Section 4M. Receiver Operating Characteristics

Victor M. Preciado, PhD MIT EECS
Dept of Electrical & Systems Engineering
University of Pennsylvania
preciado@seas.upenn.edu

Threshold Classifier

How do we choose a classifier? We should take into account the cost of a FP versus the cost of a FN (e.g., medical errors). In the medical case, the cost of FNs is much higher than the cost of FPs; hence, we would like to reduce the FNR at expense of increasing the FPR...

- Remember: A (binary) Bayesian classifier satisfies

$$C(\mathbf{x}) = \begin{cases} 1 & \text{when } \Pr(Y = 1 | X = \mathbf{x}) \geq 1/2 \\ 0 & \text{when } \Pr(Y = 1 | X = \mathbf{x}) < 1/2 \end{cases}$$

  This classifier minimizes the *classification error rate*, which is cost optimal only when the cost of a FP is the same as the cost of a FN

- To consider problems in which these costs are different, we define a classifier parameterized by a threshold $\tau \in [0, 1]$

$$C_\tau(\mathbf{x}) = \begin{cases} 1 & \text{when } \Pr(Y = 1 | X = \mathbf{x}) \geq \tau \\ 0 & \text{when } \Pr(Y = 1 | X = \mathbf{x}) < \tau \end{cases}$$

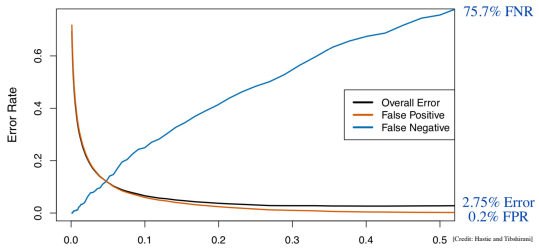
Threshold Classifier: Example

How do FNR and FPR change as a function of $\tau \in (0, 1/2)$ for the LDA classifier applied to the Default dataset?

- In the LDA binary classifier , we have that $C_\tau(\mathbf{x}) = 1$ when (Bayes' theorem)

$$\Pr(Y = 1|X = \mathbf{x}) = \frac{f_1(\mathbf{x})\pi_1}{f_0(\mathbf{x})\pi_0 + f_1(\mathbf{x})\pi_1} \geq \tau$$

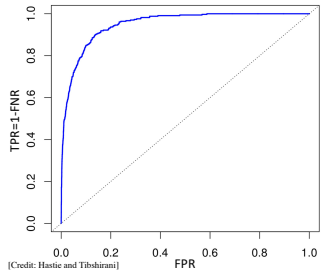
- The evolution of different error rates as we vary τ are shown below



Receiver Operating Characteristics

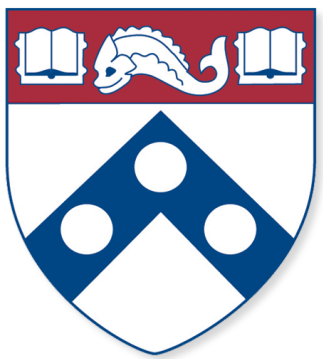
As we have seen before, we can choose τ such that out classifier obtains a lower number of FNs by sacrificing FPs. It is convenient to plot the relationship between FNR and FPR in a plot...

- **Receiver Operating Characteristics** (ROC): For historical reasons, this relationship is typically plot in terms of FPR and TPR=1-FNR.



[Credit: Hastie and Tibshirani]

The Area Under the ROC Curve (AUC) summarizes the overall performance of the parameterized classifier. A better classifier results in a higher AUC.



Penn
Engineering

Copyright 2020 University of Pennsylvania
No reproduction or distribution without permission.