Syllabus
CS 6923 Machine Learning
Fall 2018

Instructor: Prof. Lisa Hellerstein
Office: 2 Metrotech, Rm. 10.092
Phone: 646-997-3689
Email: lisa.hellerstein@nyu.edu
Office Hours: Wed. 2:30-4:30 p.m.

TEXTBOOKS

Required:
The main textbook for this course is *Introduction to Machine Learning*, by
Ethem Alpaydin, Third Edition. Published by MIT Press. (available online
through the NYU library). It covers many of the topics in the course clearly,
but is terse.

Optional/Recommended:
For a more intuitive and accessible approach to some of the topics in the course,
see

1. Daume: *A Course in Machine Learning* by Hal Daume (unfinished book
draft), http://ciml.info/

2. WIIT: *An Introduction to Statistical Learning with Applications in R*
   by James, Witten, Hastie, and Tibshirani. http://www-bcf.usc.edu/~gareth/ISL/


   by Witten, Frank, and Hall. Third Edition. Published by Morgan Kaufmann.

The following books are more comprehensive than the Alpaydin text and
assume that the reader has a stronger math/statistics background:

1. HTF: Hastie, Tibshirani, Friedman, *Elements of Statistical Learning*, Second
   Edition Published by Springer. (available online through the NYU library)

   Published by Springer.
**Homework:**
Homeworks will include written exercises as well as hands-on work with datasets and tools. All programming must be done in Python.

**Exams:**
The course will have a midterm and a final exam. Academic Honesty: Students may discuss homework with other students, but must write up their own solutions in their own words, and do their own coding. **DO NOT TAKE THIS COURSE IF YOU DO NOT THINK YOU CAN DO THE WORK YOURSELF, HONESTLY.** See below for NYU School of Engineering Conduct Policy.

**Prerequisites:**
Undergraduate level course in probability and statistics.

**Note:**
Machine Learning is more mathematical than most other graduate CS courses. Students often have difficulty with this course (and risk getting a grade of C or lower) if they know how to program but have not taken much math, or have not done well in their math courses. You should know about the probability density function (pdf), cumulative density function (cdf), continuous probability distributions, conditional probability, and expected values. It is also assumed that you know first-year calculus (derivatives and integrals) and the basics of linear algebra. You should also know about partial derivatives and gradients (or be prepared to learn about them on your own).

**Grading:**
The midterm will count 25% of your grade, the final 40%, and the homeworks/project 35%.

**Tentative Schedule:**
We may not get through all the material in this schedule.
<table>
<thead>
<tr>
<th>Lecture</th>
<th>Date</th>
<th>Topics</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9/6</td>
<td>Introduction, Bayesian Decision Theory</td>
<td>Alpaydin:3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HTF: 1</td>
</tr>
<tr>
<td>2</td>
<td>9/13</td>
<td>Naive Bayes, k-NN, ML Experiments</td>
<td>Alpaydin: 3,19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mitchell online: see below*</td>
</tr>
<tr>
<td>3</td>
<td>9/20</td>
<td>MLE, Bias/Variance</td>
<td>Alpaydin: Chap 4</td>
</tr>
<tr>
<td>4</td>
<td>9/27</td>
<td>Linear and Logistic Regression</td>
<td>Alpaydin: Chap 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mitchell online: see below*</td>
</tr>
<tr>
<td>5</td>
<td>10/4</td>
<td>Decision Trees and Random Forests</td>
<td>Alpaydin: Chap 9</td>
</tr>
<tr>
<td>6</td>
<td>10/11</td>
<td>Data Preprocessing, Dimensionality Reduction</td>
<td>Alpaydin: Chap 6</td>
</tr>
<tr>
<td>7</td>
<td>10/18</td>
<td>Perceptrons and Neural Nets</td>
<td>Alpaydin: Chap 11</td>
</tr>
<tr>
<td>8</td>
<td>10/25</td>
<td>Midterm</td>
<td>(to be determined)</td>
</tr>
<tr>
<td>9</td>
<td>11/1</td>
<td>Deep Learning</td>
<td>(to be determined)</td>
</tr>
<tr>
<td>10</td>
<td>11/8</td>
<td>Kernel Machines, SVMs</td>
<td>Alpaydin: Chap 13</td>
</tr>
<tr>
<td>11</td>
<td>11/15</td>
<td>Clustering</td>
<td>Alpaydin: Chap 7</td>
</tr>
<tr>
<td>12</td>
<td>11/29</td>
<td>Ensemble Methods</td>
<td>Alpaydin: Chap 16</td>
</tr>
<tr>
<td>13</td>
<td>12/6</td>
<td>Reinforcement Learning</td>
<td>Alpaydin: Chap 18</td>
</tr>
<tr>
<td>14</td>
<td>12/13</td>
<td>Catch-up and Review</td>
<td>Alpaydin: Chap 18</td>
</tr>
<tr>
<td>15</td>
<td>12/20</td>
<td>FINAL EXAM</td>
<td></td>
</tr>
</tbody>
</table>


**Moses Center Statement of Disability:**

If you are a student with a disability who is requesting accommodations, please contact New York University's Moses Center for Students with Disabilities (CSD) at 212-998-4980 or mosescd@nyu.edu. You must be registered with CSD to receive accommodations.

Information about CSD can be found at www.nyu.edu/csd. It is located at 726 Broadway on the 2nd floor.

**NYU School of Engineering Policies and Procedures on Academic Misconduct**

**Introduction:**

The School of Engineering encourages academic excellence in an environment that promotes honesty, integrity, and fairness, and students at the School of Engineering are expected to exhibit those qualities in their academic work. It is through the process
of submitting their own work and receiving honest feedback on that work that students may progress academically. Any act of academic dishonesty is seen as an attack upon the School and will not be tolerated. Furthermore, those who breach the School's rules on academic integrity will be sanctioned under this Policy. Students are responsible for familiarizing themselves with the School's Policy on Academic Misconduct.

Definition:

Academic dishonesty may include misrepresentation, deception, dishonesty, or any act of falsification committed by a student to influence a grade or other academic evaluation. Academic dishonesty also includes intentionally damaging the academic work of others or assisting other students in acts of dishonesty. Common examples of academically dishonest behavior include, but are not limited to, the following:

1. Cheating: intentionally using or attempting to use unauthorized notes, books, electronic media, or electronic communications in an exam; talking with fellow students or looking at another person's work during an exam; submitting work prepared in advance for an in-class examination; having someone take an exam for you or taking an exam for someone else; violating other rules governing the administration of examinations.

2. Fabrication: including but not limited to, falsifying experimental data and/or citations.

3. Plagiarism: intentionally or knowingly representing the words or ideas of another as one's own in any academic exercise; failure to attribute direct quotations, paraphrases, or borrowed facts or information.

4. Unauthorized collaboration: working together on work that was meant to be done individually.

5. Duplicating work: presenting for grading the same work for more than one project or in more than one class, unless express and prior permission has been received from the course instructor(s) or research adviser involved.

6. Forgery: altering any academic document, including, but not limited to, academic records, admissions materials, or medical excuses.