

Neural Network Implementation: Digits Classification

資工所 碩一 田少谷 P76071268

Question 1

Development environment:

```
kernel version: #39-Ubuntu SMP
Processor type: x86_64
Memory size: 31GiB System memorylsh

CPUs: 12
CPU model name: Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz
L1d cache: 32K
L1i cache: 32K
L2 cache: 256K
L3 cache: 12288K

network: Ethernet interface
```

Parametres:

iteration times: 10000

batch size : 64

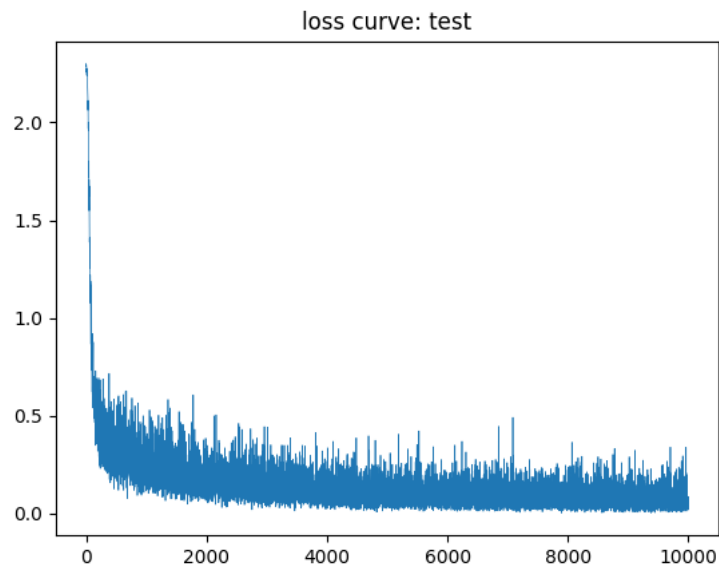
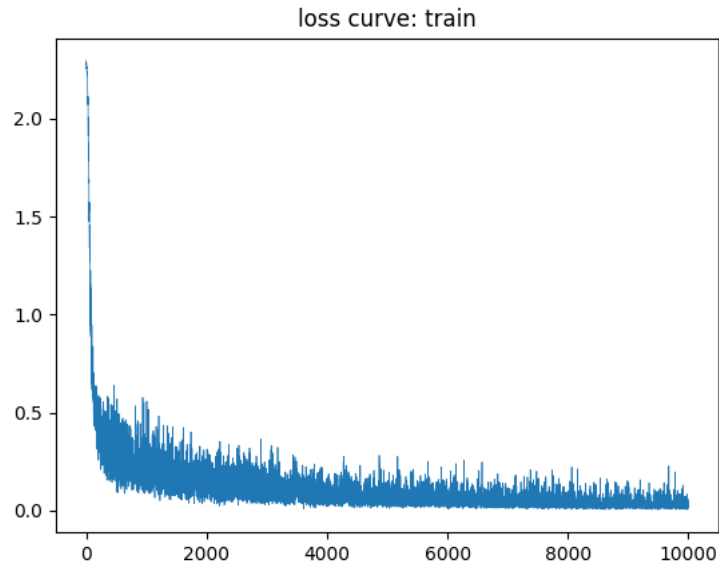
learning rate : 0.1

Result images:

result of (a) wide hidden layer:

the accuracy is at most 0.985

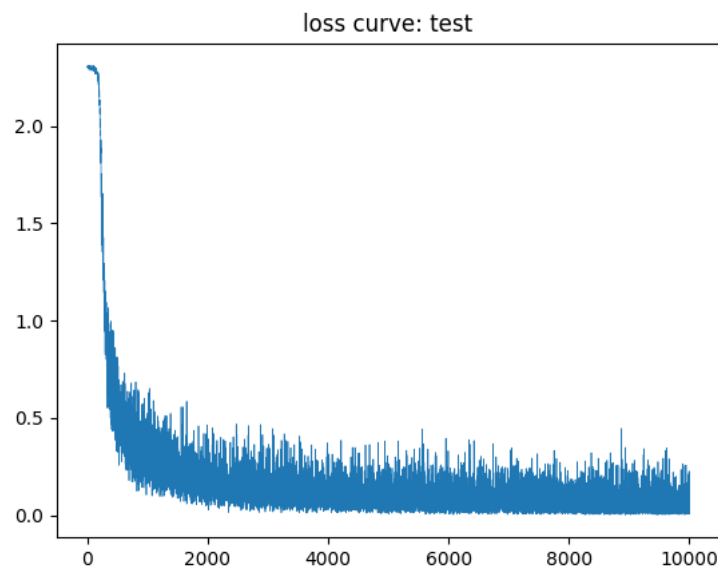
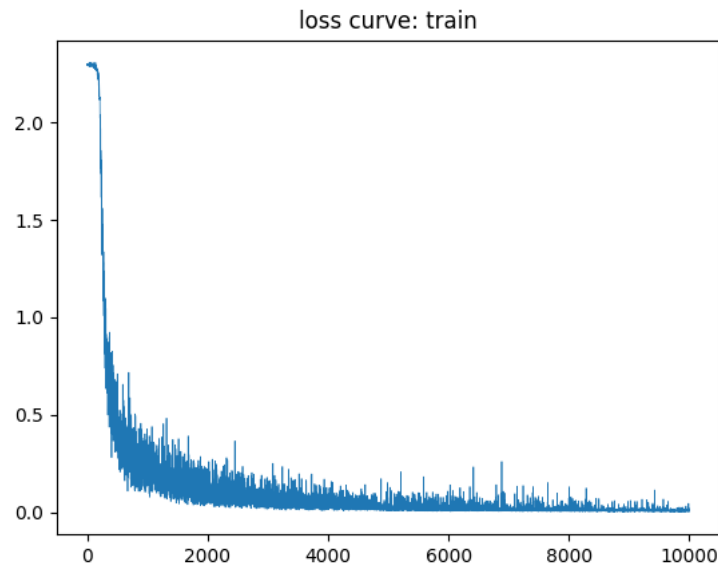
```
(venv) shao@gpuserval-System-Product-Name:~/hw4$ python3 num1a.py
train accuracy: 0.12825 ; test accuracy: 0.1336
train accuracy: 0.9437333333333333 ; test accuracy: 0.9449
train accuracy: 0.96385 ; test accuracy: 0.9611
train accuracy: 0.9741166666666666 ; test accuracy: 0.9685
train accuracy: 0.9791833333333333 ; test accuracy: 0.9712
train accuracy: 0.9850333333333333 ; test accuracy: 0.974
loss_train: [2.29639736 2.28818907 2.28981707 ... 0.03465378 0.07524023 0.019264 ]
loss_test [2.30022353 2.29811577 2.2918678 ... 0.20929411 0.02156008 0.08453439]
(venv) shao@gpuserval-System-Product-Name:~/hw4$ python3 num1a.py
train accuracy: 0.10823333333333333 ; test accuracy: 0.1075
train accuracy: 0.9447 ; test accuracy: 0.9431
train accuracy: 0.9657833333333333 ; test accuracy: 0.963
train accuracy: 0.9741833333333333 ; test accuracy: 0.9698
train accuracy: 0.9805333333333334 ; test accuracy: 0.973
train accuracy: 0.98425 ; test accuracy: 0.9743
loss_train: [2.29404952 2.29381747 2.29362879 ... 0.03690932 0.01328468 0.01671525]
loss_test [2.29938793 2.29860547 2.29211402 ... 0.0881411 0.11233221 0.05047719]
(venv) shao@gpuserval-System-Product-Name:~/hw4$ python3 num1a.py
train accuracy: 0.17778333333333332 ; test accuracy: 0.1827
train accuracy: 0.9443666666666667 ; test accuracy: 0.943
train accuracy: 0.9651166666666666 ; test accuracy: 0.9619
train accuracy: 0.97615 ; test accuracy: 0.9714
train accuracy: 0.9813666666666667 ; test accuracy: 0.9741
train accuracy: 0.9856333333333334 ; test accuracy: 0.9748
loss_train: [2.29656923 2.29703395 2.29158372 ... 0.01012397 0.04353799 0.01943344]
loss_test [2.29769568 2.29695445 2.29734655 ... 0.03480559 0.11144487 0.1679698 ]
```



result of (b) deep hidden layer:

the accuracy is at most 0.990

```
(venv) shao@gpuserval-System-Product-Name:~/hw4$ python3 num1b.py
train accuracy: 0.11236666666666667 test accuracy: 0.1135
train accuracy: 0.9457666666666666 test accuracy: 0.9446
train accuracy: 0.9732833333333333 test accuracy: 0.9671
train accuracy: 0.9786 test accuracy: 0.9712
train accuracy: 0.9874833333333334 test accuracy: 0.9766
train accuracy: 0.9869166666666667 test accuracy: 0.9743
loss_train: [2.30015595 2.3022661 2.30130096 ... 0.01296681 0.0116353 0.00513675]
loss_test [2.30280203 2.30214316 2.30229289 ... 0.0790417 0.01052876 0.02534678]
(venv) shao@gpuserval-System-Product-Name:~/hw4$ python3 num1b.py
train accuracy: 0.11668333333333333 test accuracy: 0.1147
train accuracy: 0.94705 test accuracy: 0.9467
train accuracy: 0.9615333333333334 test accuracy: 0.9591
train accuracy: 0.9799833333333333 test accuracy: 0.9698
train accuracy: 0.9876166666666667 test accuracy: 0.975
train accuracy: 0.9906666666666667 test accuracy: 0.9757
loss_train: [2.30080373 2.30082048 2.3008083 ... 0.00801993 0.00318126 0.00901452]
loss_test [2.30252582 2.3022015 2.30263693 ... 0.09111226 0.05055482 0.01100938]
(venv) shao@gpuserval-System-Product-Name:~/hw4$ python3 num1b.py
train accuracy: 0.09921666666666666 test accuracy: 0.1009
train accuracy: 0.94685 test accuracy: 0.9454
train accuracy: 0.9715833333333334 test accuracy: 0.965
train accuracy: 0.9821 test accuracy: 0.9733
train accuracy: 0.98745 test accuracy: 0.9739
train accuracy: 0.9895166666666667 test accuracy: 0.9764
loss_train: [2.29962451 2.3007029 2.30168651 ... 0.00625633 0.00432154 0.00548241]
loss_test [2.30129903 2.30215357 2.30441588 ... 0.04971366 0.05431845 0.06817228]
```

**Observations:**

1. The result of a three-layer-neural-network is slightly better than that of a two-layer-neural-network, probably because of the sufficient dataset
2. The result of train accuracy is usually better than that of test accuracy, as the model is trained with the dataset of train set. However, sometimes the test accuracy is higher than the train accuracy. This mostly occurs in the early stage of training process. As for the reason of this phenomena, I believe it's purely coincidence.
3. The values of loss decrease as the model trains better, which makes sense.
4. The parametres mentioned above are tried and considered to be leading to better results; however, as many learners of AI concern, I don't really understand why ends in this conclusion and this set of optimal parametres. I tried to print out some processes of this training, but didn't yet find out any useful information.
- 5.

Clarification:

The code of this report is constructed according to the book "Deep Learning：用Python進行深度學習的基礎理論實作". Therefore, the structure and some part of my code looks similar to the code provided in the book.

12/17/2018

Neural Network Implementation: Digits Classification - HackMD

tags: Machine Learning Neural Network Digits Classification Relu Softmax Mnist dataset