Homework Assignment 7

Due date: Wednesday, 12/7 (hardcopy in class)

1. (Variation of problem 7.7 from Mitchell’s text page 228)

Consider the hypothesis class of $H_{rd2}$ of “regular depth 2 decision trees” over $n$ Boolean variables. A “regular depth 2 decision tree” is a full depth 2 decision tree (having 3 internal nodes and 4 leaves all at depth 2) and where the left and right children of the root split on the same variable test.

(a) As a function of $n$, how many syntactically distinct trees are there in $H_{rd2}$?

(b) Give an upper bound on the number of examples needed to PAC learn $H_{rd2}$ with error $\epsilon$ and confidence $\delta$.

(c) Consider applying the Weighted Majority (WM) algorithm as follows. Each syntactically distinct tree is used as one advisor for the WM algorithm. In each step, advisors who predicted wrongly are penalized by having their weight multiplied by 0.5 (exactly as in the analysis in class).

What is the run time of the algorithm for each prediction step in terms of $n$ (please use $O()$ notation)? Provide an upper bound on the number of mistakes the WM will make in terms of $n$, and the best performance of any $H_{rd2}$ on the same sequence.