

Biometría

Ejercicio Tema 2 – Medidas de calidad

En primer lugar, para este ejercicio se partirán de cuatro listas de datos que contienen, respectivamente, los scores de clientes e impostores de dos muestras: A y B, a partir de los cuales se obtienen los umbrales para cada una de las muestras con cada uno de sus valores, el 1 y el 0, que se encontrarán en listas ordenadas.

Para las curvas ROC, se representan para cada uno de los umbrales, el valor de falsos positivos frente al complementario al de los falsos negativos relativos a la cantidad de datos en la muestra, donde se habrán creado a su vez listas ordenadas con estos valores (ya que la de umbrales también estaba ordenada).

Para la obtención de falsos negativos a partir de falsos positivos y viceversa, ya que las listas están ordenadas, bastará con encontrar el índice en la lista del valor más cercano al argumento en cuestión, y esa misma posición en la lista de umbrales y de falsos negativos o positivos será el resultado esperado. En el caso de los valores para cuando falsos positivos y falsos negativos sean iguales (o lo más parecido posible) se recorrerán las listas y se devolverán los valores con menor diferencia en valor absoluto.

Para obtener el área bajo la curva ROC, se ha utilizado la fórmula propuesta en clase, sumando 1 por cada cliente con mayor score que cada impostor, y 0.5 en caso de que el score sea el mismo, dividido por el número de comparaciones realizadas.

Finalmente, el d-prime se ha obtenido con la fórmula de las diapositivas: la diferencia de las medias de scores de clientes e impostores, dividido entre la raíz de la suma de las varianzas.