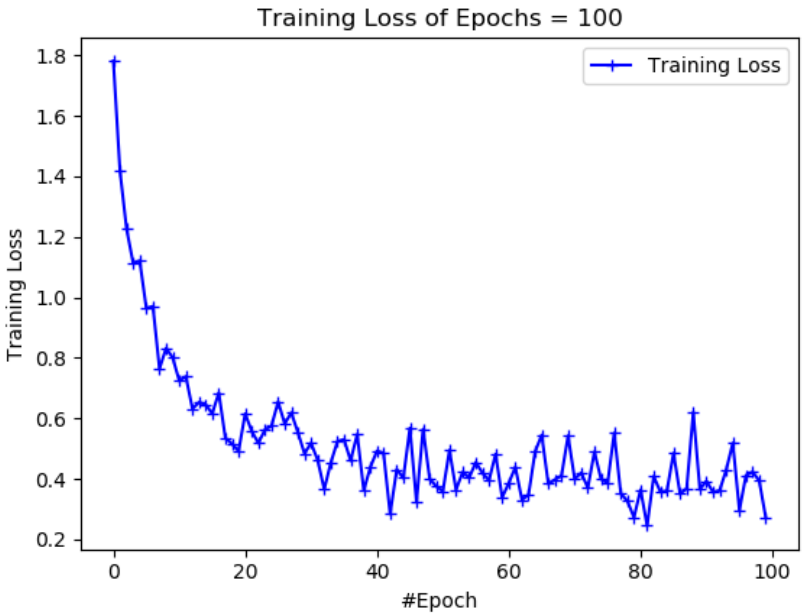
Mnist_1.py results:

Using the original parameters:

```
Finished Training
Accuracy of the network on the 10000 test images: 87 %
```

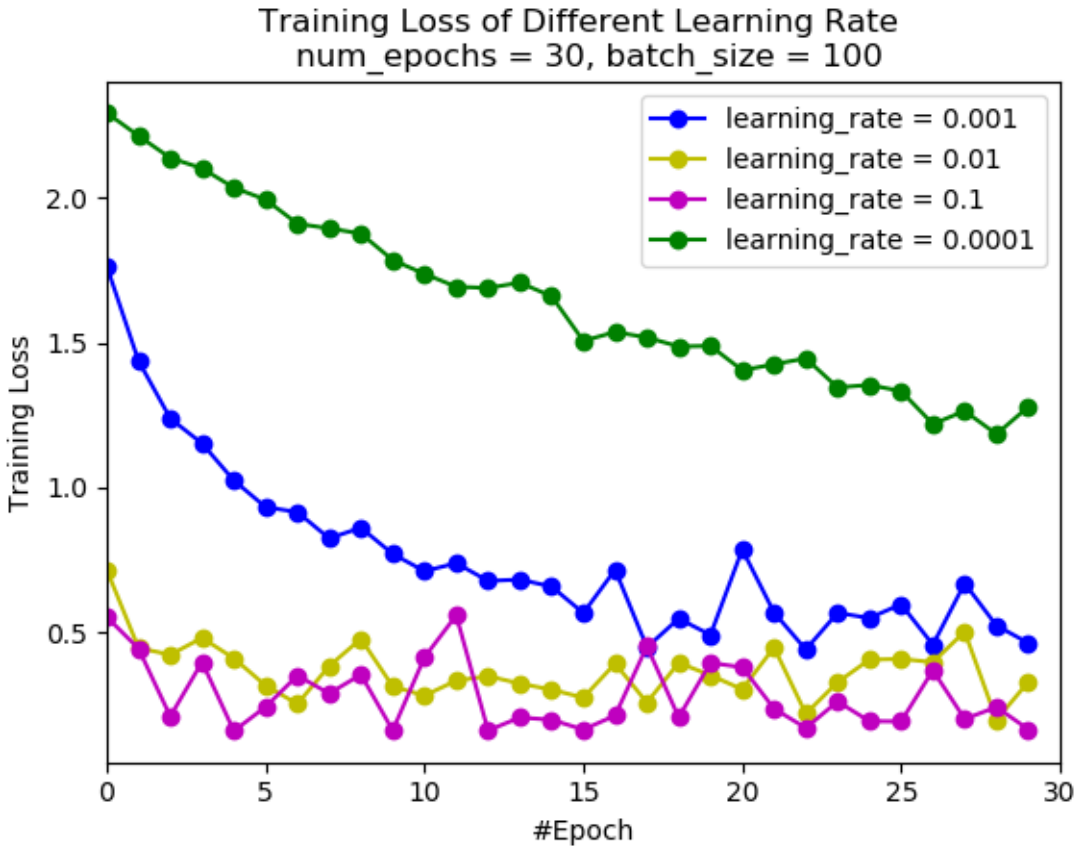
Using 100 Epochs:

```
Finished Training
Accuracy of the network on the 10000 test images: 90 %
```

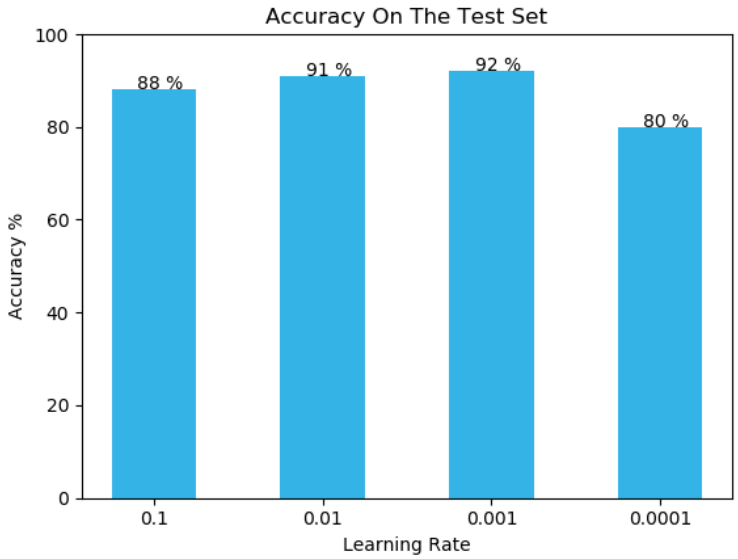


Mnist_2.py results:

Finding a better optimization configuration, trying different learning rates:

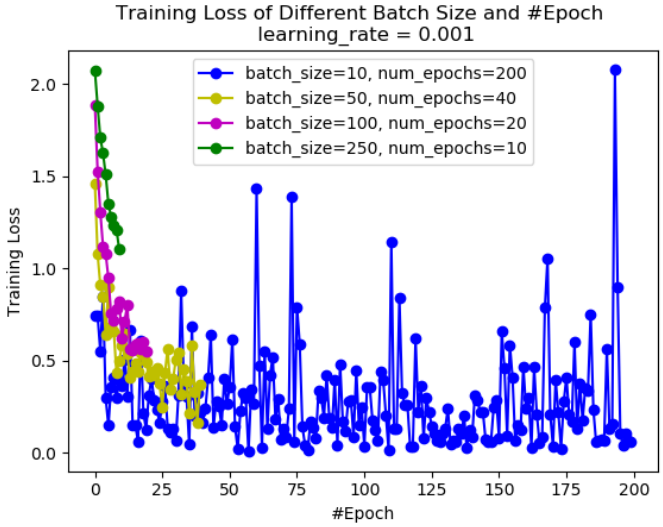


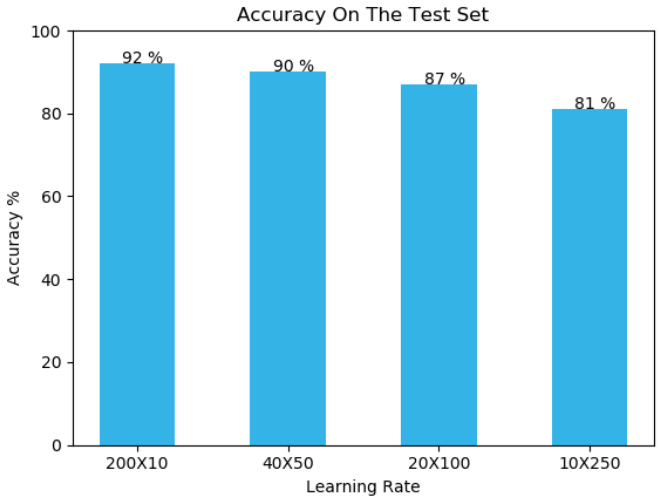
and ploting accuracy to choose the best one:



Next tried to trade-off different batch-sizes and number of epochs,

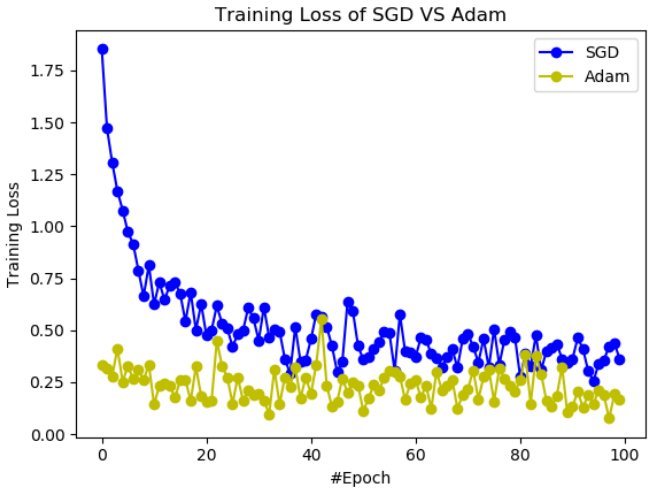
Since increasing batch-size without decreasing the number of epochs will train the NN over more ‘data’ thus providing a false comparison.

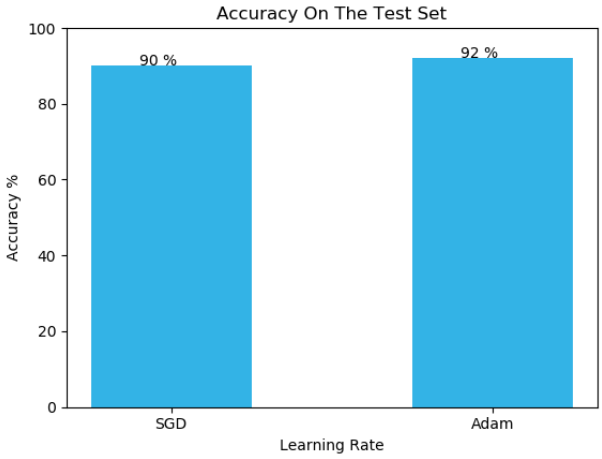




Since the results were inconclusive, I chose batch-size = 200, without changing the number of epochs.

Testing Adam optimization method:





The final parameters I chose: batch_size=200, learning_rate=0.001, with Adam algorithm as the optimizer.

Mnist_3.py results:

Testing the original vs improved vs 3-layer-ReLU: