

STT873 Homework 1. Due September 27.

1. Consider a classification problem: $X \in R$, $Y = 0$ or 1 . Denote $\eta(x) = P(Y = 1|X = x)$. Then show that under 0-1 loss the best set $C^* = \{x : \eta(x) > 0.5\}$.

2. In the previous problem show that the risk $P(Y \neq I_{\eta(X)>0.5})$ can be written as $E \min(\eta(X), 1 - \eta(X))$.

3. Consider the construction of a support vector machine in the case of 2 overlapping classes. Show that the solution has the form $\hat{\beta} = \sum_{i=1}^n \hat{\alpha}_i Y_i X_i$.

4. Given the dataset $\{(-1, 1), (0, 0), (1, 0), (3, 1), (5, 1)\}$ and the weak learners $C_0 = \{x = -1\}$, $C_1 = \{x \geq 1\}$, $C_2 = \{x = 5\}$, do 3 iterations of the AdaBoost and obtain the boosted classifier.