

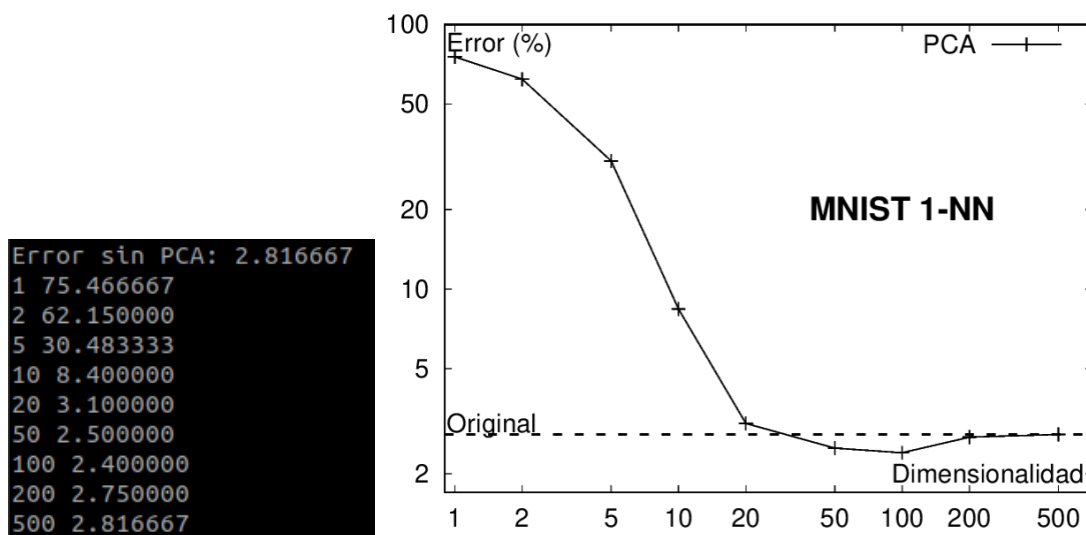
PERCEPCIÓN

Memoria de los Ejercicios Obligatorios

- Fabián Scherle Carboneres
- Lishuang Sun (María)

1. Script knn+pca-exp.m

En este apartado se tiene como objetivo obtener una gráfica en la que se puedan representar los resultados obtenidos al aplicar PCA con el 90% del conjunto de entrenamiento de dígitos manuscritos MNIST y a su vez clasificar con knn el 10% de dicho conjunto.



A la vista de los resultados obtenidos consideramos que la dimensionalidad óptima es de 100, puesto que da la mínima tasa de error, incluso en comparación con la obtenida sin PCA.

2. Script knn+pca-eva.m

Tomando como base el script knn+pca-eva.m proporcionado y la dimensionalidad óptima obtenida en el anterior apartado, se pide calcular la tasa de error de nuestro clasificador knn que clasifica el conjunto de prueba sobre el conjunto de entrenamiento, y a su vez comparar los resultados aplicando PCA y sin aplicar PCA.

Obtenemos los siguientes resultados al ejecutar nuestro script modificado:

```
Error con PCA: 2.840000
Error sin PCA: 3.090000
```

Se puede apreciar que aplicar PCA mejora en cuanto a tasa de error con respecto a la dimensionalidad original. Asimismo, reduciendo la dimensionalidad de ambos conjuntos

de muestras, a su vez reducimos el coste computacional, ya que se operan con menos datos, reduciendo también el requerimiento de espacio en memoria.

Comparando el resultado obtenido sin PCA con la MNIST web, comprobamos que coincide con la referencia de "Kenneth Wilder, U. Chicago".

En el caso del resultado obtenido al aplicar PCA notamos que es similar al de la referencia de "Kenneth Wilder, U. Chicago" con knn y distancia L3 sin preprocesado ni PCA. Cabe destacar que nuestro resultado con PCA también se acerca a la tasa de error de la referencia de "LeCun et al. 1998" con knn y distancia L2 con preprocesado de deskewing.