

Ejercicios - MNIST

El modelo que se desarrolló en clase tiene una precisión ya bastante alta. Sin embargo, hay varios ajustes que se pueden intentar para mejorarlo.

Es importante poner atención al tiempo que se tarda cada época en ejecutar.

Utilizando el código visto en clase, experimenten con los hiperparámetros del algoritmo.

1. El ancho (tamaño de la capa escondida) del algoritmo. Intenten con un tamaño de 200. ¿Cómo cambia la precisión de validación del modelo? ¿Cuánto tiempo se tardó el algoritmo en entrenar? ¿Puede encontrar un tamaño de capa escondida que funcione mejor?
2. La profundidad del algoritmo. Agreguen una capa escondida más al algoritmo. Este es un ejercicio extremadamente importante! ¿Cómo cambia la precisión de validación? ¿Qué hay del tiempo que se tarda en ejecutar? Pista: deben tener cuidado con las formas (dimensiones) de los pesos y los sesgos.
3. El ancho y la profundidad del algoritmo. Agregue cuantas capas sean necesarias para llegar a 5 capas escondidas. Es más, ajusten el ancho del algoritmo conforme lo encuentre más conveniente. ¿Cómo cambia la precisión de validación? ¿Qué hay del tiempo de ejecución?
4. Experimenten con las funciones de activación. Intenten aplicar una transformación sigmoideal a ambas capas. La activación sigmoideal se obtiene escribiendo “sigmoid” (Ojo, hay que verificar si en las versiones más recientes de Tensorflow sigue siendo así)
5. Continúen experimentando con las funciones de activación. Intenten aplicar un ReLu a la primera capa escondida y tanh a la segunda. La activación tanh se obtiene escribiendo “tanh”. (Ojo, hay que verificar si en las versiones más recientes de Tensorflow sigue siendo así)

6. Ajusten el tamaño de la tanda. Prueben con un tamaño de tanda de 10,000. ¿Cómo cambia el tiempo requerido? ¿Cómo cambia la precisión?
7. Ajusten el tamaño de la tanda a 1. Eso corresponde al SGD. ¿Cómo cambian el tiempo y la precisión? ¿Es el resultado coherente con la teoría?
8. Ajusten la tasa de aprendizaje. Prueben con un valor de 0.0001. ¿Hace alguna diferencia?
9. Ajusten la tasa de aprendizaje a 0.02. ¿Hay alguna diferencia?
10. Combinen todos los métodos indicados arriba e intenten llegar a una precisión de validación de 98.5% o más.