CSDS 440: Assignment 3

Shaochen (Henry) ZHONG, sxz517

Mingyang TIE, mxt497

Due on 09/25/2020, submitted early on 09/18/2020

Problem 14

Let R denotes the examples that are being classfied as *Positive*, and T denotes the true posive cases. We have:

$$\begin{aligned} P(R) &= P(R \mid T)P(T) + P(R \mid T^c)P(T^c) \\ &= P(R \mid T)(1 - P(T^c) + P(R \mid T^c)P(T^c) \\ &= P(R \mid T) - P(R \mid T)P(T^c) + P(R \mid T^c)P(T^c) \\ &= P(R \mid T) + [P(R \mid T^c) - P(R \mid T)] \cdot P(T^c) \\ P(R) - P(R \mid T) &= [P(R \mid T^c) - P(R \mid T)] \cdot P(T^c) \end{aligned}$$

We know that there must be $P(R) = P(R \mid T)$ as random guessing is an independent variable. We also know we should have $P(T^c) > 0$ for being a meaningful task. Subsituting these into the above equation, we have $0 = P(R \mid T^c) - P(R \mid T) \implies P(R \mid T^c) = P(R \mid T)$. This implies that the TP Rate is the same as the FP Rate and therefore $x = y$, and the ROC graph for a random gussing classifier will therefore be a diagonal line.