

# Mixtura de Gaussianas

Miguel Ángel Navarro Arenas

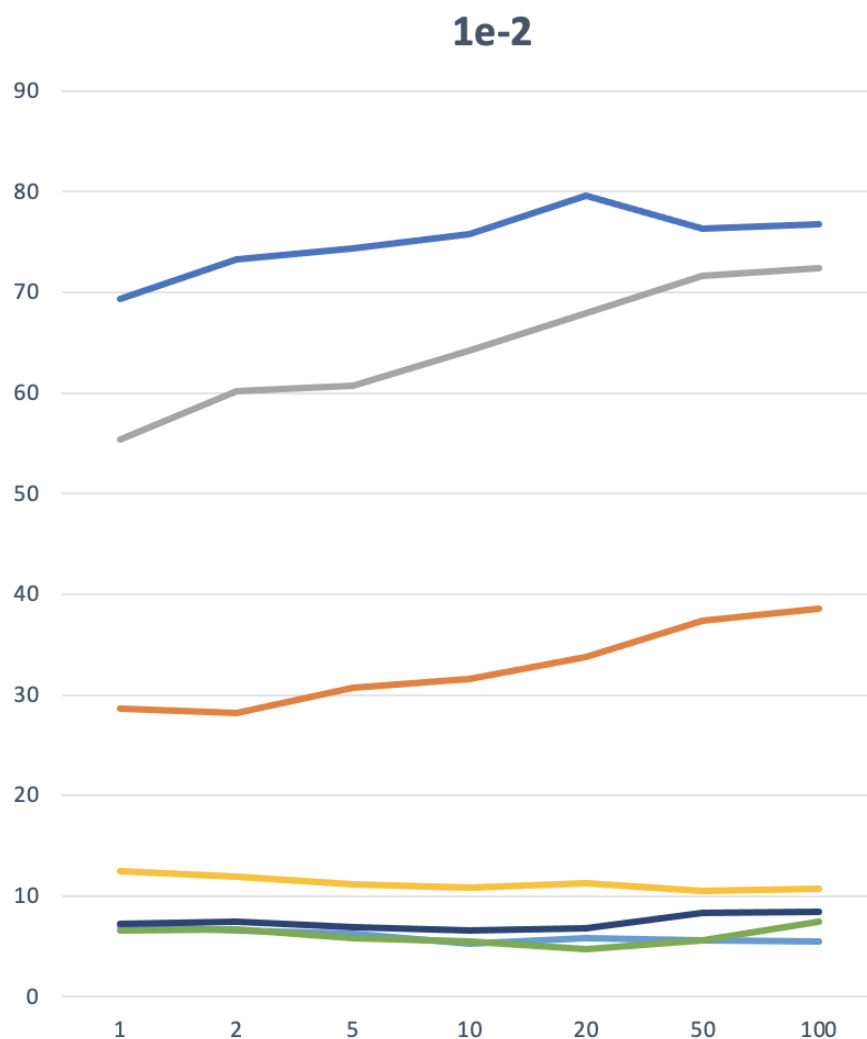
---

## EJERCICIO 1.

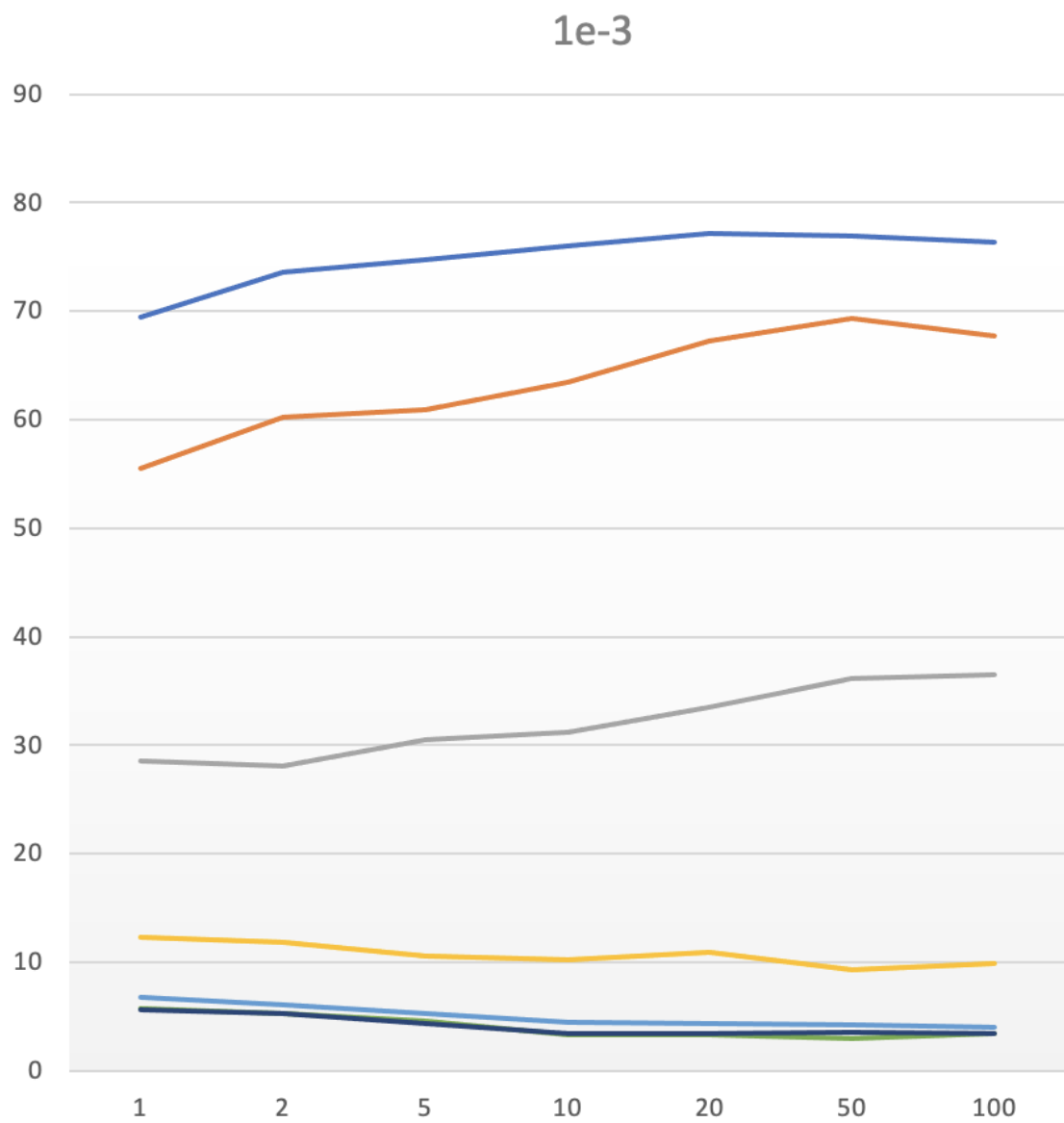
Este ejercicio, en definitiva, consiste en realizar iteraciones del PCA con diversos valores de los componentes por mixtura y el número de dimensiones de PCA.

Hemos hecho un archivo `experiment.m` en el cual realizamos lo anteriormente dicho. Obtenemos los siguientes resultados para un  $\alpha = [1e-2, 1e-3, 1e-4, 1e-5]$ :

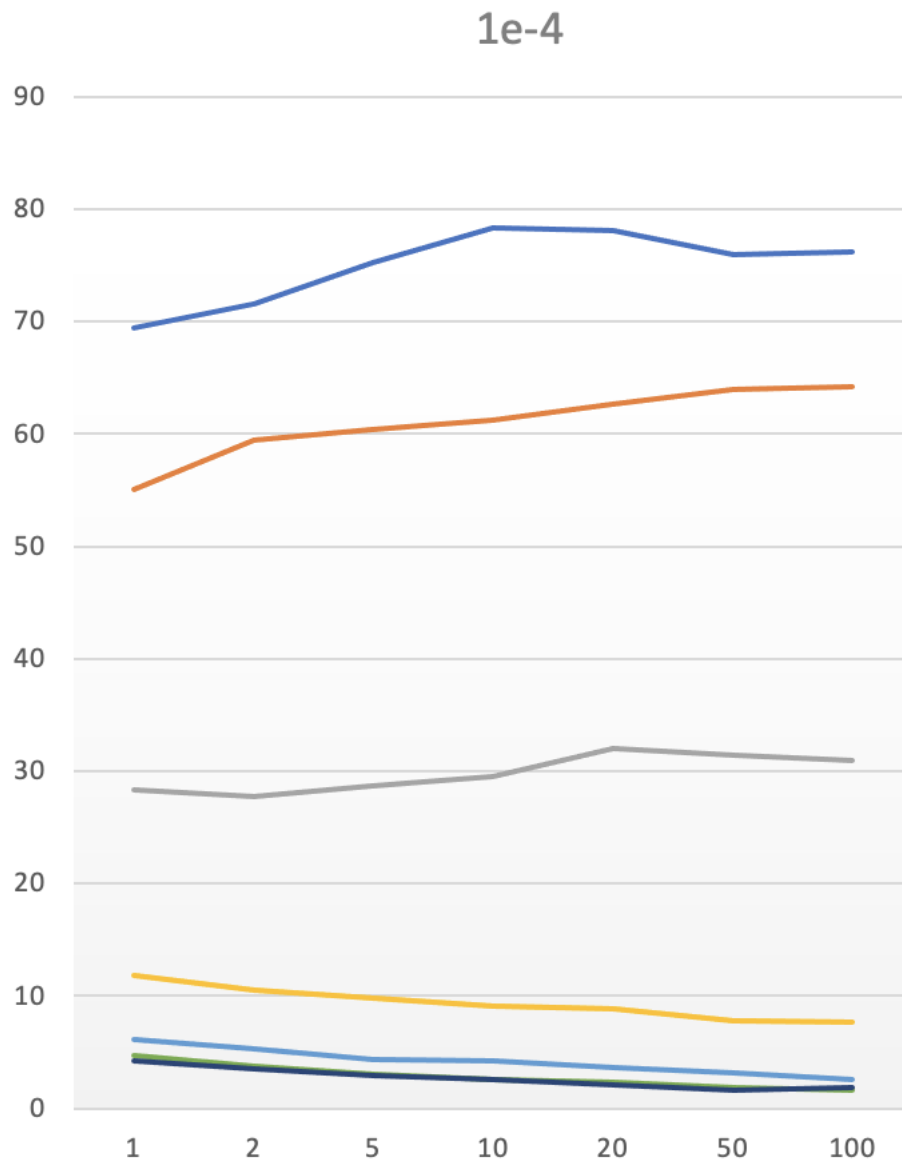
- **1e-2:**



○ 1e-3:

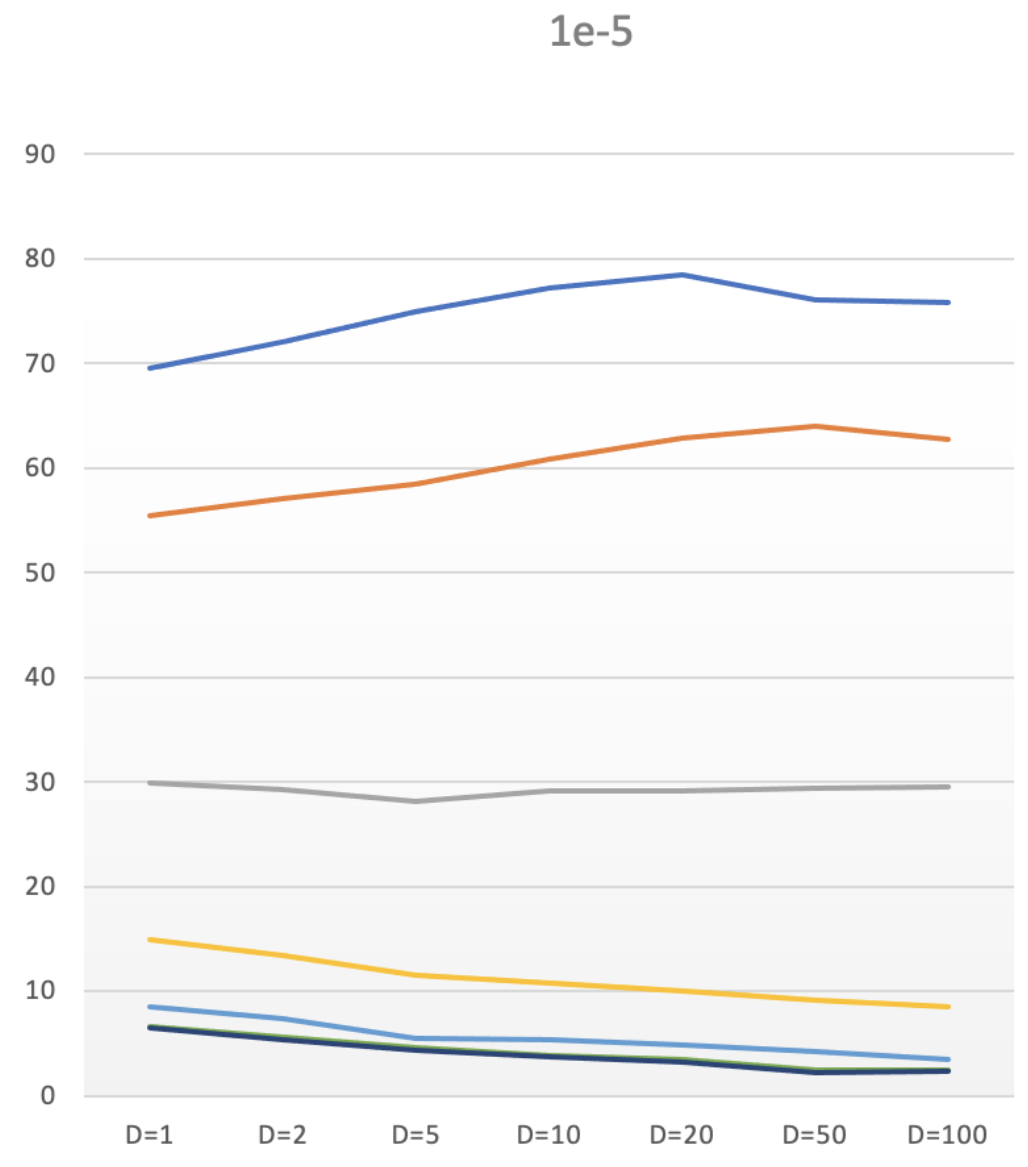


- 1e-4:



\*\*Aunque no se pueda observar correctamente con estos gráficos, en los datos obtenemos que los mejores resultados se obtienen con este alpha. En las tablas con todos los datos, el mínimo error que hemos obtenido ha sido con un alpha de 1e-4, con una K=50 y una D=100.

- 1e-5:



En los gráficos, la vertical corresponde al porcentaje de error de cada clasificador.

Las líneas, de arriba abajo, corresponden, por colores a K's ascendentes, es decir, las azules superiores corresponden a  $K = 1$ , las naranjas a  $K = 2$ , etc. Como observamos, los mejores resultados se corresponden con las K's más altas, es decir,  $K=50$  y  $K=100$ . Para poder saber cuales son las alphas mejores, debemos irnos a los datos (Los cuales se encuentran en el archivo Excel en el que se recogen todos los datos que han salido por pantalla en todas las ejecuciones).

Podemos observar que los mejores resultados se acumulan, en todas las alphas, en los valores más altos de K y de D. Podemos ver que en algunos casos se puede producir sobreentrenamiento y obtener peores resultados con K y D alto.

Por lo general, los mejores resultados se han obtenido con un K = 20, 50, 100 y un D=50, 100.

Estos últimos datos son los que utilizaremos para el siguiente ejercicio.

## EJERCICIO 2.

Para la realización de este ejercicio hemos utilizado el archivo eva.m, con los datos que hemos obtenido en el ejercicio anterior que hemos considerado como mejores.

Resultados obtenidos:

D	K	Intervalo	Error
50	20	1,6-2,2	1,91
50	50	1,6-2,1	1,85
50	100	1,38-1,88	1,67
100	20	1,8-2,3	1,98
100	50	1,65-2,2	1,90
100	100	1,48-1,97	1,72

El mínimo error que nosotros hemos conseguido es de entorno al 4,1. Hemos observado que algunos clasificadores de la página de MNIST están en torno al 3%. Podríamos mejorar estos resultados aplicando otro tipo de técnicas como veremos en siguientes ejercicios con SVM o redes neuronales.