Ethics Pledge

Consistent with the above statements, all homework exercises, tests and exams that are designated as individual assignments MUST contain the following signed statement before they can be accepted for grading.

I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination. I further pledge that I have not copied any material from a book, article, the Internet or any other source except where I have expressly cited the source.

Signature: Haodong Zhao Date: Apr 10th. 2019

Please note that assignments in this class may be submitted to

www.turnitin.com, a web-based anti-plagiarism system, for an evaluation of their originality.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Reading review**

**A Brief Introduction to Boosting**

Boosting is a way to improve the accuracy of any given learning algorithm. This article focuses on the enhancement algorithm: AdaBoost.

One of the main ideas of the AdaBoost algorithm is to maintain a distribution or set of weights on the training set. Initially, all weights are equal, but in each round, the weight of the incorrectly classified example increases, so that weak learners are forced to focus on the hard examples in the training set.

**Analyzing the training error**

The most basic theoretical feature of AdaBoost is that it can reduce training errors and is adaptable. The weak learning algorithm can be effectively converted into a strong learning algorithm.

**Generalization error**

Boosting is particularly aggressive at reducing the margin (in a quantifiable sense) since it concentrates on the examples with the smallest margins (whether positive or negative)

**Multiclass classification**

Ways to extend AdaBoost to multiple classes:

1. AdaBoost.M1
2. AdaBoost.MH
3. AdaBoost.M2

**The advantages of AdaBoost:**

1. Fast
2. Simple
3. Easy to program
4. It has no parameters to tune
5. It requires no prior knowledge about the weak learner
6. Can be flexibly combined with any method for finding weak hypotheses

In the article, Schapire and Singer's experiments prove that in nearly all of these experiments and for all of the performance measures tested, boosting performed as well or significantly better than the other methods tested. AdaBoost also has the ability to identify outliers.