15-854 Machine Learning Theory

Homework # 2 **Due: February 4, 1998**

Exercises:

1. **A bad modification to Winnow.** Suppose that we modify Winnow so that it doubles its weights on positive examples even when it did *not* make a mistake. Show how this can cause the algorithm to make an unbounded number of mistakes, even if all examples *are* consistent with some disjunction.

2. **Tracking a moving target.** Here is a variation on the deterministic Weighted-Majority algorithm, designed to make it more adaptive.

    1. Each expert begins with weight 1 (as before).
    2. We predict the result of a weighted-majority vote of the experts (as before).
    3. If an expert makes a mistake, we penalize it by dividing its weight by 2, but *only* if its weight was at least 1/4 of the average weight of experts.

    Suppose that we have two experts and suppose that on the first 50 examples the first expert is correct and the second is wrong, and then on the next 50 examples the second expert is correct and the first is wrong.

    How many mistakes does the original Weighted-Majority algorithm make (assume it always makes the wrong choice when there is a tie)?

    How many mistakes does this version of the Weighted-Majority algorithm make (assume it always makes the wrong choice when there is a tie)?

Problems:

3. **Tracking a moving target, contd.** For the algorithm of exercise 2, prove that in any contiguous block of trials (e.g., the 51st example through the 77th example), the number of mistakes made by the algorithm is at most $O(m + \log n)$, where m is the number of mistakes made by the best expert *in that block*, and n is the total number of experts.

    (If you get stuck, there is a proof in the Littlestone-Warmuth paper, but my guess is that it is easier to prove it yourself. If you do read their proof, you must still prove it in your own words.)