Homework #1

TA in charge: Chen-Wei Hung

RELEASE DATE: 09/13/2010

DUE DATE: 09/27/2010, 4:00 pm IN CLASS

TA SESSION: TBA (please check the website)

Unless granted by the instructor in advance, you must turn in a hard copy of your solutions (without the source code) for all problems. For problems marked with (*), please follow the guidelines on the course website and upload your source code to designated places.

Any form of cheating, lying, or plagiarism will not be tolerated. Students can get zero scores and/or fail the class and/or be kicked out of school and/or receive other punishments for those kinds of misconducts.

Discussions on course materials and homework solutions are encouraged. But you should write the final solutions alone and understand them fully. Books, notes, and Internet resources can be consulted, but not copied from.

Since everyone needs to write the final solutions alone, there is absolutely no need to lend your homework solutions and/or source codes to your classmates at any time. In order to maximize the level of fairness in this class, lending and borrowing homework solutions are both regarded as dishonest behaviors and will be punished according to the honesty policy.

You should write your solutions in English with the common math notations introduced in class or in the problems. We do not accept solutions written in any other languages.

1.1 Learning Exercises

In the following exercises, you need to make your arguments *convincing* to get the points.

(1) (10%) Do Exercise 1.1 of LFD.

(2) (10%) Do Exercise 1.5 of LFD.

1.2 Perceptron Learning Algorithm

(1) (5%) Do Exercise 1.3-1 of LFD.

(2) (5%) Do Exercise 1.3-2 of LFD.

(3) (5%) Do Exercise 1.3-3 of LFD.

1.3 Experiments with Perceptron Learning Algorithm (*)

(1) (10%) Generate a data set of size 20 as directed by Exercise 1.4 of LFD, and plot the examples $\{(\mathbf{x}_n, y_n)\}$ as well as the target function f on a plane. Be sure to mark the examples from different classes differently, and add labels to the axes of the plot.

(2) (10%) Run Perceptron Learning Algorithm (PLA) on the data set above. Report the number of updates that PLA takes before converging. Plot the examples $\{(\mathbf{x}_n, y_n)\}$, the target function f, and the final hypothesis g in the same figure. Comment on whether f is close to g.

(3) (5%) Repeat everything in (2) with another randomly generated data set of size 20. Compare your results with (2).

(4) (5%) Repeat everything in (2) with another randomly generated data set of size 100. Compare your results with (2).

(5) (5%) Repeat everything in (2) with another randomly generated data set of size 1000. Compare your results with (2).

(6) (10%) Modify your PLA such that it takes $\mathbf{x}_n \in \mathbb{R}^{10}$ instead of $\mathbb{R}^2$. Randomly generate a data set of size 1000 with $\mathbf{x}_n \in \mathbb{R}^{10}$ and feed the data set to the algorithm. How many updates does PLA take to converge?

(7) (10%) Repeat PLA on the same data set of (6) for 100 experiments. In the iterations of each experiment, pick $\mathbf{x}(t)$ randomly instead of deterministically. Let u_k be the number of updates that PLA takes to converge within the k-th experiment. Plot a histogram of u_k.

1.4 More on Perceptron Learning

Dr. Learn observes that PLA seems to always converge and wants to prove this fact. The doctor comes up with the following idea. Assume that the target function $f(\mathbf{x}) = \text{sign}(\mathbf{w}^* \bullet \mathbf{x})$. That is, $y_n = \text{sign}(\mathbf{w}^* \bullet \mathbf{x}_n)$ for all $n = 1, 2, \cdots, N$.

(1) (5%) Using the fact that $\mathbf{w}(t)$ classifies $\mathbf{x}(t)$ incorrectly, prove $\mathbf{w}(t+1) \bullet \mathbf{w}^* \geq \mathbf{w}(t) \bullet \mathbf{w}^*$.

(2) (5%) Aha, the doctor thinks. The results indicate that after each update, $\mathbf{w}(t+1)$ would be closer to $\mathbf{w}^*$ than $\mathbf{w}(t)$ is. That is, when taking $t \to \infty$, the resulting hypothesis g would be very close to f and hence can also classify all examples perfectly (i.e., converges). Is the argument of the doctor correct? Why or why not?