



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Máster en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital
Universitat Politècnica de València

Reconocimiento y clasificación de imágenes de ropa

Redes Neuronales Artificiales

Autor: Juan Antonio López Ramírez

Curso 2019-2020

Índice general

Índice general	III
1 Introducción	1
2 Configuración del modelo basado en MLP	3

CAPÍTULO 1

Introducción

El objetivo del trabajo consiste en probar los modelos obtenidos en la práctica de la asignatura, pero empleando esta vez otro dataset para el entrenamiento y prueba.

En nuestro caso, hemos escogido la base de datos de *Moda*, que contiene 70.000 imágenes en escala de grises con 10 categorías. Dichas imágenes muestran prendas de ropa a baja resolución (2828 píxeles).

Esto supone una evolución de la base de datos MNIST que habíamos utilizado en las sesiones de prácticas, pero el formato es el mismo: 70.000 imágenes en escala de grises de 2828 píxeles en 10 categorías coincidentes con los dígitos (0-9). Con *Moda* simplemente varía que las imágenes ahora son prendas de vestir en lugar de dígitos, y las posibles clasificaciones son:

- (0) - Camiseta
- (1) - Pantalón
- (2) - Suéter
- (3) - Vestido
- (4) - Abrigo
- (5) - Sandalia
- (6) - Camisa
- (7) - Zapatilla deportiva
- (8) - Bolso
- (9) - Botín

Al igual que en las prácticas, usaremos 60.000 imágenes para entrenar la red neuronal y 10.000 para evaluar la precisión con la que esta ha aprendido a clasificar. Entrenaremos un modelo basado en *Multi-layer Perceptron* (MLP) a partir de los resultados que se obtuvieron en las prácticas para MNIST.

Para implementar las redes neuronales, se han empleado los Jupyter Notebooks que proporciona Google Colaboratory. El entorno de ejecución por defecto que se ha usado ha sido el acelerador por GPU debido a un mejor rendimiento de tiempo.

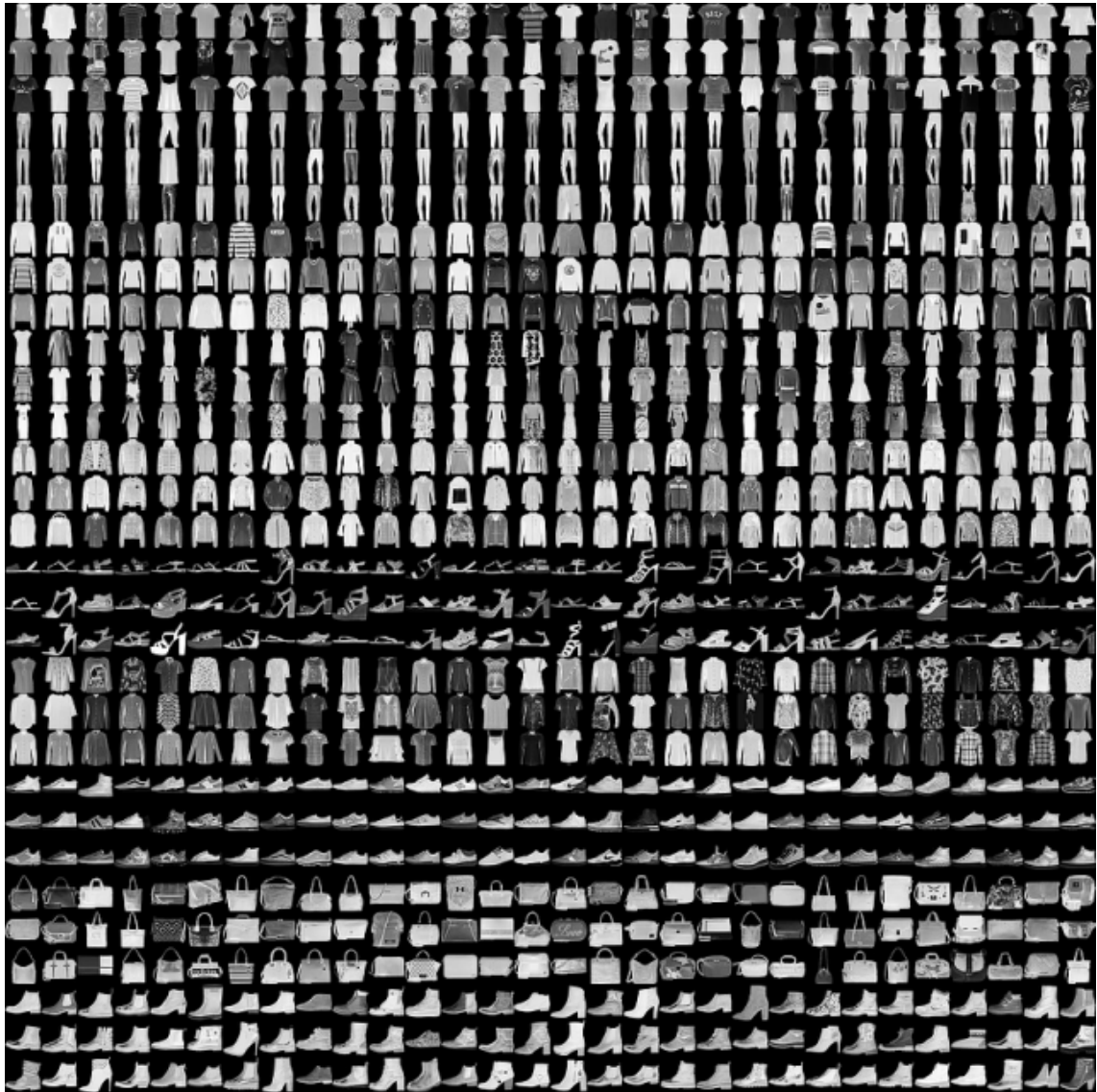


Figura 1.1: Ejemplo de muestras de la base de datos *Moda* de MNIST.

CAPÍTULO 2

Configuración del modelo basado en MLP

Para la configuración del MLP, se ha partido de los resultados que se obtuvieron en la práctica para el problema de reconocimiento de dígitos con MNIST. Esta configuración final consistía en:

- Una topología consistente en 2 capas ocultas de 512 neuronas cada una, todas con función de activación *relu*, mientras que la capa de salida cuenta con una función de activación *softmax*.
- Un tamaño de *batch* de 100 y 150 *epochs*.
- La normalización del *batch*.
- Un ruido gaussiano de 0.1 en todas las capas.
- *Learning rate annealing* para que el factor de aprendizaje empiece teniendo un valor de 0.1, se modifique en la epoch 75 a 0.01 y en la 125 a 0.001.
- *Data Augmentation* que mueve las muestras en altura y en anchura en un 10 %.

El resultado obtenido ha sido de un **90.49 %** de acierto en test, un buen resultado, aunque bastante inferior al 99.34 % de acierto que obteníamos en el problema de reconocimiento de dígitos.

Se hicieron unas pequeñas modificaciones para intentar mejorarlo, como cambiar el *Data Augmentation* para que en lugar de mover las muestras en altura y anchura un 10 % lo hiciese un 20 %, pero se obtuvo un resultado de 88.33 %. Aumentando el número de neuronas por capa de 512 a 1024, obtuvimos un 90.42 %.

Estos resultados pueden deberse a que es más sencillo reconocer un número que una prenda de vestir, ya que se trata de un objeto más complejo con más características a generalizar por parte de nuestra red neuronal.