Collab Worksheet 3

CS440/ECE448, Spring 2021

Week of 2/19 - 2/24, 2021

Question 1

You're on a phone call with your friend, trying to help figure out why their computer won't start. There are only two possibilities, $Y = \text{CPU}$, or $Y = \text{PowerSupply}$, with prior probability $P(Y = \text{CPU}) = 0.3$.

You ask your friend whether the computer makes noise when they try to turn it on. There are two possibilities, $E = \text{quiet}$, and $E = \text{loud}$. You know that a power supply problem often leaves a quiet computer, but that the relationship is stochastic, as shown:

$$P(E = \text{noise}|Y = \text{CPU}) = 0.8, \quad P(E = \text{noise}|Y = \text{PowerSupply}) = 0.4$$

(a) What is the MAP classifier function $\hat{Y}(E)$, as a function of E?

(b) What is the Bayes error rate?

(c) CPU damage is more expensive than power supply damage, so let's define a false alarm to be the case where your classifier says $\hat{Y} = \text{CPU}$, but the actual problem is $Y = \text{PowerSupply}$. Under this definition, what are the false-alarm rate and missed-detection rate of the MAP classifier?

Question 2

Consider the following binary logic function:

$$y = \neg\left((x_1 \wedge \neg x_2 \wedge \neg x_3) \vee (x_1 \wedge x_2 \wedge \neg x_3)\right)$$

Convert truth values to numbers in the obvious way: let $x_i = 1$ be a synonoym for $x_i =$ **True**, and let $x_i = 0$ by a synonym for $x_i =$ **False**. Let $x = [x_1, x_2, x_3]^T$ and $w = [w_1, w_2, w_3]^T$, let $x^T w$ denote the dot product of vectors x and w, and let $u(\cdot)$ denote the unit step function. Find a set of parameters w_1, w_2, w_3 and b such that the logic function shown above can be computed as $y = u(w^T x + b)$.

Question 3

We want to implement a classifier that takes two input values, where each value is either 0, 1 or 2, and outputs a 1 if at least one of the two inputs has value 2; otherwise it outputs a 0. Can this function be implemented by a linear classifier? If so, construct a linear classifier that does it; if not, say why not.

Question 4

Consider a problem with a binary label variable, Y, whose prior is $P(Y=1)=0.4$. Suppose that there are 100 binary evidence variables, $E=[E_1,\ldots,E_{100}]$, each with likelihoods given by $P(E_i=1|Y=0)=0.3$ and $P(E_i=1|Y=1)=0.8$ for $1\le i\le 100$.

(a) Specify the classifier function, $\hat{y}(e)$, for a naive Bayes classifier, where $e=[e_1,\ldots,e_{100}]$ is the set of observed values of the evidence variables.

(b) The naive Bayes classifier can be written as

$$\hat{y}(e)=\begin{cases}1 & w^Te+b>0\\ 0 & \text{otherwise}\end{cases},$$

where w^Te is the dot product between the vectors w and e. Find w and b (write them as expressions in terms of constants; don't simplify).