

## Course Information

January 17, 2001

Handout 1

**Course Web Page:** <http://www.classes.cec.wustl.edu/~cs527>.

**Prerequisite:** CS 201 and CS 241 or equivalent background is highly recommended. Some statistics background would also be useful but is not absolutely necessary.

**Time:** Monday and Wednesday, 2:30-4:00pm (Cupples II, 200)

**Professor:** Dr. Sally Goldman (Jolley 538, x5-7545, sg@cs)

**Graduate Teaching Assistants:** Dan Dooly, Qi Zhang and Yan Zhou.

**Office Hours:** Monday 12:30-2:30pm, Wednesday 10-11:30am, and by appointment. Office hours will be held in Jolley 538.

**Text Book:** The required text book for this course is:

Tom M. Mitchell *Machine Learning*. McGraw-Hill, 1997.

The reading assignments are given in the Course Calendar.

**Homework Assignments:** There will be four homework assignments where you will have 3 weeks for each. For each homework there will be a set of 10 point, 20 point and 40 point problems to choose from. You are to pick a set of problems that totals 40 points. There is a lot of flexibility provided in the homeworks. The 10 point problems are more traditional homework problems including some more theoretical problems. There will also be problems in which you can implement or modify a provided implementation of a machine learning algorithm and see what happens as changes are made in key components. These problems will be 20 or 40 point problems. In addition, you will have the option to read a paper related to one of the topics being studied for a 20 point problem. Finally, you will always have a chance to design your own homework option under the “Choose you own adventure” choice. I expect you to spend 15-20 hours on each homework.

**Late Policy** Homeworks are to be handed in at class. If for some reason you are unable to attend class, your homework must be submitted either under my door (Jolley 538) or in my box (S. Goldman) in Bryan 509 by 2:15pm on the day it is due. There will be a one point penalty per work day for late homeworks with a due date of 4:00pm for that day.

**Policy on Collaboration:** You are encouraged to discuss course material with other students, however, everything you submit should be your own work and your own writing. On the cover sheets for the 20 point and 40 point problems you are explicitly asked to acknowledge sources (people, websites, books, ...) you used for the problem. For the 10 point problems you should include this information at the top of your write-up. **Violations carry severe penalties.** If you have questions about any of this, or if you need advice for specific situations, please ask Professor Goldman.

**Exams:** There will be one in class exam and a final exam. They will be designed to be sure that everyone has a fundamental understanding of all of the topics covered.

**Final Grade Formula:** The homework assignments and exams will be weighted as follows:

- homeworks—75%
- midterm—10%
- final—15%

**Homework Write-up and Grading:** Here's some guidelines on what is expected in your homeworks.

**10 point problems:** These problems will be submitted and graded in the traditional manner. Please submit each 10 point problem on a separate sheet of paper clearly labeled with your name, the homework number, and problem number. They can be handwritten (neatly) or typed.

**20 point paper critique:** If you choose to read a paper then you will write a paper critique. Attach to the front of your critique the paper critique cover sheet. On this sheet you will see exactly how the points will be assigned. Each critique should be 2-3 pages long with reasonable font size, margins and spacing. The purpose of a paper critique is not to summarize, rather you should choose one or two points about the work that you found interesting. In your critique avoid making unsupported value judgments, like “I liked...” or “I disagreed with...” If you make judgments of this sort, explain why you liked or disagreed with the point you describe. Also, be sure to distinguish comments about the writing of the paper from the approach to the problem and the specific work that is done toward the approach. Examples of questions that you might address are:

- How does this paper advance the field of machine learning? Which important issues in the field does this paper illuminate and how?
- What problem does this paper solve, and what are the strengths and limitations of its approach?
- Does the method described seem mature enough to use in applications? Why or why not? What applications seem particularly amenable to this approach?
- What good ideas does the problem formulation, the solution, the approach or the research method contain that I could apply elsewhere?
- What would be good follow on projects and why?

You may discuss papers with fellow students, but the critique must be written by yourself. Critiques that paraphrase sections of the paper (e.g., the introduction or conclusion) will be returned with a failing grade – this is a form of plagiarism. The critique must be written in your own words.

**20 point and 40 point project:** Attach to the front of your project report the appropriate (20 or 40 point) cover sheet. On this sheet you will see exactly how the points will be assigned. For a 20 point project you are expected to submit a

2-3 page paper and for a 40 point project you are expected to submit a 4-6 page paper. Please use a reasonable font size, margins and spacing.

The goal of these projects is to learn more about the machine learning techniques we are studying by experimenting with variations of the algorithms. These projects will provide an opportunity to learn how to do publishable research in machine learning. You should consider the following criteria for evaluation as you do your project. These criteria are extracted from the instructions for reviewers at a major conference and hence have higher expectations, but nonetheless give good guidance:

**Quality:** Is the submission technically sound or are there fundamental flaws? Is the submission vacuous, polished, or sloppy? How general are the results? How far have the authors pushed them? Is the submission just a proposal or is it a careful research corpus of results? How promising is the direction?

**Significance:** How important is the work reported? Does it attack an important or difficult problem or a peripheral or simple one? Does it advance the state of the art?

**Novelty:** Has this or similar work been reported? Are the problems and approaches completely new? Is this a novel combination of familiar techniques? Are the authors aware of relevant existing research or are they reinventing the wheel? Is this research well-known in another community?

**Clarity:** Is the submission well organized and well-written? Is there sufficient detail presented so that it is clear what has actually been done? From the details in the paper could one replicate the results?

When writing your paper you should include:

1. An introduction that defines the problem, motivates its importance, sketches the approach, and previews the results/conclusions of the paper.
2. A description of the approach in detail that illustrates it by means of clear concrete examples and points out whatever issues are important about your method. This portion may consist of a few sections (i.e., method, results, discussion). The organization will depend on the project, but should be well motivated and easy to follow.
3. A related work description that surveys one or two related papers and explains what they have to do with the project. Perhaps you were using an idea from them, or maybe it looked at the same problem but used a different approach. Whether this material fits before, after, or interspersed with the project description will depend on the project. (This is not expected in the 20 point project and is worth 3 points in the 40 point project.)
4. A conclusion that summarizes your results, evaluates the approach that you used and identifies areas for future work.
5. A bibliography that lists the references cited in the paper (if applicable).