CMPS 142 Syllabus, Spring 2017

Instructor: David Helmbold, dph@soe.ucsc.edu
TA: Kuan-Sung Huang , kuhuang@ucsc.edu
Lecture: TuTh 9:50 – 11:25 in Porter Acad. 148
Sections: Wed 9:20–10:25 and Th 3:25-4:25, both in Porter Acad. 148, optional sections first week
Office Hours: (tentitive) Tuesdays 11:45-12:45 (instructor); TBD (TA)
Webpage: https://courses.soe.ucsc.edu/courses/cmps142/Spring17/01
Discussion Forum: on Piazza
Text: Andrew Ng's lecture notes at http://cs229.stanford.edu/materials.html

I have not been happy with any of the published machine learning texts. This quarter we will refer to Andrew Ng's course notes from Stanford. They can be accessed from: http://cs229.stanford.edu/materials.html, Alex Holehouse has created notes from an earlier version of Andrew Ng's course available at www.holehouse.org/mlclass/. The texts:

- *Introduction to Machine Learning* by Alpaydin (easier),
- *Machine Learning* by Flach (easier),
- *Machine Learning* by Mitchell (an older book, but still with great material),
- *Artificial Intelligence: A Modern Approach* by Russell and Norvig (an AI book with some very good sections on machine learning),
- *Pattern Recognition and Machine Learning* by BIshop (a more mathematically advanced book)

are on 1-day reserve at the Science Library and can be used for additional reading.

Prerequisites: CMPS 101 (advanced data structures and algorithms), Math 23a (post-calculus linear algebra), and either AMS 131 or CE 107 (probability). Any exceptions made are at the student's own risk.

Here is a tentative syllabus for the Machine Learning class. Additional topics may be inserted and/or some topics may be skipped based on the interests of the class. The syllabus is aggressive, and it is unlikely that we will get to everything on this list.

Planned Topics:

1. Introduction and overview of machine learning and key concepts,
2. Supervised learning:
   (a) Linear regression, including instance-based techinques (Ng Part I)

(b) Regularization and Bias-Variance tradeoff (Ng Part I)

(c) Logistic regression (Ng Part II)

(d) Probability review

(e) Generative learning models, Naive Bayes (Ng Part IV)

(f) Perceptron Algorithm

(g) Support Vector Machines (Ng Part V)

(h) Decision Trees

(i) Neural networks

(j) Model selection and feature selection

(k) Perceptron Algorithm

(l) Ensemble Methods – Bagging and Boosting, random forests

3. Unsupervised learning (clustering, Mixture of Gaussians)

4. (Possibly) On-line learning

Evaluation will be based on regular group homework assignments, including a multi-week project towards the end of the course, and the final exam (at the scheduled final exam time: Tuesday June 13 at 4–7pm). Exam and homework will be weighted about equally, although I reserve the right to not pass students with very poor performance on the final regardless of their other scores.

Other Points:

- Students are responsible for their own understanding. If anything is unclear, ask questions in lecture, sections, office hours, or the class forum.
- Students should check the forum regularly (daily or at least every other day) for announcements and clarifications.
- Both lectures and the reading are important. It is important to keep up with the reading, and reading ahead is often helpful. Lectures are mandatory, and students are responsible for the material covered there.
- Due dates are firm, and it is each student's responsibility to manage their time and complete the assignments on time. Students should read and think about the assignments the day they are assigned so they can ask questions and get the help they need well before the due date.

- Written homework assignments will be done in groups of 2-3 students and each group should turn in a single set of solutions with all member's names and email accounts. Group members should rotate – *do not work with the exact same group twice!*. *All* members of the group must attempt each problem and fully understand the group's solution. It is inappropriate to simply split up the assigned problems among the group members. **All help from outside the group (from the web, books other than text, or people other than the TA or instructor) *must* be clearly acknowledged.** Presenting other's work as your own is dishonest and is called **plagiarism**. If a group is not functioning well, inform the instructor.

- Academic Honesty violations, such as submitting the un-attributed work of others, are serious issues and will result in a zero on the assignment, a lowered grade in the course, and a report to the department, and Dean of Graduate Studies. Improperly borrowed work can be as large as an entire solution or as small as a single sentence, figure, or idea. See also
  `http://www.ucsc.edu/academics/academic_integrity/undergraduate_students`

- If you qualify for classroom accommodations because of a disability, please get an accommodation Authorization from the Disability Resource Center (DRC) and submit it to me in person during my office hours or by appointment within the first two weeks of the quarter. Contact DRC by phone at 831-459-2089, or by email at drc@ucsc.edu for more information.

- If you need accommodation due to conflicts, family emergencies, illness/injury, or other difficulties, inform the instructor as soon as possible. An "incomplete" is only given by request when there is a medical, family, or similar emergency that prevents a student who has been doing clearly passing work from finishing the course.