**PERCEPCIÓN**

**Memoria de los Ejercicios Opcionales**

* Fabián Scherle Carboneres
* Lishuang Sun (María)

1. Distancia L0, L1 y L3

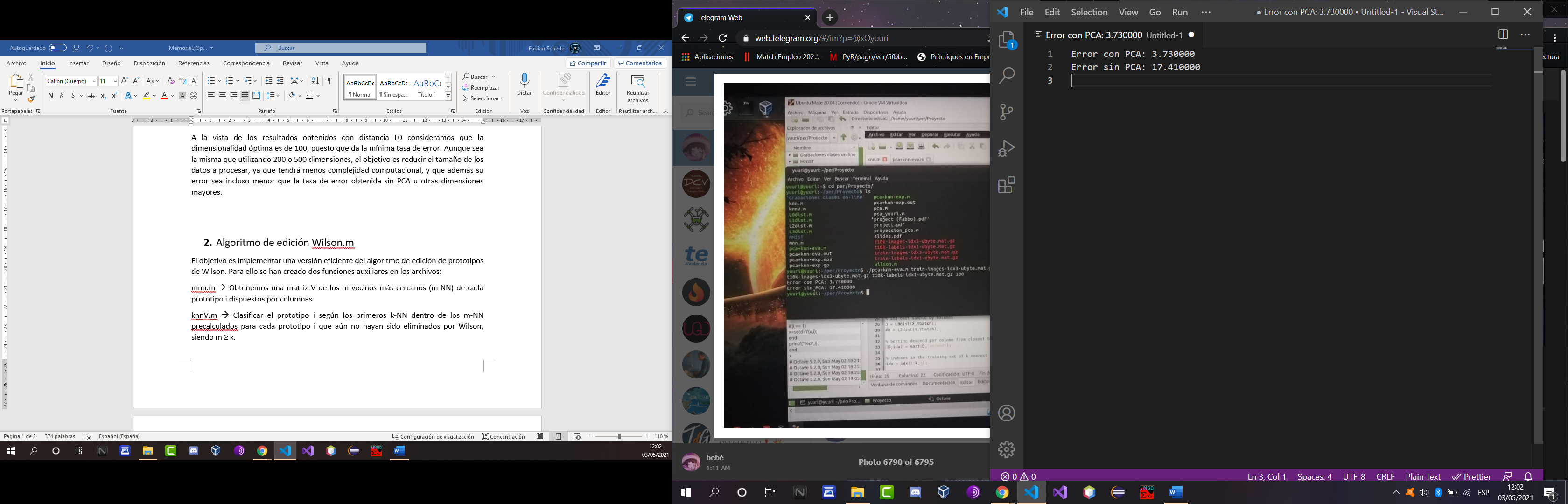
En este apartado se nos pide implementar otras familias Lp y estudiar su comportamiento en un clasificador knn con k=1 combinado con PCA. Hemos considerado estudiar el comportamiento del clasificador usando la distancia L0. Adicionalmente, hemos añadido el código necesario para probar las distancias L1 y L3.

Captura de pantalla de computadora

Descripción generada automáticamente Captura de pantalla de computadora

Descripción generada automáticamente

A la vista de los resultados obtenidos con distancia L0 y el script pca-knn-exp.m, consideramos que la dimensionalidad óptima es de 100, puesto que da la mínima tasa de error. Aunque sea la misma que utilizando 200 o 500 dimensiones, el objetivo es reducir el tamaño de los datos a procesar, ya que tendrá menos complejidad computacional, menos espacio en memoria, y que además su error sea incluso menor que la tasa de error obtenida sin PCA u otras dimensiones mayores.



En relación con el resultado de ejecutar pca+knn-eva.m con una dimensionalidad de 100, vemos que tanto la tasa de error sin aplicar PCA y aplicando PCA es similar a las tasas de error obtenidas en el script anterior. Si comparamos L0 con la distancia L2, podemos concluir que L2 ofrece menores tasas de error y por ende mejores resultados.

1. Algoritmo de edición Wilson.m

El objetivo es implementar una versión eficiente del algoritmo de edición de prototipos de Wilson. Para ello se han creado dos funciones auxiliares en los archivos:

mnn.m 🡪 Obtenemos una matriz V de los m vecinos más cercanos (m-NN) de cada prototipo i dispuestos por columnas.

knnV.m 🡪 Clasificar el prototipo i según los primeros k-NN dentro de los m-NN precalculados para cada prototipo i que aún no hayan sido eliminados por Wilson, siendo m ≥ k.

Si la clase a la que pertenece el prototipo i (dada por la función knnV) no es igual a la clase a la que realmente pertenece, se elimina este prototipo de la lista de prototipos resultantes.

Texto

Descripción generada automáticamenteGráfico, Gráfico de líneas

Descripción generada automáticamente

En vista de los resultados obtenidos con knn+pca-exp-wilson.m, podemos apreciar que el error de clasificación aplicando una reducción de dimensionalidad PCA con 50 dimensiones es incluso menor que clasificando con la dimensionalidad original. Por lo tanto, este clasificador parece funcionar bien incluso sin el formato original de los prototipos a clasificar.

Captura de pantalla de computadora

Descripción generada automáticamente

Ahora vamos a llevar el clasificador al mundo real con knn+pca-eva-wilson.m, esto es, clasificar prototipos de prueba con respecto a los prototipos de entrenamiento, dos conjuntos diferentes. En este caso, para PCA hemos usado 50 dimensiones, dado que es la dimensión óptima obtenida en el apartado anterior. A partir de los resultados (figura de arriba) podemos ver que, efectivamente, PCA reduce la dimensionalidad de los conjuntos pero que además conserva o mejora su capacidad discriminativa eliminando información innecesaria y eliminando prototipos ruidosos con Wilson.