ENGG*6500: Introduction to Machine Learning
Fall 2015

University of Guelph
School of Engineering

(Revision 0: August 6, 2015)

1 INSTRUCTIONAL SUPPORT

1.1 Instructor

Instructor: Graham Taylor, Ph.D.
Office: THRN 2413
Email: gwtaylor@uoguelph.ca
Office hours: Monday 16:00-17:00

1.2 Lab Technician

Technician: Joel Best
Office: RICH 3501
Email: jbest@uoguelph.ca

2 LEARNING RESOURCES

2.1 Course Website

Course material, news, announcements, and grades will be regularly posted to the ENGG*6500 Courselink site. You are responsible for checking the site regularly.

2.2 Required Resources

There is no required textbook for the course.
2.3 Recommended Resources

There are many excellent resources both online and in print. While there is no required textbook, you may find the following to be complimentary to the course content:


Notes to accompany lectures will be posted on the course website (CourseLink) throughout the semester. You will be granted access to the website when you register for the course.

2.4 Additional Resources

Lecture Information: This course will use an “inverted classroom” model. This means that instead of being introduced to the material in a largely one-way lecture, you’ll watch the lecture as a video at home before class, and then in class, we can have a much more dynamic discussion about it. This will also allow us to focus the discussion on topics which you feel most interesting, or confusing!

I will assign recorded lectures from leading experts in the field. Details about how to view the videos will be given in the first lecture period. Each lecture period will consist of a class discussion around the content of the videos, so it is mandatory to watch the videos before the corresponding lecture period. It is expected that all students in the class contribute to the discussion. I will also occasionally assign background readings. We will discuss this material during the lecture period, so it is mandatory that you read the assigned papers before the corresponding lecture period.

Lab Information: There is no lab for the course, but the required software for assignments is installed on the computers in THRN 2336 and THRN 1313. You can use these machines when there is no lab or class scheduled in that room. If you need to verify lab availability, please contact the lab technician, Joel Best.

Assignments: Download the assignments according to the schedule given in this handout. Submission will be in hard copy and details of the required deliverables will be included in the assignment description.

Tests: The instructor will provide more information related to the midterm two weeks in advance.

Miscellaneous Information: Other information related to Machine Learning will be posted on the course website.

2.5 Communication & Email Policy

Please use lectures and lab help sessions as your main opportunity to ask questions about the course. Major announcements and/or changes will be posted to the course website. It is your responsibility to check the course website regularly.

Electronic communication should be limited to the course forum, however topics of a personal and confidential nature (e.g. marks) should be emailed to the instructor: gwtaylor@uoguelph.ca. As per university
regulations, all students are required to check their mail.uoguelph.ca e-mail account regularly: e-mail is the official route of communication between the University and its students.

3 ASSESSMENT

3.1 Dates and Distribution

Assignments: 40%
- A1 (10%), due Oct 7, in class
- A2 (10%), due Oct 26, in class
- A3 (10%), due Nov 16, in class
- A4 (10%), due Nov 30, in class

I expect you to try solving each assignment on your own. However, when you are stuck on a problem, I encourage you to collaborate with other students in the class, subject to the following rules:

1. You may discuss a problem with any student in this class, and work together on solving it. This can involve brainstorming and verbally discussing the problem, going together through possible solutions, but should not involve one student telling another a complete solution.

2. Once you solve the problem, you must write up your solutions on your own, without looking at other people’s write-ups or giving your write-up to others.

3. In your solution for each problem, you must write down the names of any person with whom you discussed it. This will not affect your grade.

4. Do not consult solution manuals or other people’s solutions from similar courses.

Detailed instruction on the content of each assignment will be distributed during the term.

Midterm Test: 20%

The midterm test will be held during the regular lecture period on Oct 28. It will test all material covered in lectures and assignments to-date.

Project: 40%

Details on the final project will be given shortly after the Thanksgiving holiday. Students can choose from one of two project options:

1. The instructor will provide a data-analysis problem and dataset. Students will solve the problem using a machine learning approach of their choosing.

2. The students will propose a data-analysis problem and dataset. The students will solve the problem using a machine learning approach of their choosing. Selecting this option requires the students to submit a 1 page proposal, and the instructor to approve this proposal (within 1 week of submission). If the instructor does not approve the proposal, students will complete the first option.
In both cases, the deliverable will be a report in the same format as a machine learning conference paper (e.g. International Conference on Machine Learning format) – you are highly encouraged to use this \LaTeX{} style file. The report (pdf only) is due Dec 3 at 23:59 and must be sent by email to gwtaylor@uoguelph.ca.

Projects are to be completed individually or in pairs. In the case of pair projects, both students will receive the same grade.

### 3.2 Course Grading Policies

**Passing Grade:** The passing grade is 65%.

**Missed Assessments:** If you are unable to meet an in-course requirement due to medical, psychological, or compassionate reasons, please email the course instructor. See the undergraduate calendar for information on regulations and procedures for Academic Consideration:
http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-ac.shtml

**Accommodation of Religious Obligations:** If you are unable to meet an in-course requirement due to religious obligations, please email the course instructor. See the undergraduate calendar for information on regulations and procedures for Academic Consideration of Religious Obligations:
http://www.uoguelph.ca/registrar/calendars/undergraduate/current/c08/c08-accomrelig.shtml

**Missed Midterm Test:** If you miss the midterm test due to grounds for granting academic consideration or religious accommodation, the weight of the missed test will be added to the project. There will be no makeup midterm.

**Late Assignments:** Due dates for the assignments are firm. Assignments must be submitted in person, within ten minutes of the start of the lecture on the specified date (i.e. by 9:40). Late submissions of assignments will not be accepted.

**Late Projects:** Similar to the assignments, the due date for the project is firm. Late projects will not be accepted.

**Clarification About Grades:** If you have questions about the grade your assignment received, please ask your instructor. Any item that is re-marked will be re-marked entirely. Therefore it is strongly suggested that you thoroughly review your entire document before making a re-marking request. Pencil-written works will not be re-marked. Re-marking requests will not be honoured more than one week after the document has been returned.

### 4 AIMS, OBJECTIVES & GRADUATE ATTRIBUTES

#### 4.1 Calendar Description

The aim of this course is to provide students with an introduction to algorithms and techniques of machine learning particularly in engineering applications. The emphasis will be on the fundamentals and not specific approach or software tool. Class discussions will cover and compare all current major approaches and
their applicability to various engineering problems, while assignments and project will provide hands-on experience with some of the tools.

4.2 Course Aims

Companies face a profound shortage of talent with deep analytical skills. This course serves as a foundation for further academic or industry work in the age of big data.

We will place special emphasis on the area of deep learning and representation learning, however we will give a broad overview of the field of machine learning. You are encouraged to explore practical applications of these techniques across a wide variety of domains.

4.3 Learning Objectives

At the successful completion of this course, the student will have demonstrated the ability to:

1. Identify and discuss general machine learning concepts such as generalization, overfitting, preprocessing, feature extraction, training/validation/test sets
2. Recognize the most common machine learning paradigms: classification, regression, density estimation, clustering, etc.
3. Implement basic machine learning algorithms
4. Debug, visualize, and evaluate a running algorithm
5. When appropriate, choose among and apply off-the-shelf machine learning toolkits to problems in your field
6. Use a compute cluster for high-throughput experimentation

4.4 Instructor’s Role and Responsibility to Students

The instructor’s role is to develop and deliver course material in ways that facilitate learning for a variety of students. Selected lecture notes will be made available to students on Coursera/D2L but these are not intended to be stand-alone course notes. During lectures, the instructor will expand and explain the content of notes and provide example problems that supplement posted notes. Scheduled classes will be the principal venue to provide information and feedback for tests and project.

4.5 Students’ Learning Responsibilities

Students are expected to take advantage of the learning opportunities provided during lectures and tutorials. Students, especially those having difficulty with the course content, should also make use of other resources recommended by the instructor. Students who do (or may) fall behind due to illness, work, or extra-curricular activities are advised to keep the instructor informed. This will allow the instructor to recommend extra resources in a timely manner and/or provide consideration if appropriate.

Presentations which are made in relation to course work—including lectures—cannot be recorded or copied without the permission of the presenter, whether the instructor, classmate or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.

1McKinsey Global Institute. Big data: The next frontier for innovation, competition, and productivity. (2011)
4.6 Relationships with other Courses & Labs

Previous Courses:

- **Probability & Statistics** (STAT*2120 or equivalent)
- **Calculus** (MATH*1210 or equivalent)
- **Linear algebra** (ENGG*2400 or equivalent)
- **Data structures & algorithms** (CIS*2520 or equivalent)

5 Teaching and Learning Activities

5.1 Timetable

Lectures:
- Monday 9:00 - 10:20 THRN 1126
- Wednesday 9:00 - 10:20 THRN 1126

5.2 Course Topics and Schedule

<table>
<thead>
<tr>
<th>Lectures</th>
<th>Lecture Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Orientation</td>
</tr>
<tr>
<td>2</td>
<td>Scientific Computing with Python</td>
</tr>
<tr>
<td>3-4</td>
<td>Perceptrons, Probability Review</td>
</tr>
<tr>
<td>5</td>
<td>Linear and logistic neurons, Linear algebra review</td>
</tr>
<tr>
<td>6</td>
<td>Backpropagation</td>
</tr>
<tr>
<td>7</td>
<td>Softmax units, Language models</td>
</tr>
<tr>
<td>8</td>
<td>Ways to deal with large # of possible outputs</td>
</tr>
<tr>
<td>9-10</td>
<td>Convolutional networks, Object recognition</td>
</tr>
<tr>
<td>11</td>
<td>Gradient descent, Practical learning strategies</td>
</tr>
<tr>
<td>12</td>
<td>Modeling sequences, Recurrent neural networks</td>
</tr>
<tr>
<td>13</td>
<td>Midterm test</td>
</tr>
<tr>
<td>14</td>
<td>Ways to improve generalization</td>
</tr>
<tr>
<td>15-16</td>
<td>Full Bayesian learning, Dropout</td>
</tr>
<tr>
<td>17</td>
<td>Mixtures of Gaussians</td>
</tr>
<tr>
<td>18</td>
<td>Hopfield networks</td>
</tr>
<tr>
<td>19-20</td>
<td>Restricted Boltzmann machines</td>
</tr>
<tr>
<td>21</td>
<td>Sigmoid belief networks</td>
</tr>
<tr>
<td>21</td>
<td>Deep belief networks</td>
</tr>
<tr>
<td>22</td>
<td>Autoencoders</td>
</tr>
<tr>
<td>23</td>
<td>Semantic hashing, Deep autoencoders for image retrieval</td>
</tr>
<tr>
<td>24</td>
<td>Course review</td>
</tr>
</tbody>
</table>
5.3 Other Important Dates

Drop Date: The last date to drop one-semester courses, without academic penalty, is Friday November 6. Two-semester courses must be dropped by the last day of the add period in the second semester. Refer to the Graduate Calendar for the schedule of dates: http://www.uoguelph.ca/registrar/calendars/graduate/current/sched/sched-dates-f10.shtml

6 Lab Safety

Safety is critically important to the School and is the responsibility of all members of the School: faculty, staff and students. As a student in a lab course you are responsible for taking all reasonable safety precautions and following the lab safety rules specific to the lab you are working in. In addition, you are responsible for reporting all safety issues to the laboratory supervisor, GTA or faculty responsible.

7 Academic Misconduct

The University of Guelph is committed to upholding the highest standards of academic integrity and it is the responsibility of all members of the University community faculty, staff, and students to be aware of what constitutes academic misconduct and to do as much as possible to prevent academic offences from occurring. University of Guelph students have the responsibility of abiding by the University’s policy on academic misconduct regardless of their location of study; faculty, staff and students have the responsibility of supporting an environment that discourages misconduct. Students need to remain aware that instructors have access to and the right to use electronic and other means of detection. The Academic Misconduct Policy is detailed in the Graduate Calendar: http://www.uoguelph.ca/registrar/calendars/graduate/current/genreg/sec_d0e1687.shtml

8 Recording of Materials

Presentations which are made in relation to course work—including lectures—cannot be recorded or copied without the permission of the presenter, whether the instructor, classmate or guest lecturer. Material recorded with permission is restricted to use for that course unless further permission is granted.

8.1 Resources

A tutorial on Academic Misconduct produced by the Learning Commons can be found at: http://www.academicintegrity.uoguelph.ca/

The School of Engineering has adopted a Code of Ethics that can be found at: http://www.uoguelph.ca/engineering/undergrad-counselling-ethics
The Graduate Calendar is the source of information about the University of Guelph’s procedures, policies and regulations which apply to graduate programs:
http://www.uoguelph.ca/registrar/calendars/graduate/current/

9 ACCESSIBILITY

The University of Guelph is committed to creating a barrier-free environment. Providing services for students is a shared responsibility among students, faculty and administrators. This relationship is based on respect of individual rights, the dignity of the individual and the University community’s shared commitment to an open and supportive learning environment. Students requiring service or accommodation, whether due to an identified, ongoing disability for a short-term disability should contact the Centre for Students with Disabilities as soon as possible.

For more information, contact CSD at 519-824-4120 ext. 56208 or email csd@uoguelph.ca or see the website: http://www.uoguelph.ca/csd/