Problem 8.1

Consider a nearest neighbors classifier with just two training tokens:

\[
\vec{x}_0 = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \quad y_0 = -1, \quad \vec{x}_1 = \begin{bmatrix} u_1 \\ v_1 \end{bmatrix}, \quad y_1 = +1
\]

Prove that a nearest neighbor classifier, constructed from this training dataset, gives you a linear dichotomizer. Find the vector \( \vec{w} \) and the offset \( b \) in terms of the variables \( u_0, v_0, u_1, v_1 \).

Problem 8.2

Consider the following Bayesian classifier:

\[
p_Y(0) = \pi_0, \quad p_Y(1) = 1 - \pi_0
\]

Suppose that \( \mathcal{X} = (\mathbb{R}_+)^d \), that is, \( \vec{x} \) is a \( d \)-dimensional vector, all of whose elements are non-negative. Within this domain, the likelihoods are determined by the parameter vectors \( \vec{u} = [u_1, \ldots, u_d]^T \) and \( \vec{v} = [v_1, \ldots, v_d]^T \) as

\[
p_{X|Y}(\vec{x}|0) = c_0 e^{-\vec{u}^T \vec{x}}, \quad p_{X|Y}(\vec{x}|1) = c_1 e^{-\vec{v}^T \vec{x}}
\]

where \( c_0 = \prod_{k=1}^{d} u_k \) and \( c_1 = \prod_{k=1}^{d} v_k \) are normalizing constants. Show that this Bayesian classifier is actually a linear dichotomizer. Find \( \vec{w} \) and \( b \) in terms of \( \vec{u}, \vec{v}, \) and \( \pi_0 \).

Problem 8.3

A minimum-risk classifier is a generalized Bayesian classifier which, instead of minimizing the probability of error, minimizes some other type of expected loss function (“risk” means “expected loss”). For example, consider the following loss function:

\[
\mathcal{L}(y, \hat{y}) = \begin{cases} 
0 & y = \hat{y} \\
1 & y = -1 \text{ but } \hat{y} = +1 \quad \text{(false alarm)} \\
C & y = +1 \text{ but } \hat{y} = -1 \quad \text{(miss)}
\end{cases}
\]

The minimum-risk classifier is defined by

\[
h(x) = \arg \min E [\mathcal{L}(y, h(x))]
\]

Consider the table \( p_{X,Y}(x, y) \) on lecture slide 21. Depending on the value of \( C \), the minimum-risk classification rule might result in up to five different classification functions \( h(x) \). List all five of the classification functions, and say for what values of \( C \) each one is a minimum-risk classifier.

Problem 8.4

The Bayes risk is defined by

\[
\mathcal{R}_{\text{Bayes}} = \min E [\mathcal{L}(y, h(x))]
\]

Find the Bayes risk, as a function of \( C \), for each of the five classifiers from problem 3.