

Exercise 3

Probabilities of characters in normal text

$$\Pr(Y = j) = 0.015$$

$$\Pr(Y = k) = 0.045$$

$$\implies \Pr((Y = j) \vee (Y = k)) = \Pr(Y = j) + \Pr(Y = k) = 0.06$$

Part (a)

The most basic classifier we can choose is the classifier on based only on prior probabilities. Since $\Pr(Y = j) < \Pr(Y = k)$, let this classifier be $f(x) = k \forall x \in X$.

$$\text{Calculating risk: } \mathcal{R}(f) = \Pr(Y = j \mid ((Y = j) \wedge (Y = k))) = \frac{\Pr(Y=j)}{\Pr((Y=j) \vee (Y=k))} = \frac{1}{4} < 0.3.$$

Thus, the most basic classifier has a risk lower than the classifier being sold. So we would not buy the classifier being sold.

The lowest probability of error if we do not look at the input data is 0.25.

Part (b)

Again let us choose the most basic classifier, which just predicts $\sim j$ without regard to the input data.

$$\mathcal{R}(f) = \Pr(Y = j) = 0.015 < 0.2$$

By the same logic of Part (a), we would not buy the classifier being sold.

Exercise 5

Notations

S = message is actually spam

L = message is actually legitimate

\hat{S} = message is classified as spam

\hat{L} = message is classified as legitimate

Part (a)

From the given data, we fill up the confusion matrix.

	\hat{S}	\hat{L}	
S	0.51 (= 0.85×0.6)	0.09 (= $0.6 - 0.51$)	0.6
L	0.02 (= 0.4×0.05)	0.38 (= $0.4 - 0.02$)	0.4 (= $1 - 0.6$)
	0.53 (= $0.51 + 0.02$)	0.47 (= $1 - 0.53$)	1

From the above table, we can conclude the following¹:

- False positive = $\Pr(\hat{S} \cap L) = 0.02$
- False negative = $\Pr(\hat{L} \cap S) = 0.09$
- Average error = $\Pr(\hat{S} \cap L) + \Pr(\hat{L} \cap S) = 0.02 + 0.09 = 0.11$

Part (b)

- A classification algorithm with zero false positive rate would be a classifier which simply classifies everything as legitimate.
- A classification algorithm with zero false negative rate would be a classifier which simply classifies everything as spam.

¹Definitions from "Counting performance measures (2)" page of the slides

Part (c)

