CS074/CS174, Winter 2015

Machine Learning and Statistical Data Analysis

Course description

This course provides an introduction to statistical modeling and machine learning. Topics include learning theory, supervised and unsupervised machine learning, statistical inference and prediction. A wide variety of algorithms will be presented, including K-nearest neighbors, naive Bayes, decision trees, support vector machines, logistic regression, K-means, mixtures of Gaussians, principal components analysis, Expectation Maximization. The course will also discuss modern applications of machine learning such as image segmentation and categorization, speech recognition, and text processing.

Administrative information

Instructor

Lorenzo Torresani | Sudikoff 109 | office hours: Mondays 1-3pm and Wednesdays 1-3pm

Teaching assistant

Haris Baig | office hours: Thursdays 1-3pm in Sudikoff 222

Course staff email

cs174@cs.dartmouth.edu

Lectures

Tue&Thu 10-11:50am | x-hour (used occasionally, see schedule page) Wed 3-3:50
Life Sciences Center 100

Lab

Sudikoff 001: Linux machines with Matlab. As an alternative, you can use Matlab on your machine by following the instructions provided here.

Textbook (recommended but not required)

Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer 2006

Grading and policies

Grading scheme

The course grade will be based 33% on the homework assignments (each of the three homework assignments will count for 11% of the final grade), 34% on the term project, and 33% on the final exam. The homeworks will require answering questions and implementing some algorithms in Matlab, but prior knowledge of Matlab is not required. During x-hours on September 16th, we will present a tutorial covering the basics of Matlab.

Late homeworks

Each student has 3 free late days to be used over the course of the term as he/she likes. Once these days are used up, any homework turned in late will be penalized 25% per late day. *No exception!* The late days can be used only for the homeworks, not for the project submissions. Any portion of a late day is counted as one full day (i.e., even one minute late counts as a full day, no exceptions!). Assignments are typically due at 11:59 pm of the due date. The code portion of each homework submission must be turned in via Canvas. The answers to technical questions must be provided in paper form and dropped in the course mailbox near the Sudikoff entrance.

Homework and Project Schedule

Howework 1: out on 1/20/2015; due on 2/5/2015.
Howework 2: out on 2/5/2015; due on 2/19/2015.
Howework 3: out on 2/19/2015; due on 3/5/2015.
Project proposal & spotlight presentation: 1/22/2015.
Project milestone: 2/17/2015.
Project final presentation: 3/10/2015.
Project final write-up: 3/15/2015.
Final exam: 3/13/2015, 8am-11am.

No-laptop policy

We have a no-laptop policy in class (texting, sleeping or engaging in other activities unrelated to the lecture is also forbidden). This policy will be strictly enforced so as to encourage active participation by all students and to avoid distracting people that are focusing on the lecture. Class attendance is highly recommended and considered necessary to do well in the course. However, it is not mandatory. If you come to class you are expected to obey this policy. A penalty of 5% will be applied to the final grade every time this policy is violated.

Auditing

Please contact the instructor if you would like to audit the course.

Academic integrity

You may discuss the assignments with other current CS074/174 students, but your submission must be entirely your own work. That is, your code and any other solutions you submit must be created, written/typed, and documented by you alone. You may not copy anything directly from another student's work. For example, memorizing or copying onto paper a portion of someone else's solution would violate the honor code, even if you eventually turn in a different answer. Similarly, e-mailing a portion of your code to another student, or posting it on-line for them to see would violate the honor code. We do encourage discussion of assignments between students, subject to these rules.

You cannot make use of any code taken from outside references for your homeworks, unless explicitly authorized to do so by the instructor. As a rule of thumb, you should treat any external code as software written by another CS074/174 student: you are not allowed to copy it or to use it as a template to implement your solution.

You are allowed to use external software for your project. However, you should clearly report the use of external code and include pointers to such software in your project write-up. The project grade will be based on the novelty of your solution/application but also on the amount of new code written by you to implement the idea. So keep this in mind when considering to use software written by someone else.

You cannot collaborate or copy in any way during the exams. The exam will be will be closed-book, closed-notes, closed laptop.

These rules will be strictly enforced and any violation will be treated seriously