Homework2

Problem 1:

Error rate is: 21.89 %

Problem 2:

At any given x we may incur one of the 3 possible costs if the prediction is incorrect:

1. $P(h(x) = 0, Y = 1 \mid x) = P(h(x) = 0 \mid x) * \eta(x)$
2. $P(h(x) = 1, Y = 0 \mid x) = P(h(x) = 1 \mid x) * (1 - \eta(x))$
3. $\theta * P(h(x) = NULL \mid x) = \theta * P(h(x) = NULL \mid x)$

The classifier should minimize the expected cost of prediction. In other words, the classifier should classify in such a way so that at any given x, cost is, $\min(\theta, \eta(x), 1 - \eta(x))$. Since θ lies between 0 and 0.5, the classifier can be written as:

$$h(x) = \begin{cases} 0 \; if \; \eta(x) < \; \theta \\ 1 \; if \; (1 - \eta(x)) < \; \theta \\ NULL \; if \; \theta < \; \min(\eta(x), 1 - \eta(x)) \end{cases}$$