

CSE 417A: Homework 6

Due: December 2, 2014

Notes:

- Please keep in mind the collaboration policy as specified in the course syllabus. If you discuss questions with others you **must** write their names on your submission, and if you use any outside resources you **must** reference them. **Do not look at each others' writeups, including code.**
- Instructions for how to get files from the SVN repository are available on the course website and on Piazza.
- Homework (in hardcopy) is due **at the beginning of lecture**. In addition, your code submissions must also be timestamped before lecture begins.
- Please comment your code properly.
- There are 6 problems on 1 page in this homework.

Problems:

1. (50 points) Implement AdaBoost using decision stumps learned using information gain as the weak learners, and apply this to the one-vs-three and ~~one-vs-five~~ three-vs-five problems described in HW5 on the `zip.train` and `zip.test` data. In order to do this, you should complete the stub `AdaBoost` function available in your SVN repository. Graphically report the training set error and the test set error as a function of the number of weak hypotheses, and summarize and interpret your results.
2. (10 points) AIMA Problem 18.3
3. (10 points) AIMA Problem 18.11
4. (10 points) AIMA Problem 18.17
5. (10 points) AIMA Problem 18.19
6. (10 points) Suppose your input data consists of the following (x, y) pairs:

$(3, 5); (5, 6); (7, 9); (2, 11); (3, 8)$

What value of y would you predict for a test example where $x = 3.2$ using (a) the 3-nearest neighbors average and (b) the 3-nearest neighbors linear regression?