Midterm 2 Practice Problems

1. AdaBoost with Decision Stumps. Say I am at iteration $t$ of AdaBoost with $n = 6$. I train a classifier with the current weights (shown for each example below) and this is the resulting decision tree:

(a) If I use a threshold of 0.5, what are the labels of each leaf? In other words, if a new example had `weather="sun"`, would I classify it as $+1$ or $-1$? And if the example had `weather="rain"`?

(b) What is the weighted training error $\epsilon_t$ for this classifier?

(c) We compute the score of this model as $\alpha_t = \frac{1}{2} \ln \left( \frac{1-\epsilon_t}{\epsilon_t} \right)$. What is the intuition behind using this scoring function?

(d) (outside class) What challenges would we need to overcome to use AdaBoost in a multi-class classification setting?
2. *Ensembles reducing error.* Say I have $T = 5$ classifiers and each one has an $r = \frac{1}{3}$ chance of being incorrect. If I run all of them on a new example, the number of votes for the wrong class could be $R = 0, 1, 2, 3, 4,$ or $5$. Of these options, which would result in an incorrect classification overall? Compute $P(R = k)$ for each of these options, then use this result to compute the overall probability of an incorrect classification.

3. *Discriminative vs. generative models.* Out of the methods: $k$-Nearest Neighbors, Logistic Regression, Naive Bayes, Random Forests, AdaBoost, Support Vector Machines, Neural Networks, $K$-means, and Gaussian Mixture Models, which ones are discriminative and which ones are generative?

4. In which situations would you use a ROC curve as an evaluation metric?

5. With $n = 2$ training examples, how many unique datasets can I generate with bagging? What about $n = 3$?