CSE 598E/ STAT 598A
Project One
Due 19 February 2007

The goal of this project is to construct and apply a boosted tree structured classifier such as those discussed in Chapter 10 and in §9.2 of the text. You are strongly encouraged to use a data set related to your research, but you may also choose one of the six data sets listed on the class website http://www.stat.psu.edu/~barlow/csestat598/index.html. You may implement it in the programming language of your choice--I will be able to understand your project best if it is implemented in MATLAB, C,C++, FORTRAN, or a language with similar structures.

These are the same six data sets used by Professor Li last semester. In fact, this should be built upon the Decision Trees developed for Project Three from last semester. You are to form groups of two or three. If you were not in the class last semester, I strongly encourage to have a partner that was in the class. If you were in the class last semester, I may encourage you to welcome a co-worker who did not.

The basic requirements of the project are:

1. Implement some version of boosting trees. The most straightforward way would be to start with your Project Three from last semester and build AdaBoost.M1 on top of it.

2. Apply your classifier to a data set and evaluate it.

3. Write a report. The report should include: (1) a psuedo-code of your algorithm (in English is fine); (2) Describe the results from your data set; (3) Explain the contributions of each group member.