

宿題 3

MNIST を識別するニューラルネットワークを実装する。宿題 1 と同等の構造を利用する。

中間層のユニット数を 128, 学習率を 0.001, ミニバッチサイズを 32, エポックを 50 とした。訓練データのうち 10,000 個を validation 用に分け, 学習状況の確認に利用した。

混同行列は表 1 のようになり, また, 各カテゴリごとの正解率等は表 2 のようになった。結果としては, 学習は進まず, 全て 3 と予測してしまうモデルとなってしまった。ハイパラを振ったりもしてみたが, 1 つの数字を予測するのみのモデルとなってしまった。原因としては, まず実装ミスを疑ったが, 宿題 1 と同じモジュールを使っているのであまり考えにくいと思われた。現状, ハイパラにかなり敏感なのかという考察となった。

プログラムは??ページの Listing ??に示した。関数についての説明は宿題 2 に示した通りである。また, モデルについては宿題 1 で示した通りである。

表 1: NN に対する混同行列

	0	1	2	3	4	5	6	7	8	9
0	0	0	0	980	0	0	0	0	0	0
1	0	0	0	1135	0	0	0	0	0	0
2	0	0	0	1032	0	0	0	0	0	0
3	0	0	0	1010	0	0	0	0	0	0
4	0	0	0	982	0	0	0	0	0	0
5	0	0	0	892	0	0	0	0	0	0
6	0	0	0	958	0	0	0	0	0	0
7	0	0	0	1028	0	0	0	0	0	0
8	0	0	0	974	0	0	0	0	0	0
9	0	0	0	1009	0	0	0	0	0	0

表 2: NN に対する各カテゴリごとの結果

Category	#Data	#Correct	Accuracy
0	980	0	0.000
1	1135	0	0.000
2	1032	0	0.000
3	1010	1010	1.000
4	982	0	0.000
5	892	0	0.000
6	958	0	0.000
7	1028	0	0.000
8	974	0	0.000
9	1009	0	0.000
All	10000	1010	0.101