

다층신경망

배치 및 은닉층 크기에 따른 영향

산업인공지능대학원
2020254011 윤재웅

1. 프로그램 4-4를 수행하여 결과를 정리하고, 프로그램의 동작을 설명하시오

결과] Iteration 85, loss = 0.00074631
Training loss did not improve more than tol=0.000100 for 10 consecutive epochs. Stopping.
[[969 0 3 0 0 1 5 1 3 2]
[0 1123 3 0 1 0 3 1 1 2]
[0 4 1000 1 8 0 2 8 7 0]
[1 1 4 989 0 8 1 4 5 3]
[1 0 3 1 958 0 5 5 3 11]
[1 1 1 6 0 871 5 0 4 3]
[2 2 3 0 3 6 936 0 4 0]
[1 1 7 2 2 0 0 999 5 5]
[3 3 7 8 2 4 1 3 938 1]
[2 0 1 3 8 2 0 7 4 982]]
테스트 집합에 대한 정확률은 97.65 %입니다.
수행시간: 128.13413429260254초
<matplotlib.axes._subplots.AxesSubplot at 0x7f04346cabd0>

- 진행 횟수 : 85번
- 최종 loss : 0.00074631
- 수행시간 : 약 128.1초
- 정확률 : 97.65%



1. 프로그램 4-4를 수행하여 결과를 정리하고, 프로그램의 동작을 설명하시오

동작설명]

```
from sklearn.datasets import fetch_openml
from sklearn.neural_network import MLPClassifier
import matplotlib.pyplot as plt
import numpy as np
import time

start = time.time()
mnist=fetch_openml('mnist_784')
x_train=mnist.data[:60000]; x_test=mnist.data[60000:]
y_train=np.int16(mnist.target[:60000]); y_test=np.int16(mnist.target[60000:])
mlp=MLPClassifier(hidden_layer_sizes=(100),
                  learning_rate_init=0.001,
                  batch_size=128,
                  max_iter=300,
                  solver='adam',
                  verbose=True)

mlp.fit(x_train,y_train)

res=mlp.predict(x_test)
conf=np.zeros((10,10),dtype=np.int16)
for i in range(len(res)):
    conf[res[i]][y_test[i]]+=1
print(conf)
no_correct=0
for i in range(10):
    no_correct+=conf[i][i]
accuracy=no_correct/len(res)

print("테스트 집합에 대한 정확률은", accuracy*100, "%입니다.")
print("수행시간: {0}초".format(str(time.time() - start)))
import matplotlib.pyplot as plt
import seaborn as sns
plt.figure(figsize=(10,7))
sns.set(font_scale=2)
sns.heatmap(conf, annot=True) #heatmap graph
```

- start time으로 수행시간 측정
- MNIST dataset을 다운아 사용
- 학습용으로 6만개의 데이터를 분할하여 확보
- 테스트용으로 나머지 데이터 분할
- 다층 퍼셉트론 분류기 세팅 후 학습

- 테스트 집합으로 예측
- 혼동행렬 구성 및 출력
- 정확률 계산 및 출력

2. Batch size를 128로 하고, 은닉층 size가 50인 경우

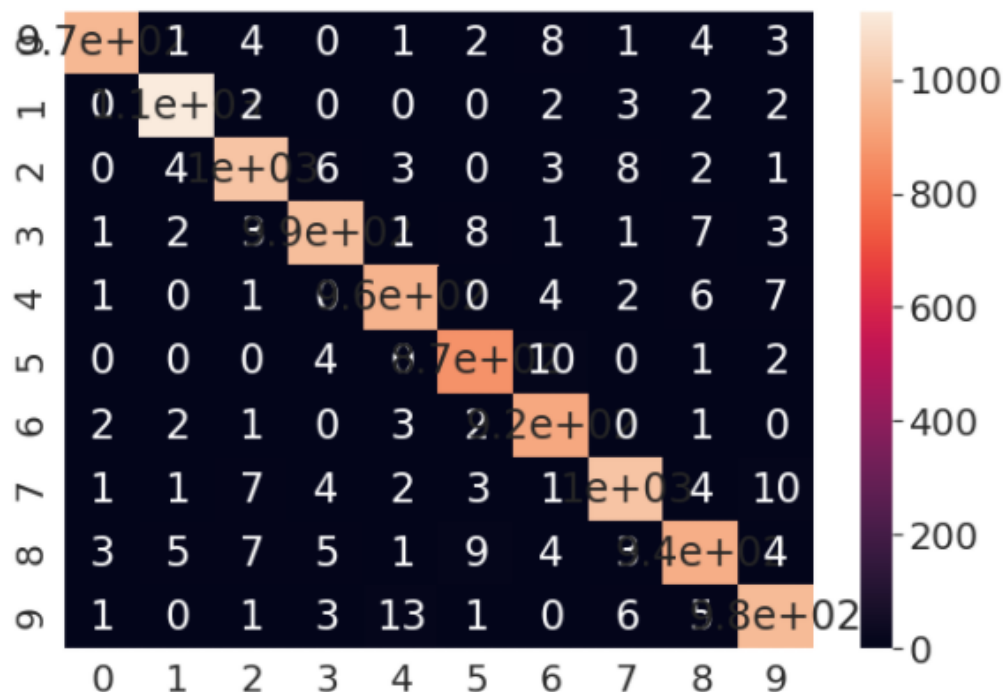
결과] Iteration 52, loss = 0.01497001
Training loss did not improve more than tol=0.000100 for 10 consecutive epochs. Stopping.

```
[[ 971  1  4  0  1  2  8  1  4  3]
 [  0 1120  2  0  0  0  2  3  2  2]
 [  0  4 1006  6  3  0  3  8  2  1]
 [  1  2  3 988  1  8  1  1  7  3]
 [  1  0  1  0 958  0  4  2  6  7]
 [  0  0  0  4  0 867 10  0  1  2]
 [  2  2  1  0  3  2 925  0  1  0]
 [  1  1  7  4  2  3  1 1004  4 10]
 [  3  5  7  5  1  9  4  3 942  4]
 [  1  0  1  3 13  1  0  6  5 977]]
```

테스트 집합에 대한 정확률은 97.58 %입니다.

수행시간: 172.1223030090332초

<matplotlib.axes._subplots.AxesSubplot at 0x7f043cee7e50>



- 진행 횟수 : 52번
- 최종 loss : 0.00125469
- 수행시간 : 약 172.1초
- 정확률 : 97.58%

< 비교 > [512, 100]

- 진행 횟수 : 85번
- 최종 loss : 0.00074631
- 수행시간 : 약 128.1초
- 정확률 : 97.65%

2. Batch size를 128로 하고, 은닉층 size가 50인 경우

분석]

설정값		학습 횟수	정확률	소요시간
Batch size	Hidden layer			
512	100	85	97.65%	128.1 초
512	50	110	97.38%	125.7 초
128	100	50	97.94%	142.3 초
128	50	52	97.58%	172.1 초

- batch와 hidden의 size에 따라 차이가 나타남
- 정확률의 차이는 미비한 것으로 생각됨
- batch size에 의해 소요시간이 많이 달라지는 것을 확인

결론]

- batch size와 hidden layer size에 의해 정확률과 소요시간이 달라지나, 무조건 크다고 하여 좋은 결과를 내는 것은 아니다.