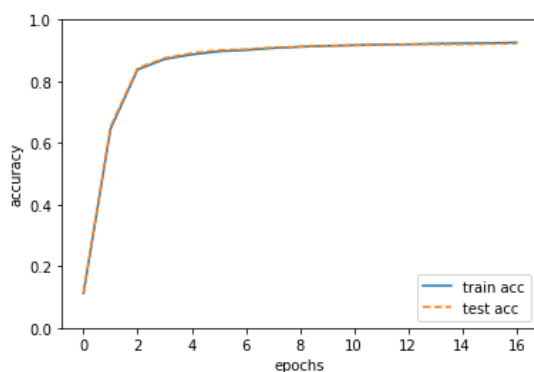


# 딥러닝의 실제\_10주차 과제

2023254025 김혜영

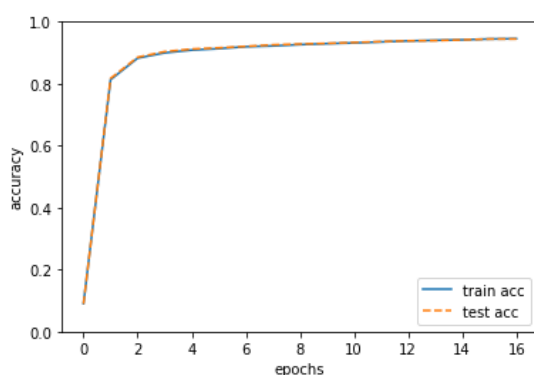
교재 p.143에 있는 프로그램을 수정하여 다음의 경우에 대한 학습 결과 그래프를 출력하십시오.

(1) hidden\_size=10으로 하였을 때 (learning\_rate = 0.1)



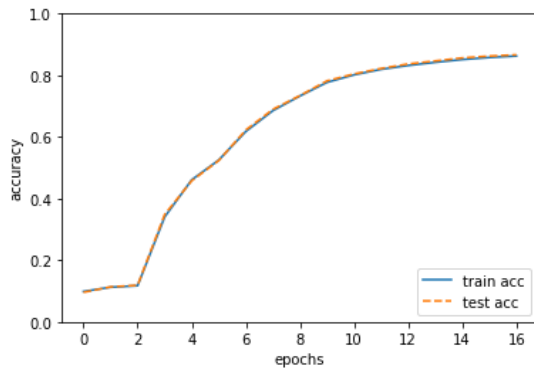
```
In [1]: runfile('C:/Users/User/Desktop/deepLearning/deepLearning_env')
train acc, test acc | 0.11236666666666667, 0.1135
train acc, test acc | 0.6226833333333333, 0.6279
train acc, test acc | 0.82085, 0.8215
train acc, test acc | 0.8730666666666667, 0.8769
train acc, test acc | 0.8880666666666667, 0.8898
train acc, test acc | 0.8976166666666666, 0.8985
train acc, test acc | 0.9039, 0.9058
train acc, test acc | 0.9077, 0.909
train acc, test acc | 0.91165, 0.9124
train acc, test acc | 0.9144, 0.9168
train acc, test acc | 0.9169833333333334, 0.918
train acc, test acc | 0.9178833333333334, 0.9202
train acc, test acc | 0.9207333333333333, 0.9213
train acc, test acc | 0.9223833333333333, 0.9208
train acc, test acc | 0.9236333333333333, 0.9227
train acc, test acc | 0.9251166666666667, 0.924
train acc, test acc | 0.9263666666666667, 0.9252
```

(2) hidden\_size=100으로 하였을 때 (learning\_rate = 0.1)



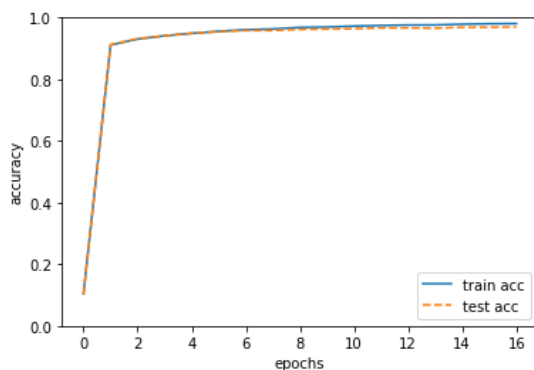
```
In [1]: runfile('C:/Users/User/Desktop/deepLearning/deepLearning_env')
train acc, test acc | 0.09035, 0.0892
train acc, test acc | 0.8129333333333333, 0.817
train acc, test acc | 0.88255, 0.8854
train acc, test acc | 0.8996333333333333, 0.9036
train acc, test acc | 0.90865, 0.9119
train acc, test acc | 0.91315, 0.9158
train acc, test acc | 0.9195, 0.9205
train acc, test acc | 0.92205, 0.9256
train acc, test acc | 0.9262333333333334, 0.9281
train acc, test acc | 0.92945, 0.9302
train acc, test acc | 0.9318166666666666, 0.9332
train acc, test acc | 0.9349666666666666, 0.936
train acc, test acc | 0.93815, 0.9385
train acc, test acc | 0.9402, 0.9391
train acc, test acc | 0.9418, 0.9417
train acc, test acc | 0.9444166666666667, 0.9443
train acc, test acc | 0.9452666666666667, 0.9444
```

(3) learning\_rate= 0.01일 때 (hidden\_size = 50)



```
In [1]: runfile('C:/Users/User/Desktop/deepLearning
deepLearning_env')
train acc, test acc | 0.09863333333333334, 0.0958
train acc, test acc | 0.1124, 0.1135
train acc, test acc | 0.11731666666666667, 0.1187
train acc, test acc | 0.3404, 0.3482
train acc, test acc | 0.46086666666666665, 0.4579
train acc, test acc | 0.5250833333333333, 0.5235
train acc, test acc | 0.6178333333333333, 0.6226
train acc, test acc | 0.6853, 0.6896
train acc, test acc | 0.7329333333333333, 0.7343
train acc, test acc | 0.7771333333333333, 0.782
train acc, test acc | 0.8015666666666666, 0.8045
train acc, test acc | 0.8194333333333333, 0.8222
train acc, test acc | 0.8319833333333333, 0.8369
train acc, test acc | 0.8421833333333333, 0.8466
train acc, test acc | 0.8512166666666666, 0.8565
train acc, test acc | 0.8579, 0.8624
train acc, test acc | 0.8626833333333334, 0.8666
```

(4) learning\_rate=0.5 일 때 (hidden\_size = 50)



```
In [1]: runfile('C:/Users/User/Desktop/deepLearning
deepLearning_env')
train acc, test acc | 0.10441666666666667, 0.1028
train acc, test acc | 0.9111, 0.9123
train acc, test acc | 0.9307166666666666, 0.9321
train acc, test acc | 0.9413333333333334, 0.9414
train acc, test acc | 0.9491333333333334, 0.9495
train acc, test acc | 0.9558, 0.9542
train acc, test acc | 0.9608333333333333, 0.9584
train acc, test acc | 0.9633333333333334, 0.9585
train acc, test acc | 0.9682833333333334, 0.962
train acc, test acc | 0.9695666666666667, 0.9632
train acc, test acc | 0.9726833333333333, 0.9641
train acc, test acc | 0.9745333333333334, 0.9669
train acc, test acc | 0.9761333333333333, 0.9667
train acc, test acc | 0.9764166666666667, 0.9658
train acc, test acc | 0.9787333333333333, 0.9689
train acc, test acc | 0.9803333333333333, 0.9687
train acc, test acc | 0.9807333333333333, 0.9699
```

(5) 위의 시뮬레이션 결과를 이용하여, hidden\_size 와 learning\_rate가 학습에 미치는 영향을 기술하시오.

- **hidden\_size (hidden state의 벡터 크기)**

신경망의 은닉층에 존재하는 뉴런의 수로, 모델의 용량(capacity)을 결정

⇒ 모델이 학습할 수 있는 복잡성의 정도를 나타냄

- hidden\_size 값이 클 때 → 모델이 더 복잡한 함수를 모델링할 수 있어 학습 데이터에 대한 세밀한 패턴을 찾아낼 수 있지만, 과적합(overfitting) 발생 위험이 있어서 테스트 데이터에 대해 일반화 성능이 떨어질 수 있음
- hidden\_size 값이 작을 때 → 모델이 너무 단순해져 학습 데이터의 패턴을 충분히 학습하지 못할 수 있음 (=과소적합)

- **learning\_rate (학습률)**

한 번 학습할 때 얼마만큼 학습해야 할 지 학습 양을 의미하며, 한 번의 학습량으로 학습한 이후 가중치 매개변수가 갱신됨

경사하강법에서 사용되는 스텝 크기 조절

- learning\_rate 값이 클 때 → 파라미터 업데이트가 크게 이루어져 빠른 수렴이 가능하지만, 너무 높으면 학습 과정에서 발산할 수 있음
- learning\_rate 값이 작을 때 → 수렴 속도가 느리며 많은 학습 반복을 필요로 하므로 최적의 수렴에 도달하지 못할 수 있음