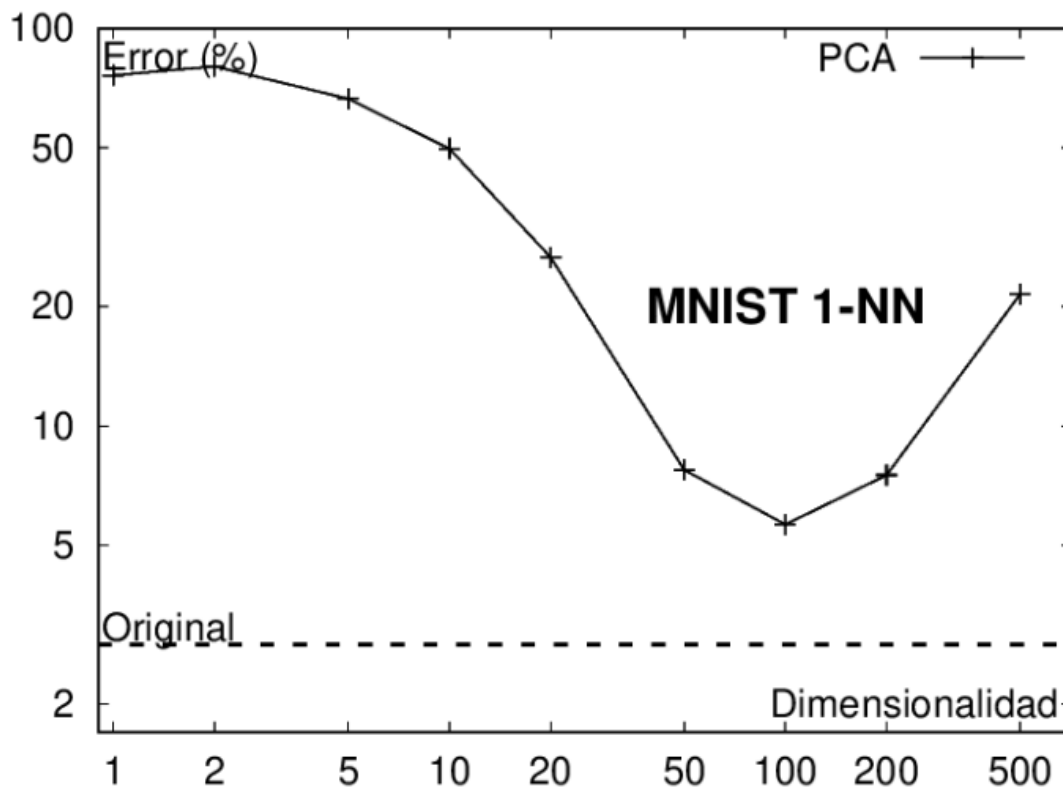


Análisis de la implementación de la distancia Mahalanobis-diagonal por clase

Presentación y justificación

A continuación, se presentarán las conclusiones obtenidas de la implementación y ejecución de la distancia Mahalanobis-diagonal por clase.

Análisis de resultados



Esta gráfica corresponde a los datos obtenidos tras la ejecución del archivo implementado Maha_pca+knn-exp.m.

La representación muestra la relación entre la tasa de error del clasificador empleando la distancia Mahalanobis-diagonal por clase y la dimensionalidad empleada.

De esta prueba se puede concluir que la dimensión de los datos afecta en gran medida a la eficacia del clasificador sea cual sea la distancia de cálculo empleada. Uno de los hechos observables más importantes es que con la distancia Mahalanobis-diagonal por clase la tasa de error al reducir de manera exagerada la dimensionalidad de los datos aumenta en mayor medida que con la distancia L2.

En nuestro caso concreto, en la gráfica se puede observar que la menor tasa de error se consigue con una dimensionalidad igual a 100, aunque en ningún momento se consigue llegar a la tasa de error del clasificador knn con cálculo de distancia L2 y sin empleo de la técnica PCA.

Análisis de la influencia de la PCA con la distancia Mahalanobis-diagonal por clase

Clasificador	Tasa de error
knn(Mahalanobis-diagonal por clase)	97.57
knn(Mahalanobis-diagonal por clase) +PCA(100)	5.9

Esta tabla muestra los resultados obtenidos de la ejecución tras el aprendizaje del clasificador.

Se puede observar que la aplicación del PCA sobre los datos permite una mejora excepcional a la hora de clasificarlos.

Esta mejora refleja que el PCA tiene una gran influencia sobre el clasificador knn mediante la utilización de la distancia Mahalanobis-diagonal por clase.

Comparación con la base de datos MNIST

Con la ejecución de nuestro clasificador hemos obtenido resultados similares a los obtenidos por LeCun et al en 1998 mediante un clasificador knn y el empleo de la distancia euclídea (L2).

Conclusiones

Para el caso concreto de los dígitos manuscritos y la utilización del clasificador de los k vecinos más cercanos se observa que el uso de la distancia Mahalanobis-diagonal por clase no presenta una mejora en comparación a la distancia L2.

Además con la comparación de los datos obtenidos y los presentes en la base MNIST se demuestra que la implementación llevada a cabo para este trabajo no es óptima ya que se obtienen resultados no similares.