CMPS 242 Syllabus, Fall 2018

Instructor: David Helmbold, `dph@ucsc.edu`
TA: Andrei Ignat, `aignat@ucsc.edu`; and Thomas Schmitz; `tschmitz@ucsc.edu`
Lecture: TuTh 1:30 – 3:05 in Merrill Acad. 102
Sections: TBD
Instructor office hours: (tentative) Mondays 2-3 and Thursdays 3:30-4:30
TA Office hours: TBD
Webpage: `https://courses.soe.ucsc.edu/courses/cmps242/Fall18/01`
Text: *Pattern Recognition and Machine Learning* by Bishop
Discussion Forum: on Piazza

Students are expected to have some background in the analysis of algorithms, probability, and linear algebra, as well as fluent programming skills in at least one modern programming language.

Planned Topics: 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.

1. Introduction and overview of machine learning and key concepts (ch 1), including probability, decision theory, and generative models/discriminative models/discriminants.
2. Bayesian learning and parameter estimation (ch 2)
3. Instance based learning (nearest neighbor) (ch 2.5)
4. Linear Regression (ch 3)
5. Linear classification and the Perceptron algorithm (ch 4)
6. Batch learning: Decision Trees (ch 14.4) and Artificial Neural Networks (ch 5)
7. Graphical Models (ch 8)
8. Kernels and Support vector machines (ch 6-7)
9. Clustering, EM Algorithm and K-means (ch 9)
10. Boosting (AdaBoost) (ch 14.3)
11. On-line prediction (Blum survey)
12. Concept learning, PAC model and generalization bounds (if time permits)
13. Reinforcement Learning (if time permits)

Evaluation will be based on regular group homework assignments and the **final exam** (at the scheduled final exam time: Thursday December 13 from 12 – 3pm). Exam and homework will count about equally, although I reserve the right to not pass students with very poor performance on the final regardless of their other scores.

Additional Resources: There are many additional sources of information, including:

- *All of Statistics: A Concise Course in Statistical Inference* by Larry Wasserman
- *Deep Learning* by Goodfellow, Bengio, and Courville (free web version available, has a good introduction to machine learning)
- *The Elements of Statistical Learning* by Hastie, Tibshirani, Friedman (research standard, available on-line for free)
- *Pattern Classification* by Duda, Hart and Stork, or the earlier *Pattern Classification and Scene Analysis* by Duda and Hart
- *Machine Learning* by Mitchell (a much older, but standard reference)
- *Introduction to Machine Learning* by Alpaydin (more elementary)
- *Machine Learning, an Algorithmic Perspective* by Marsland (more elementary)
- *Machine Learning a probabilistic perspective* by Murphy
- *Artificial Intelligence A Modern Approach* by Russell and Norvig has some very pertinent chapters
- Andrew Ng's Stanford course notes: `http://cs229.stanford.edu/materials.html`

Note that these other sources often use different notation. I have requested that the text, Mitchell's book, and Hastie et.al.'s book be on reserve in the Science Library.

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" can only be 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.