

Practice Problems for Boosting. Advanced Machine Learning

1. Consider the following dataset.

$x = (x_1, x_2)$	y	weights
(0, 0)	-1	2
(0, 1)	-1	2
(1, 1)	-1	3
(1, 3)	1	1
(2, 1)	1	2

- (a) Compute the weighted classification error of the following classifier. Which points are misclassified?

$$g(x) = \begin{cases} 1, & \text{if } x_2 \geq 0.5 \\ -1, & \text{otherwise} \end{cases}$$

- (b) Find a stump that minimizes the weighted classification error. What is the best weighted classification error?

2. Consider the following table with 4 iterations of Adaboost and 4 observations. Write the **expression** for the prediction that the final algorithm gives to x_1 .

m	err	α	$T_m(x_1)$	$T_m(x_2)$	$T_m(x_3)$	$T_m(x_4)$
1	0.25	1.01	-1	-1	1	-1
2	0.167	1.6	-1	1	-1	-1
3	0.1	2.2	1	1	1	-1
4	0.056	2.8	-1	1	1	1

3. Consider the following (exponential) loss function for binary classification with $y \in \{-1, 1\}$.

$$L(y, f) = e^{-yf}$$

- (a) What is the value of the loss function for (x, y) when y has a negative label and $f(x) = -\frac{1}{2}$.
- (b) Write the pseudo-residual in gradient boosting for this loss.
- (c) Why is binary cross entropy loss considered to be better than the exponential loss?
4. What is stagewise additive modeling? What is its relationship with gradient boosting?
5. Write the pseudocode for gradient boosting for MSE and MAE.
6. What is the β that minimizes the function $G(\beta) = (1 - err)e^{-\beta} + err \cdot e^{\beta}$