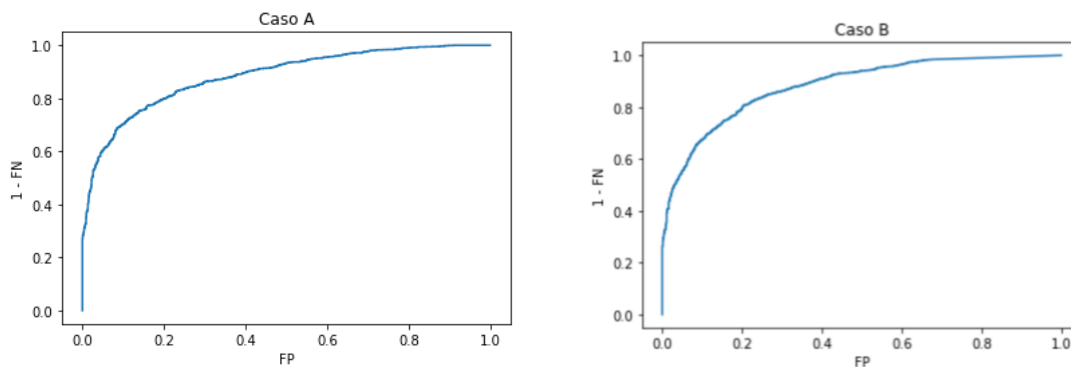


Biometría

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.

Ejercicios Tema 3 – PCA y LDA

En esta parte se llevarán a cabo los ejercicios correspondientes a las PCA (eigenfaces) y LDA (fisherfaces).

En el primer caso, el de PCA, consiste en reducir la dimensionalidad de una imagen desde d dimensiones (siendo $d = n \times m$) a d' dimensiones mediante proyección a componentes principales. Tras esto, el reconocimiento se realiza en el nuevo espacio de representación reducido, en este caso, a partir del vecino más cercano usando el toolkit scikit-learn.

Para ello se utilizan los d' primeros eigenvectores, que son vectores de d componentes con módulo unitario y que se pueden representar como caras eigenfaces, por ejemplo:

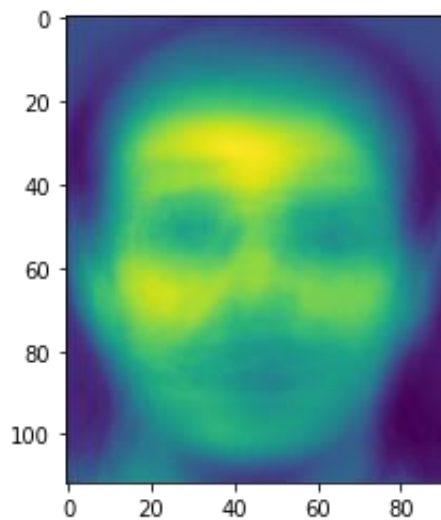


Ilustración 1 - Ejemplo eigenface

A continuación, se muestra la gráfica con los valores de accuracy obtenidos al reducir la dimensionalidad con la técnica de PCA desde 1 hasta 200 dimensiones (tamaño original), donde se puede ver que los Mejores resultados de clasificación por vecino más cercano se obtienen cuando se realiza una reducción a 78 dimensiones, con una precisión del 90%.

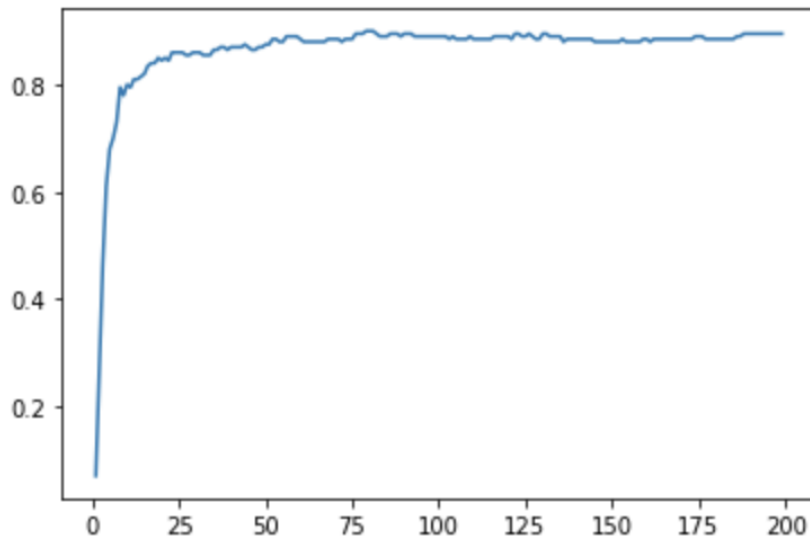


Ilustración 2 - Resultados de PCA

Por otro lado, para aplicar la técnica de LDA se suele aplicar antes PCA para evitar singularidades, por lo que la reducción de dimensionalidad del espacio original suele ser mayor en LDA con respecto a PCA, ya que primero se reduce a d' dimensiones con PCA.

Para este ejercicio se ha aplicado la técnica de LDA para reducir hasta 50 dimensiones y al igual que en el caso anterior se han obtenido valores de accuracy con vecino más cercano.

A continuación se pueden ver los resultados obtenidos gráficamente, donde se puede ver que el mejor resultado es de 82.5% reduciendo a 46 dimensiones y, como era de esperar, obtiene peores resultados que PCA.

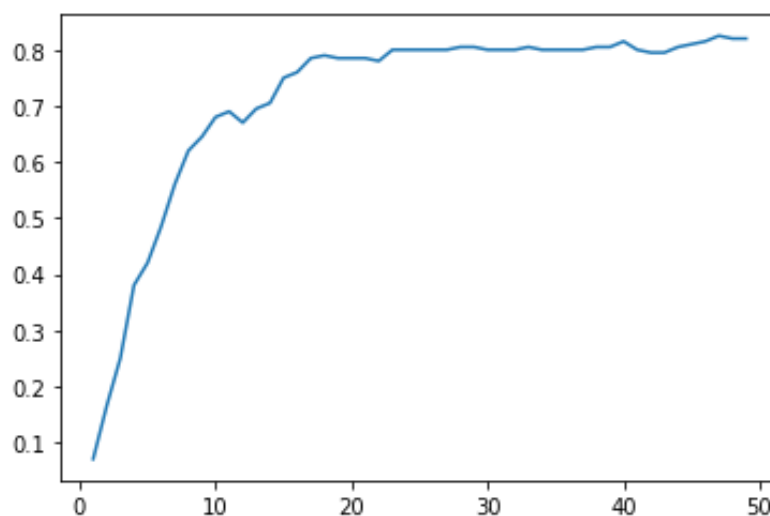


Ilustración 3 - Resultados de LDA