Machine Learning

V. Adamchik CSCI 567 Spring 2019

Discussion 9 University of Southern California

Boosting

Problem 1

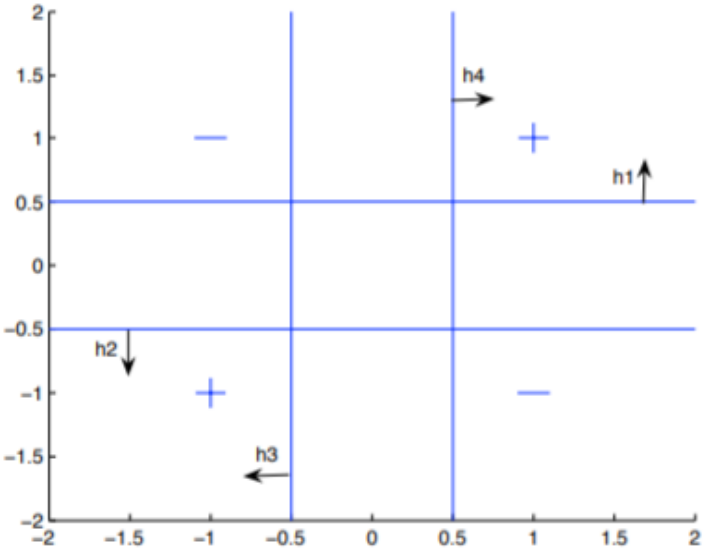
T/F. AdaBoost will eventually give zero training error regardless of the type of weak classifier it uses, provided enough iterations are performed.

T/F. In AdaBoost weights of the misclassified examples go up by the same multiplicative factor.

T/F. A weak learner with less than 50% accuracy does not present any problem to the AdaBoost algorithm.

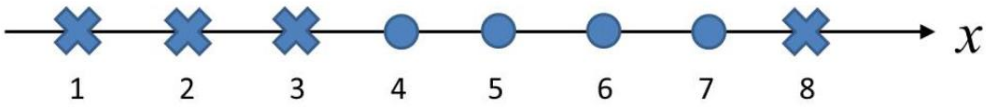
Problem 2

Consider the four binary classifiers below. The arrow means that the corresponding classifier classifies every data point in that direction as +. Prove that there are no weights $\beta_1, \ldots, \beta_4$, that make the ensemble $\Sigma\beta_i h_i$ classifier consistent with the data.



Problem 3

Imagine running AdaBoost with a 1-dimensional training set of 8 examples as shown



Circles mean y = +1 and crosses mean y = -1. The number under each example is its x coordinate.

The base classier set H consists of all decision stumps parameterized by a pair (s, b) such that

$$h_i(x) = \begin{cases} s, & \text{if } x > b \\ -s, & \text{otherwise} \end{cases}$$

Problem 3 (continue)

1. Which of the following is the possible parameter for h_1?

(A) (s, b) = (+1, 3.5)
(B) (s, b) = (-1, 3.5)
(C) (s, b) = (+1, 7.5)
(D) (s, b) = (-1, 7.5)

2. Suppose we run AdaBoost for two rounds and observe that β_1 and β_2 are both positive but not equal. Is it possible that the final classier H after these two rounds has zero training error? Why or why not?

Problem 4

Assume that the weighted classification error

$$\varepsilon_t \leq \frac{1}{2} - \gamma$$

for some constant $\gamma \in \left(0, \frac{1}{2}\right)$. Prove that the normalization factor Z_t satisfies:

$$Z_t = \sum_{n=1}^{N} D_t(n) \exp(-\beta_t y_n h_t(x_n)) \leq \sqrt{1 - 4\gamma^2}$$