" exam in early November (10), projects (4*20 = 80). COMP has other 5xx courses that are 4 credits, as well as other lower level courses that are 1 or 2 credits, so students have flexibility in filling their credit requirements.

Responsible Instructor
Joelle Pineau

Course Description
Selected topics in machine learning and data mining, including clustering, neural networks, support vector machines, decision trees. Methods include feature selection and dimensionality reduction, error estimation and empirical validation, algorithm design and parallelization, and handling of large data sets. Emphasis on good methods and practices for deployment of real systems.

Teaching Dept.
0155 : Computer Science

Administering Faculty/Unit
SC : Faculty of Science

Prerequisites
MATH-323 or ECSE-205 or ECSE-305 or equivalent Web Registration Blocked? : N

Corequisites

Restrictions

Supplementary Calendar Info
1. Some background in Artificial Intelligence is recommended, e.g. COMP-424 or ECSE-526, but not required.

Additional Course Charges

Campus
Downtown

Projected Enrollment
100

Requires Resources Not Currently Available
N

Explanation for Required Resources

Required Text/Resources Sent To Library?

Library Consulted About Availability of Resources?

Consultation Reports Attached?

Effective Term of Implementation
201609

File Attachments
No attachments have been saved yet.

To be completed by the Faculty

For Continuing Studies Use

Approvals Summary
Show all comments
<table>
<thead>
<tr>
<th>Version No.</th>
<th>Departmental Curriculum Committee</th>
<th>Departmental Meeting Chair</th>
<th>Departmental Curric/Academic Committee</th>
<th>Faculty SCTP</th>
<th>Version Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Submitted to Faculty Meeting for approval</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Submitted to Faculty Meeting for approval</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Submitted to Faculty Meeting for approval</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Submitted to Faculty Meeting for approval</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Submitted to Faculty Meeting for approval</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Submitted to Faculty Meeting for approval</td>
</tr>
</tbody>
</table>
COMP-551: Applied Machine Learning (4 credits)

Syllabus

Instructor

Prof. Joelle Pineau, School of Computer Science
Email: jpineau@cs.mcgill.ca
Office: McConnell 106N
Office hours: T.b.d.

Course Description

The course will cover selected topics and new developments in Data mining and 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 clustering, neural networks, support vector machines, decision trees. We will also discuss methods to address practical issues such as feature selection and dimensionality reduction, error estimation and empirical validation, algorithm design and parallelization, and handling of large datasets.

Course content (subject to minor changes):

1. Linear regression. Linear classification.
2. Performance evaluation, overfitting, cross-validation, bias-variance analysis, error estimation.
3. Naive Bayes.
4. Decision trees. Regression trees and ensemble methods.
9. Online / streaming data.
10. Data structures and Map-Reduce.

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:

Prerequisites / Antequetesites

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. If you have doubts regarding your background, please contact the instructor to discuss this.

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 exercises - 5%
- One in-class written midterm examination - 35%
- Four projects (data analysis case studies + peer reviews) - 60%

The weekly exercises will consist of quizzes (in class) or practical work (take-home) designed to develop basic understanding of the course material as we progress through the topics. These are designed to provide some practice for the midterm.

The midterm is designed to assess in-depth understanding of fundamental methods and algorithms. It will be scheduled towards the later end of the semester (mid-November). There is no final exam.

The four projects (data analysis case studies) will require reading, writing, programming and experimenting 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. Some case studies may be individual, most will be done in groups (usually of 3 or less).

We will use a online peer-review system to evaluate the data analysis case studies. Each student will be asked to read and evaluate submissions of their colleagues. The emphasis will be placed on providing constructive feedback on the methodology and presentation. Specific questions are posed to guide the evaluation. There is 1 numerical item and 1 field for comments.

Each student's project grade is follows: 90 percent from the TA's evaluation of their project (same grade for all three members of a team), 10 percent based on their submitted peer-review (full marks if they submitted, minimum 100 words, no marks otherwise; this component is assessed individually, not per team.)

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 30% penalty, and can be submitted up to 1 week after the deadline.

No make-up quizzes or midterm will be given. There are at least 8 quizzes, worth 1 percent each, for a
maximum of 5 percent (i.e. they only have to write 5). There is no make-up midterm, but in some circumstances (e.g. doctor's note), the instructor will make special arrangements (e.g. an oral exam).

**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.
Hi Josie,

As you can see from his response below, the ECSE chair suggested I write "ECSE 205 or 305". But since it already says "or equivalent", I thought that listing 205 would be sufficient. I can switch to his suggestion if you think that would be better. Thanks.

-ML

-------- Forwarded Message --------

Subject: RE: ECSE curriculum committee related question (COMP 551 and ECSE 305 prereq)
Date: Thu, 22 Oct 2015 18:47:38 +0000
From: Roni Khazaka, Prof. <roni.khazaka@mcgill.ca>
To: Michael Langer <langer@cim.mcgill.ca>

Hello Michael,

I will circulate to our curriculum committee but I expected no fundamental objection. I assume the reason for this is so that our students can take the course?

There is one point however that we need to address. ECSE305 is slotted to be retired soon and will be replaced with ECSE205 Probability and Statistics. Would it be possible make the pre-requisite ( ECSE305 OR ECSE205 )?

I will get back to you shortly with a more official answer.

Roni
present to the Faculty of Science.

Thanks very much.

Mike Langer
Hi Josie,
Here is the consultation from MATH.
Thanks.
-ML

-------- Forwarded Message --------
Subject: Re: proposed course COMP 551 (prereq MATH 323)
Date: Thu, 22 Oct 2015 14:03:32 -0400
From: jaksic@math.mcgill.ca <jaksic@math.mcgill.ca>
To: Michael Langer <langer@cim.mcgill.ca>
CC: Angela White, Ms. <angela.white@mcgill.ca>, Axel Hundemer <hundemer@math.mcgill.ca>

Hi Michael, of course, we confirm.

Vojkan
> On Oct 22, 2015, at 1:45 PM, Michael Langer <langer@cim.mcgill.ca> wrote:
> 
> Hi Vojkan,
> 
> Our department is proposing a new course next week, which lists MATH 323 as a
> prerequisite. I am attaching the proposal and syllabus. I would appreciate it if
> you could confirm that your department has no objections to us including MATH 323 as a
> prerequisite. Josie tells me that I need at least an email confirmation by next
> Tuesday.
> 
> Thanks very much.
> 
> Michael Langer
> 
> <COMP551-syllabus.html><course proposal - COMP 551 - Applied Machine Learning.pdf>