# Advance Machine Learning
# Quiz: Boosting
## February 14th, 2017

**NAME:**_____

1. Consider the following dataset $(x_1 = (0, -1), y_1 = -1), (x_2 = (1, 0), y_2 = 1), (x_3 = (-1, 0), y_3 = 1)$ and $(x_4 = (0, 1), y_4 = -1)$ with corresponding weights $w_1 = \frac{2}{7}, w_2 = \frac{1}{7}, w_3 = \frac{2}{7}, w_4 = \frac{2}{7}$.

   (a) **(2 pts)** Compute the weighted misclassification rate of the following classifier:
   $$g_{1,0.5}(x) = \begin{cases} 1, & \text{if } x^1 > 0.5 \\ -1, & \text{otherwise} \end{cases}$$
   This stump misclassifies $x_3$ and gives $\frac{w_3}{\sum_{i=1}^{4} w_i} = \frac{2}{7}$ misclassification rate.

   (b) **(2 pts)** Find a stump that minimizes the weighted misclassification rate.
   Here is one solution that misclassification $x_2$ and gives $\frac{1}{7}$
   $$g(x) = \begin{cases} 1, & \text{if } x^1 < -0.5 \\ -1, & \text{otherwise} \end{cases}$$

2. **(1.5 pts)**What are the strategies for dealing with overfitting in gradient boosting? Explain at least three strategies.

   I accepted many answers here. I was hopping to hear about how you do regularization: shrinkage, number of trees, subsampling, tree depth parameter.

3. Consider the following training set as input to adaboost: $(A, -1), (B, -1), (C, 1), (D, 1), (E, -1)$, and $(F, 1)$. After the first iteration the base classifier misclassifies points $E$ and $F$. After the second iteration the base classifier misclassifies points $A$ and $F$ (See the Adaboost algorithm)

   (a) **(2 pts)** What are the weights of each point in the training set after one iteration?
   $err_1 = \frac{1}{3}$, $\alpha_1 = \log \frac{2/3}{1/3} = \log 2$ so
   $w_A = w_B = w_C = w_D = \frac{1}{6}$ and $w_E = w_F = \frac{1}{3}$

   (b) **(2.5 pts)** What is the prediction the algorithm would give to point $A$ using two rounds of Adaboost (M=2)?
   $err_2 = \frac{1/6+1/3}{8/6} = \frac{3}{8}$ , $\alpha_2 = \log \frac{1-3/8}{3/8} = \log \frac{5}{3}$
   $G(A) = sign(\log 2 G_1(A) + \log \frac{5}{3} G_2(A)) = sign(-\log 2 + \log \frac{5}{3}) = -1$