

MNIST test 보고서(이창욱)

데이터:

train data 60000개, test data 10000개로 이루어진 손글씨 숫자 데이터

➔ 주어진 모델에서 hyper parameter들을 변경해 모델의 accuracy를 올리기

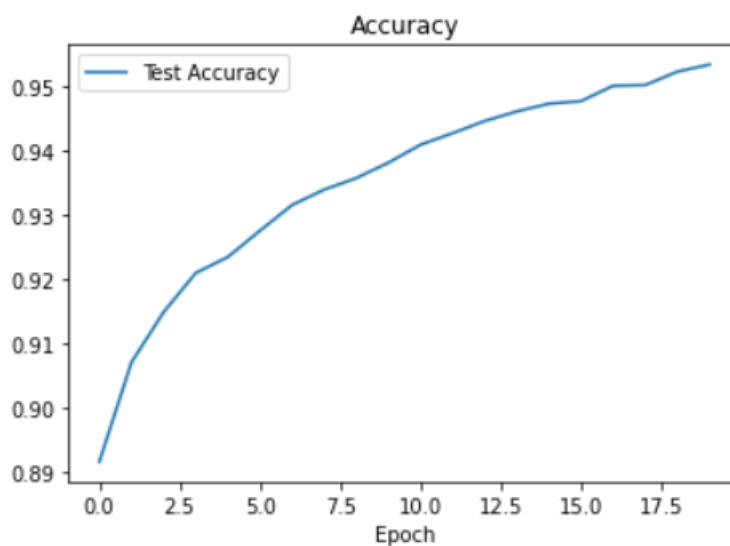
기존 모델:

```
nn.Linear(28 * 28, 64),  
nn.ReLU(),  
nn.Linear(64, 10),  
nn.Sigmoid()
```

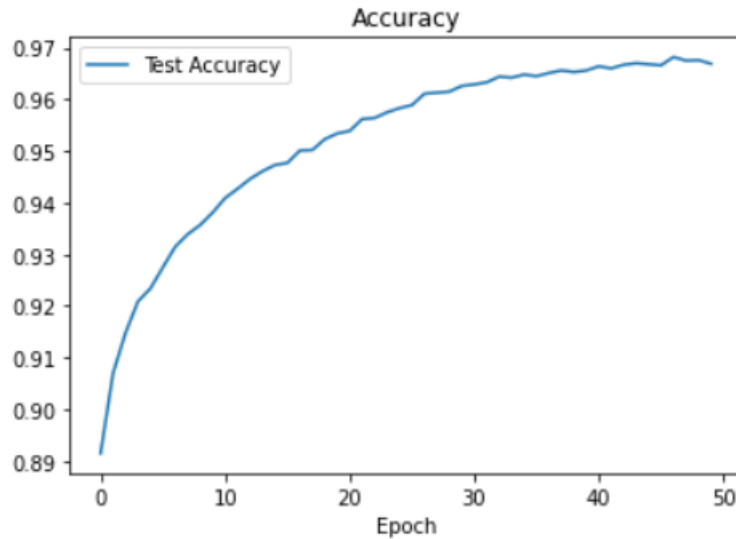
batch size: 16

epoch: 20

1. Epoch이 커지면서 Test Accuracy 역시 커지는 모습이 보인다.

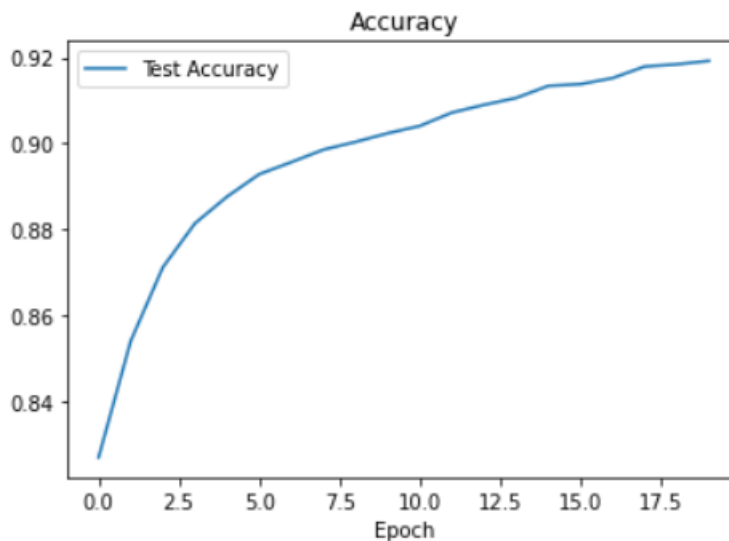


Overfitting의 위험에서는 아직 자유롭다고 볼 수 있다. 따라서 epoch을 키워서 accuracy를 높여보자.

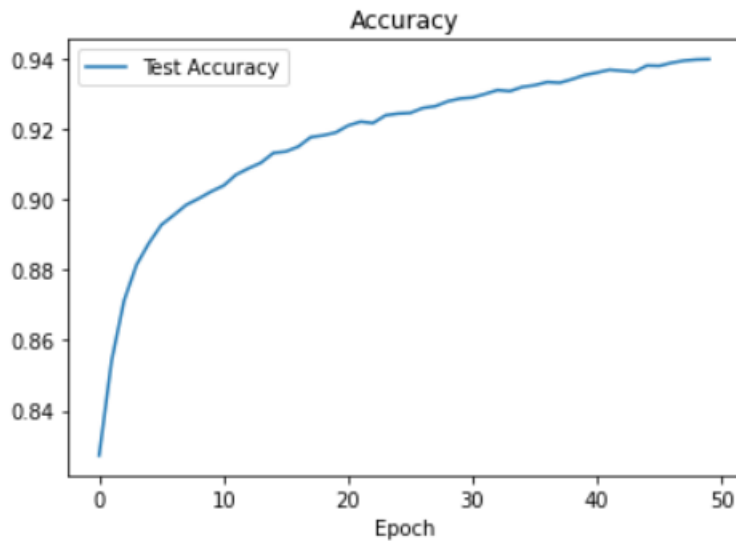


대략 45 정도에서 더 이상 높아지지 않는 것을 확인할 수 있다.

2. Batch size 256으로 크게 하기



20번의 epoch 후의 test accuracy는 0.9192로 오히려 낮아진 것을 확인할 수 있다. 이는 batch size가 너무 커져서 20번의 epoch만으로는 학습이 충분히 진행되지 못했음을 보여 준다. 따라서 epoch를 50번으로 키워주면



여전히 accuracy가 이전 모델보다 못한 것을 확인할 수 있다.

따라서 최종 모델은 epoch 45, batch는 그대로 16으로 하는 모델로 선정했다.