Assignment 1

Assigned: Tue Aug 28, 2001
Due: Thu Aug 30, 2001

Before we dive into the technical details of machine learning, I would like us to get a better appreciation for the sorts of applications where machine learning techniques are useful. To introduce you to the field, your first assignment is to report on a problem that has been solved using machine learning techniques. I do not want you to report on the techniques themselves or any details of the solution. I simply want you to report on the problem. You should address the following points:

- **What is the task?** (You can characterize the task in terms of the input-output behavior of the system. For example, the speech recognizer I will demonstrate in class extracts input features from an acoustic signal, and produces as output a probability distribution over a set of predefined alternative classifications of the signal.)
- **What is the training data?** (How much data is used to train the system? Where does the data come from?)
- **How was the system's performance evaluated?** (Was the system tested in the real world? Are there quantitative reports of accuracy of the system?)
- **What do you think about the work?** (Do you see this as a sensible application? Does it have practical implications? Can you think of ways in which the work could be extended—to other domains, to bigger problems, to related problems?)
- **What is your information source?** (Give a web URL or citation to a paper.)

You may not be able to answer all of these questions, depending on how detailed your source of information is. I would like you to hand in a *brief* report on your findings, no more than one page in length. One or two paragraphs should suffice. As with all assignments for this course, I expect the report to be typed and spell checked.

The assignments are due *at the start of class* on August 30. The reason for this is that we will spend one or more class periods on oral presentations of your findings. You are expected to come to class prepared to give a short report (2-5 minutes) on your findings. Beyond the sharing of information, the goal of these oral presentations is to allow me to get to know you, to allow you to get to know one another, and as a stimulus for class discussion.

For this assignment, as with all assignments for this course, you should work on your own. If you have friends in the class, please avoid picking the same example problem that they have picked.

Fortunately, the WWW is a rich source of information about machine learning. I have links to many different resources on the web concerning machine learning on the class home page (www.cs.colorado.edu/~mozer/courses/ugrad-ml). Start there. You can also look through journals at the library that publish papers on machine learning. The most promising journals are: *Machine Learning*, *Neural Computation*, *Neural Networks*, and *AI Magazine*. You should find links to some of these journals (most of which make the papers available electronically on the web if you are accessing the web from within CU) on the class home page. If you find a brief description of a research project on the web, you may also be able to find a technical paper on the web site. You can download and print these papers to find more details of the work.