

Aprendizaje automático

- práctica 4 -

Por:

Diego Caballé Casanova (738712)

Apartado 2

En este caso clasificaremos números escritos a mano, para esto necesitaremos hacer la versión de la regresión regularizada multiclase.

Lo primero ha sido encontrar el mejor lambda usando k-fold en mi caso, el mejor lambda es: 0.002222996482526.

Con este lambda, podemos reentrenar con todos los datos y conseguir la Theta grande, es decir, la matriz de las tetas de cada clase.

Aprovechando esto y como corrección he utilizado este entrenamiento para ver la tasa de error con los datos de entrenamiento y de test. Las tasas de error son respectivamente:

Tetr =

0.0728

Tecv =

0.0988

Unos valores muy muy aceptables.

Apartado 3

Después de haber entrenado con todos los datos, usamos sólo los datos de test para calcular la matriz de confusión y los valores de precision y recall. Los resultados son:

matrizDeConfusion =

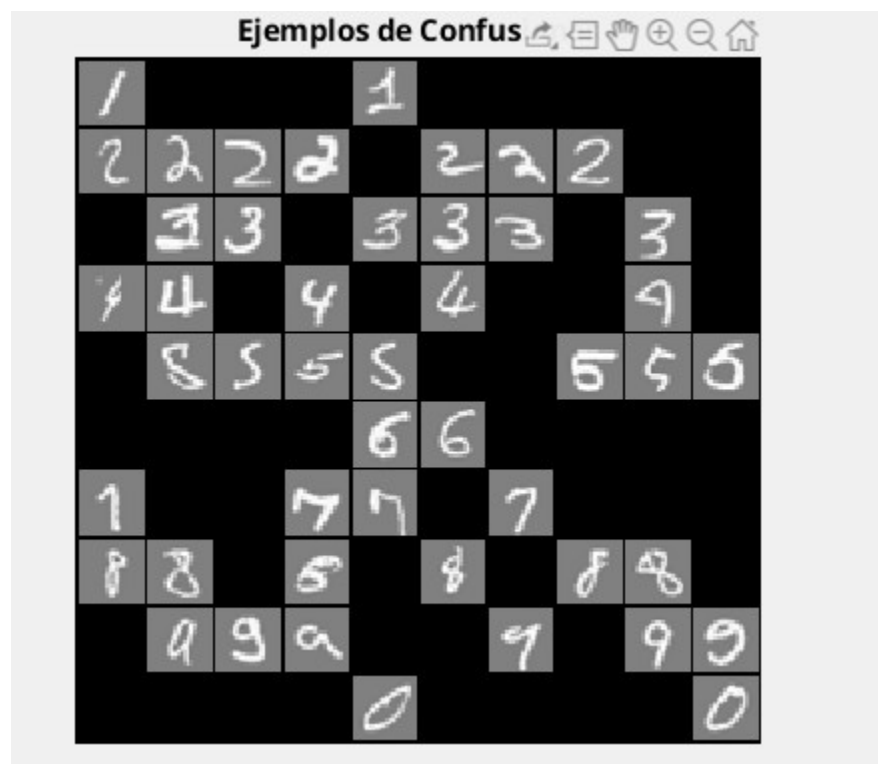
| | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|-----|
| 81 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 84 |
| 3 | 83 | 1 | 1 | 0 | 1 | 2 | 3 | 0 | 0 | 94 |
| 0 | 2 | 64 | 0 | 7 | 1 | 4 | 0 | 1 | 0 | 79 |
| 1 | 1 | 0 | 74 | 0 | 3 | 0 | 0 | 5 | 0 | 84 |
| 0 | 1 | 3 | 2 | 69 | 0 | 0 | 1 | 3 | 1 | 80 |
| 0 | 0 | 0 | 0 | 3 | 73 | 0 | 0 | 0 | 0 | 76 |
| 2 | 0 | 0 | 3 | 1 | 0 | 66 | 0 | 0 | 0 | 72 |
| 1 | 1 | 0 | 2 | 0 | 3 | 0 | 77 | 4 | 0 | 88 |
| 0 | 1 | 1 | 3 | 0 | 0 | 2 | 0 | 62 | 1 | 70 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 72 | 73 |
| 88 | 89 | 69 | 85 | 84 | 81 | 74 | 81 | 75 | 74 | 800 |

recall =

0.9384

precision =

0.9228



Los datos que peor reconoce serían los 5 y los 8. Pero los datos, en general son bastante buenos.