Advanced Machine Learning (GR5242)
Fall 2019

John P. Cunningham
jpc2181
Lydia Hsu, Elliott Gordon Rodriguez, Ding Zhou, Peter Lee
yh2692, eg2912, dz2336, jl4303

Course Syllabus

Description

The third course in the Machine Learning sequence, culminating the skills and knowledge from Statistical Computing & Introduction to Data Science (GR5206) and Statistical Machine Learning (GR5241). The course will go into depth on probabilistic models and deep learning, and will include homeworks and a substantial course project in the `python` language.

Administrative

Lectures

- Section 001:
  Monday and Wednesday, 4:10PM–5:25PM
  Location: 402 Chandler
- Section 002:
  Tuesday and Thursday, 4:10PM–5:25PM
  Location: 428 Pupin

TA Office Hours

- Mondays 5:30pm-7:00pm, Room 1025, Department of Statistics
- Tuesdays 5:30pm-7:00pm, Room 1025, Department of Statistics
- Office: Department of Statistics, 10th Floor School of Social Work, 1255 Amsterdam

Grading and Academic Integrity

- **Homework (20%).** Homework will contain both written and `python` data analysis elements. This is due online at `https://courseworks2.columbia.edu` on Wednesdays at 4:10pm. **Your two lowest homework grade will be automatically dropped at the end of the term.**
- **Midterm Exam (40%).** This will be given in class during midterm week. **Tentative midterm dates: October 17 Thursday (Section 002) and October 21 Monday (Section 001).**
- **Final Project (40%).** A substantial course project will be executed and submitted by teams of 1-4 students. **Final Project due date: December 09 Monday (both sections).**

Late Work and Regrading Policy: No late work or requests for regrades are accepted. To accommodate unexpected circumstances, we have implemented two important features:

- Your lowest two homework grades will be automatically dropped at the end of the term.
- You may submit and resubmit your homework as many times as you like up until the deadline. This means that you should submit any partial solutions as you complete them, to make sure you receive as much credit as possible for the work you have done. After the deadline, the system will not allow you to submit your homework. If you do not submit anything by the deadline, you will get a 0. **There will be no exceptions to this rule. Submit your homework early.**

Reading Material

No explicit readings will be assigned. Rather, students should use the following two books as supporting references.

- Bishop, C. *Pattern Recognition and Machine Learning.* Springer-Verlag, 2006.
- Murphy, K. *Machine Learning: a Probabilistic Perspective.* MIT Press, 2012.

Topics

The course will be taught in two halves, with the following topics:

- Part I: Probabilistic models
  - Markov and hidden markov models
  - Graphical models
  - Sampling and MCMC algorithms
  - Variational inference
  - Neural network basics
- Part II: Deep learning
  - NN software
  - Neural networks: convolutional, recurrent, etc.
  - Reinforcement learning
  - Dimension reduction and autoencoders