

Memoria

REDES NEURONALES

Mireia Pires State y Darío Mesas Martí
Aprendizaje automático | 20/12/2022

Contenido

DESCRIPCIÓN DEL TRABAJO RELIZADO	2
DESCRIPCIÓN DEL EXPERIMENTO	3

DESCRIPCIÓN DEL TRABAJO RELIZADO

Para el desarrollo de esta práctica se han realizado diferentes experimentos para entrenar un clasificador basado en redes neuronales que minimice el error de clasificación de FashionMNIST en test.

El código empleado ha sido proporcionado por el profesorado y el entorno que se ha usado para ejecutarlo y modificarlo ha sido Google colab. En cuanto a las modificaciones, se han tenido en cuenta los siguientes parámetros:

- **N.º de épocas:** Este parámetro indica el número de veces que se entrena la red neuronal con un conjunto de datos. En general, se necesitan varias épocas para entrenar una red neuronal, pero si se usan demasiadas podría dar lugar a un modelo sobre ajustado a los datos de entrenamiento y su precisión en conjuntos de datos desconocidos puede disminuir.

En esta práctica se ha realizado el experimento con hasta 20 épocas.

- **Factor de aprendizaje:** Este valor se utiliza para determinar qué tan rápido o lento se actualizan los pesos de la red neuronal durante el proceso de entrenamiento. Un valor demasiado elevado puede dar lugar a un modelo sobre ajustado y un valor muy pequeño puede hacer que el entrenamiento sea demasiado largo y no converger a una solución óptima.

Los valores usados para este parámetro han sido {0.0001, 0.001, 0.01, 0.1}.

- **N.º de neuronas:** Es el número de unidades de procesamiento que se encuentran en cada capa de la red. Un gran número de neuronas puede hacer que el modelo aprenda y generalice mejor, pero es más costoso.

Se han usado las siguientes cantidades de neuronas {128, 256, 1024}.

- **N.º de capas:** Se refiere al número total de capas ocultas que hay en la red, ya que, la capa de entrada es de tamaño 28x28 (que es la resolución de las imágenes del conjunto de datos Fashion-MNIST) y la capa de salida es de tamaño 10 (que corresponde a las 10 clases de ropa del conjunto de datos).

En esta práctica se ha experimentado con diferentes números de capas, concretamente 1, 2 y 3.

- **Tamaño de batch:** Es la cantidad de datos de entrenamiento que se utilizan en cada iteración del proceso de entrenamiento.

Se han usado los siguientes tamaños {32, 64, 128}.

DESCRIPCIÓN DEL EXPERIMENTO

En primer lugar, se ha probado con diferentes épocas (15 y 20) y se ha elegido el valor 15 porque era la época donde se conseguía la mejor precisión (en otros experimentos se han usado 30 épocas para ver la evolución).

	teacc	teloss
Epoch = 10	88,5%	0.35
Epoch = 15	89,6%	0.35
Epoch = 20	89,2%	0.38

Tabla 1

Respecto al factor de aprendizaje, se han probado los valores mencionados anteriormente con 15 épocas.

	teacc	teloss
F. aprend. = $1e-4$	89,1%	0.32
F. aprend. = $1e-3$	89,6%	0.35
F. aprend. = $1e-2$	85,9%	0.42
F. aprend. = $1e-1$	10,0%	2.32

Tabla 2

En cuanto a la cantidad de neuronas, se ha experimentado con diferentes valores teniendo en cuenta la mejor época y factor de aprendizaje ya calculados.

	teacc	teloss	tiempo
Neuronas = 128	89,1%	0.34	2min 30s
Neuronas = 256	89,6%	0.33	3min 38s
Neuronas = 1024	89,3%	0.34	14min 51s

Tabla 3

Cabe destacar que se ha notado que cuanto mayor cantidad de neuronas, mayor era el tiempo de entrenamiento. En este caso, los tiempos han sido para 128 neuronas 2min 30s, para 256 neuronas 3min 38s y para 1024 neuronas 14min 51s. Además, el mejor resultado que se observa en la tabla se consigue con 14 épocas.

Luego, se han probado diferentes cantidades de capas ocultas todas ellas con 256 neuronas.

	teacc	teloss	tiempo
Capas = 1	88,9%	0.32	2min 42s
Capas = 2	89,6%	0.33	3min 38s
Capas = 3	88,9%	0.34	3min 54s

Tabla 4

Como se puede observar, el mejor valor para el entrenamiento de esta red neuronal es de dos capas ocultas con 256 neuronas. También, se ha visto que el tiempo de entrenamiento aumenta si aumentan las capas.

Por último, se ha entrenado al modelo con diferentes tamaños de batch para los mejores valores de parámetros que se han calculado.

	teacc	teloss	tiempo
Tamaño = 32	88,9%	0.35	4min 43s
Tamaño = 64	89,6%	0.33	3min 38s
Tamaño = 128	89,5%	0.32	2min 38s

Tabla 5

Finalmente, como se ha observado que el tamaño de batch de 128 daba mejores resultados, se han aumentado el numero de capas a tres y las épocas a 30. En la época 24 se ha conseguido mejorar el valor de la precisión a 89,7%, teloss = 0.39 y tiempo = 4min 47s.