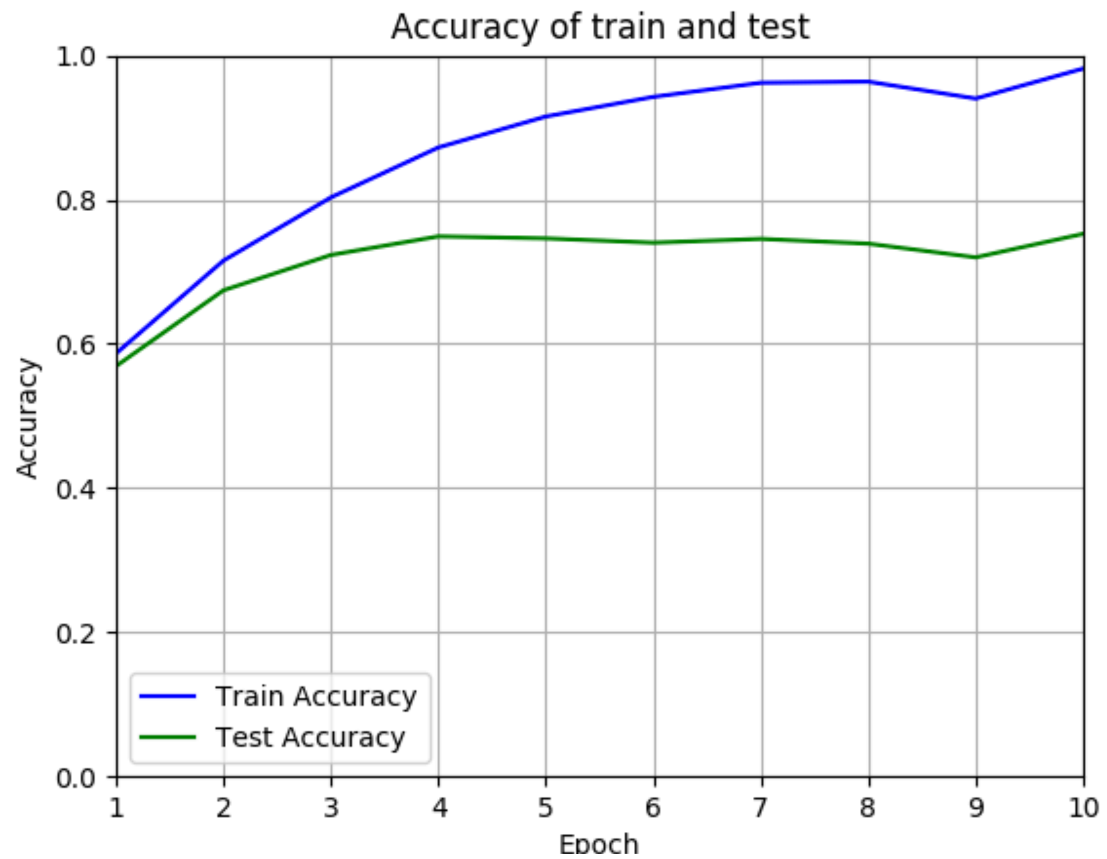
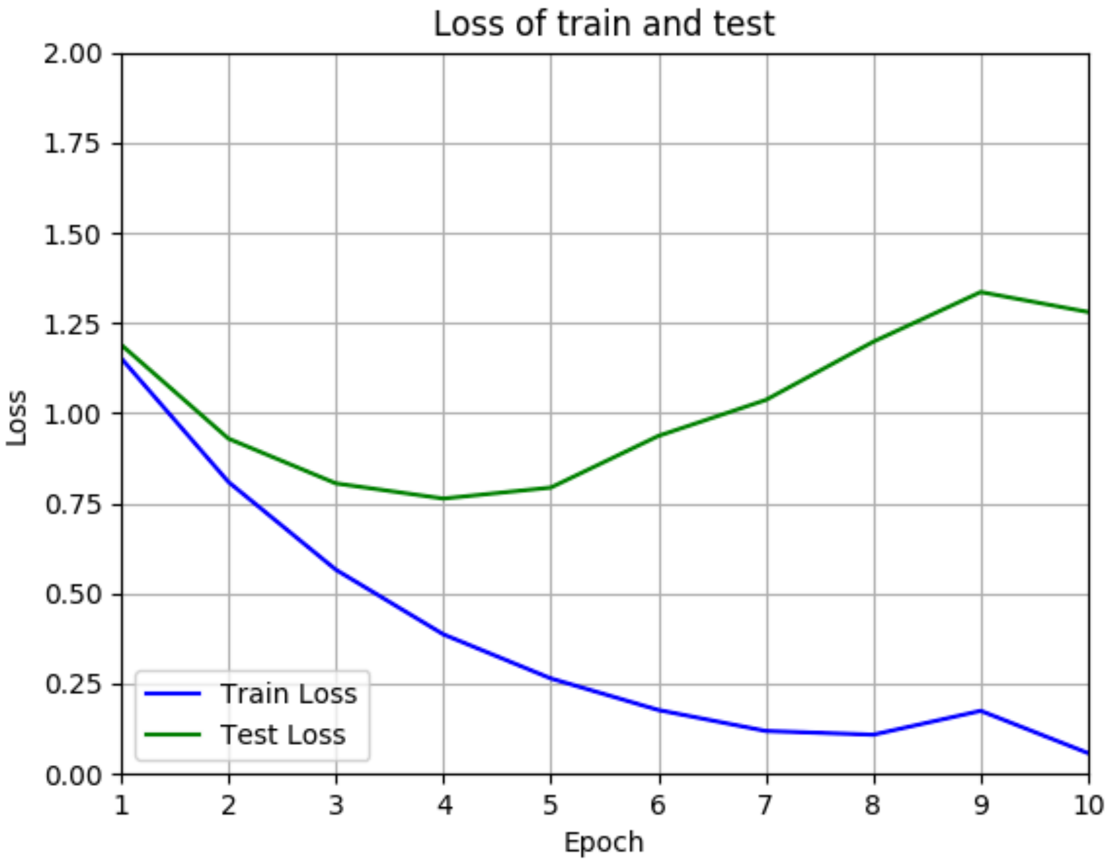
CS 535 Deep Learning
Assignment 3

Mohamad Hosein Danesh
ID: 933619683

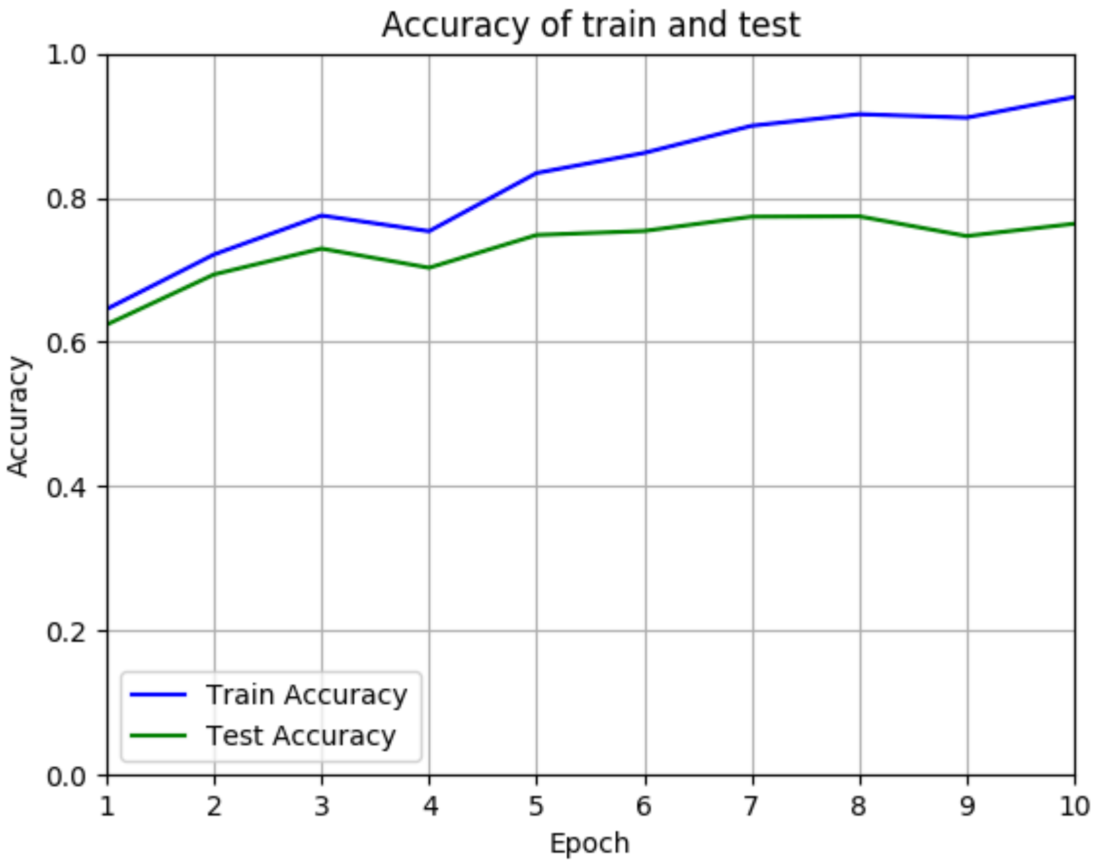
March 4, 2019

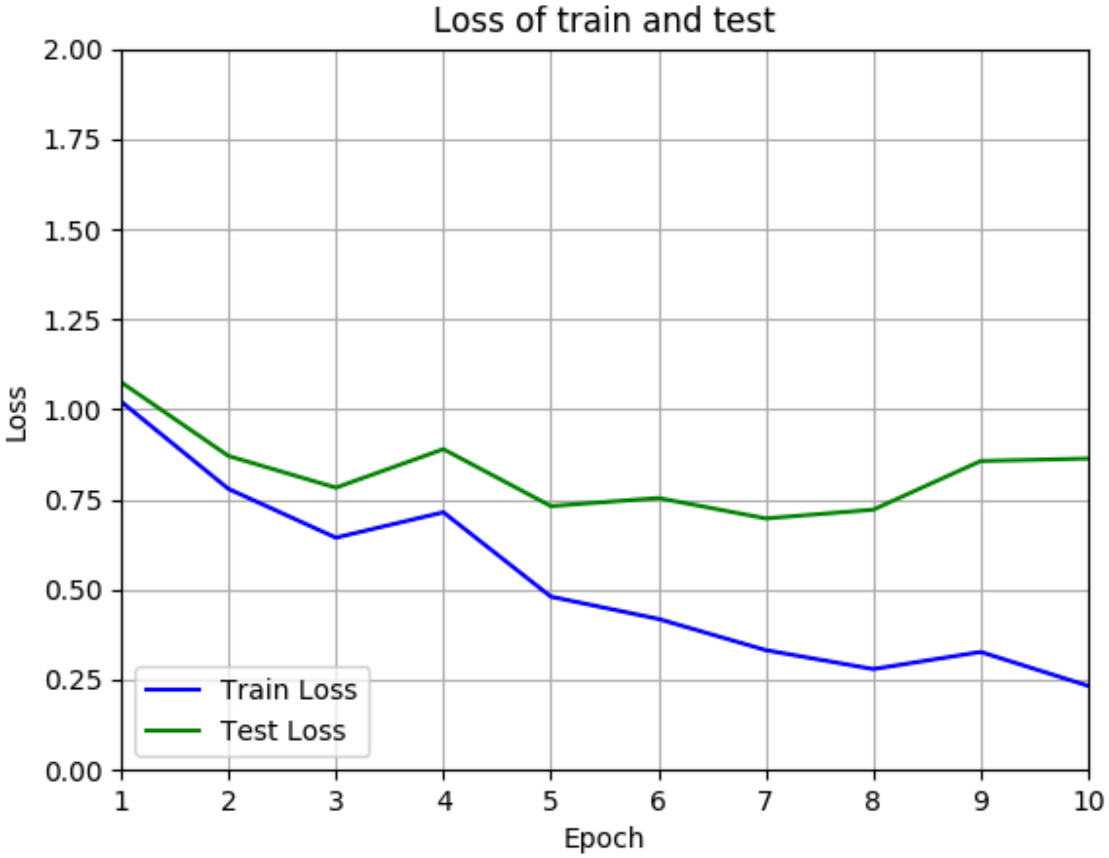
Setting 0) Just run the template code and plot the results.



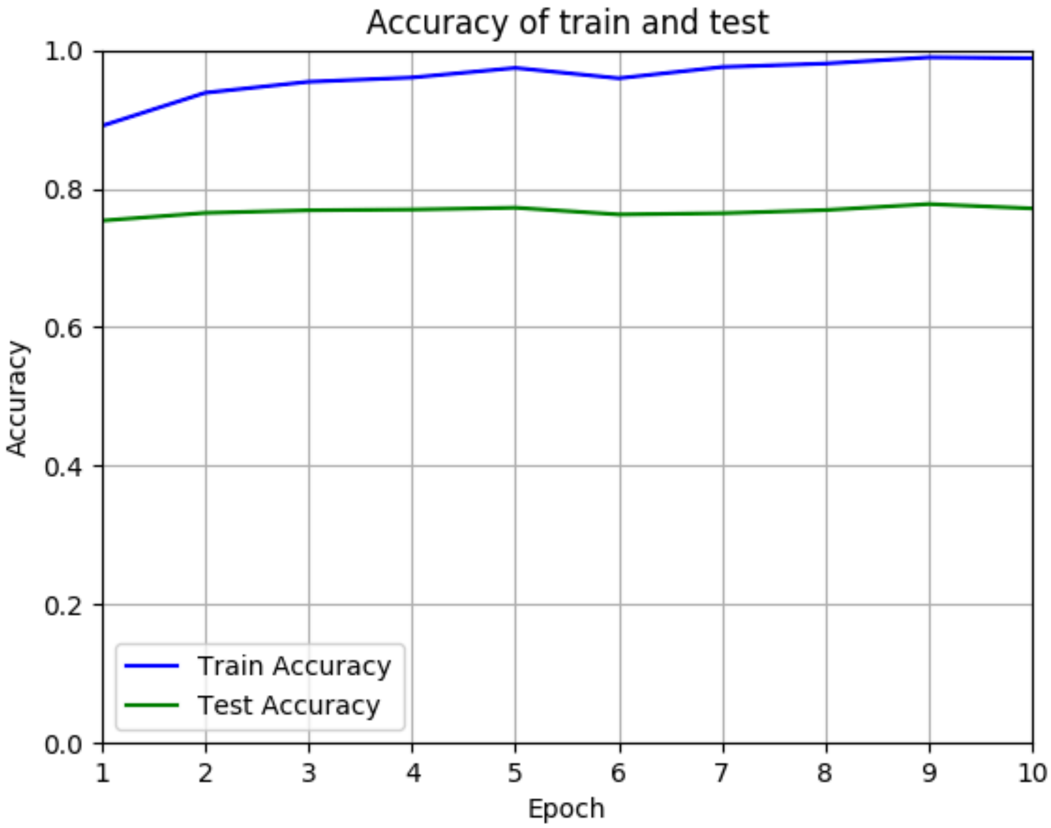


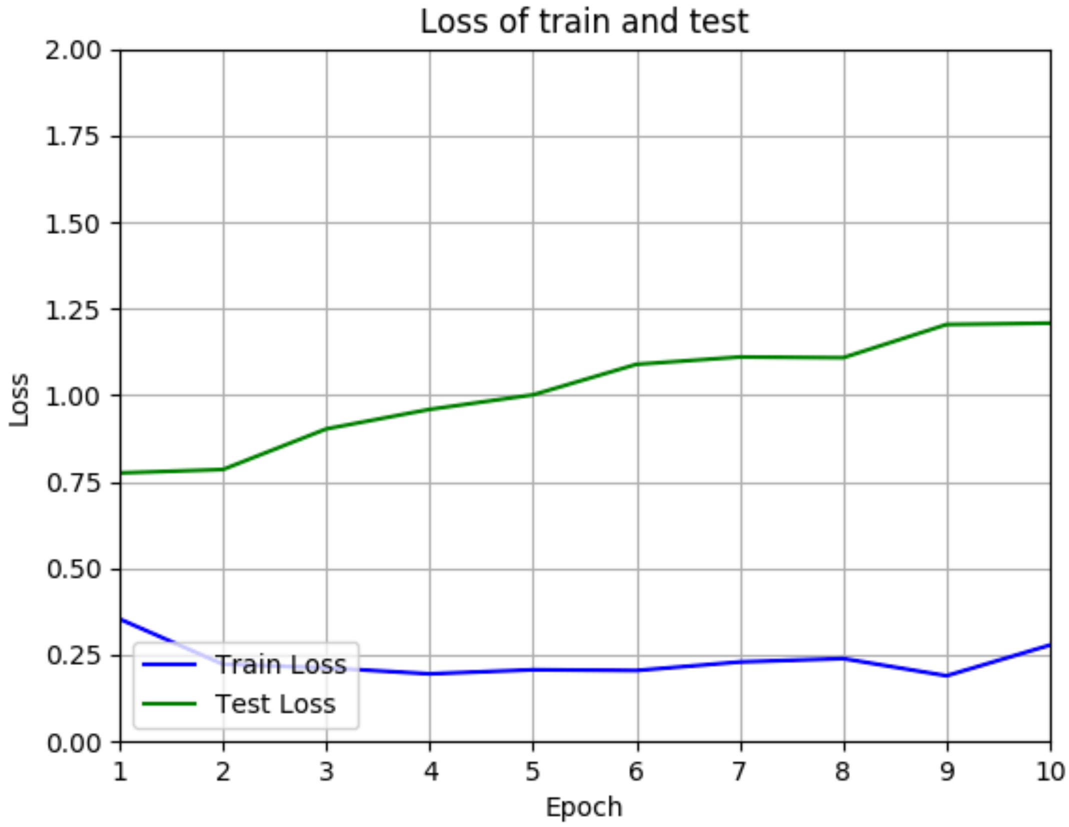
Setting 1) A batch normalization layer has been added right after the first fully connected layer. It performs not as good as the template code, which its results are shown in setting 0.





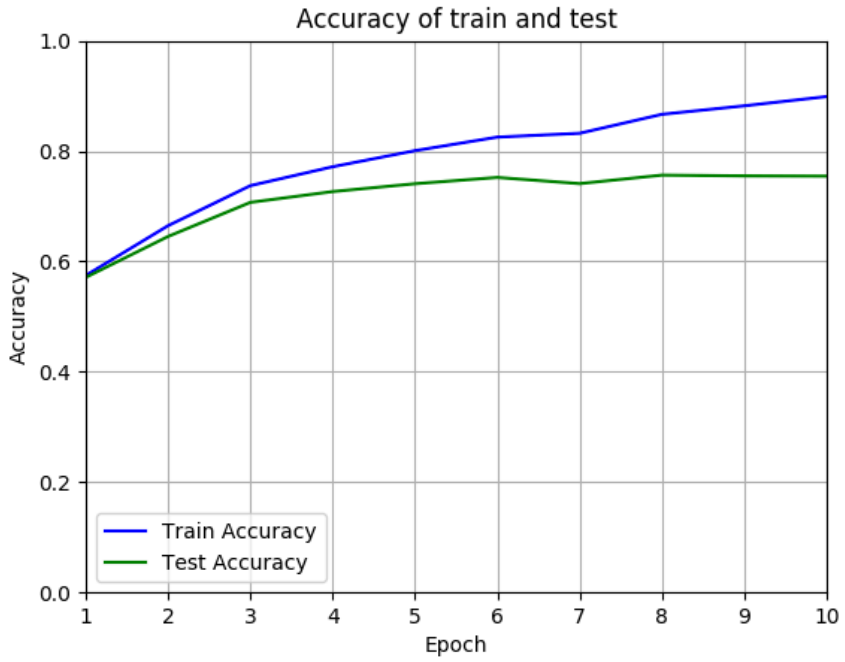
Setting 2) A new fully connected layer has been added. Comparing the results of setting 2 to previous ones, it can be inferred that the network performs better when it uses a pre-trained model and fine tuned based on that model

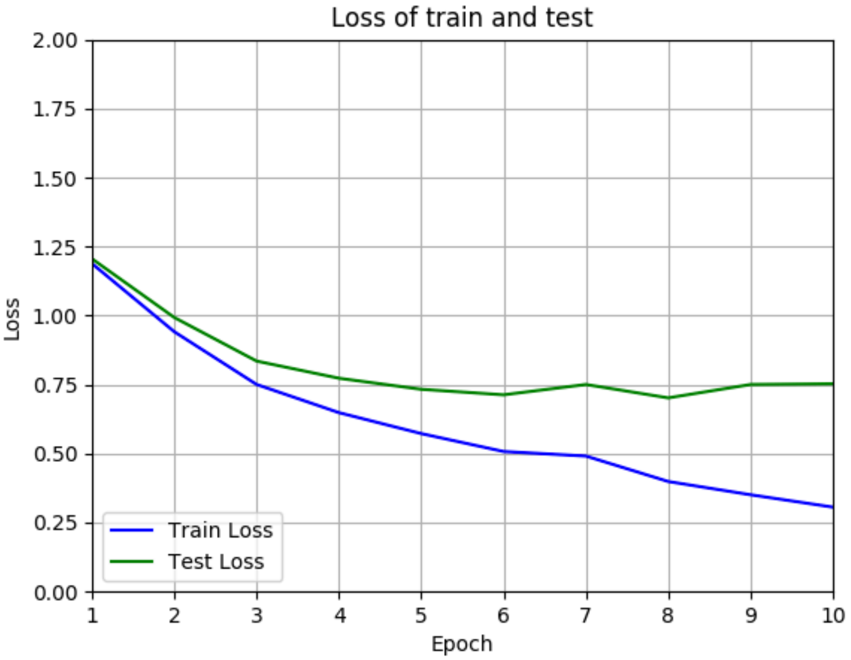




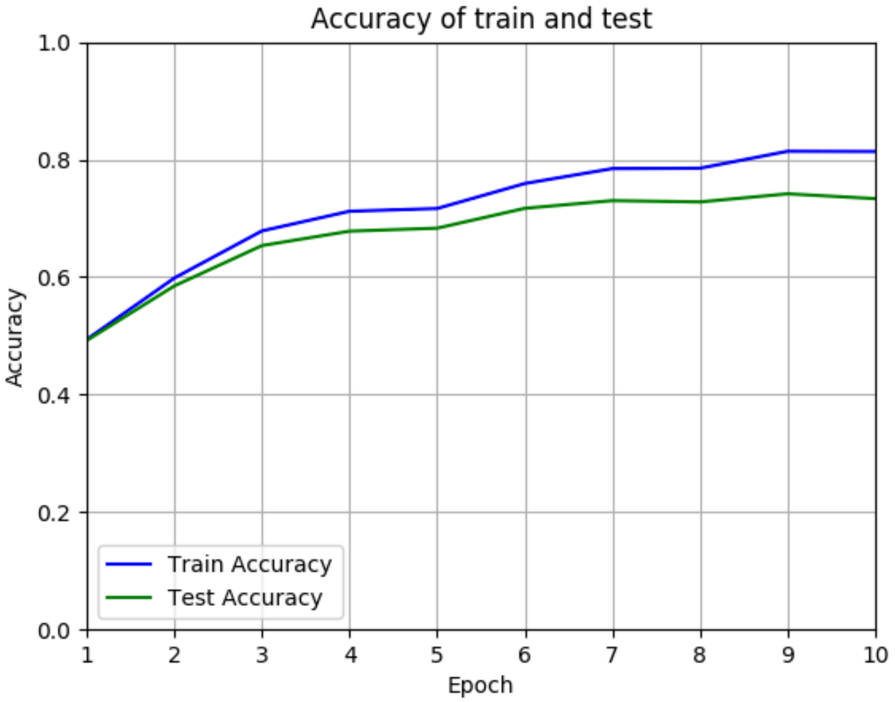
Setting 3) Testing different adaptive schedule algorithms.

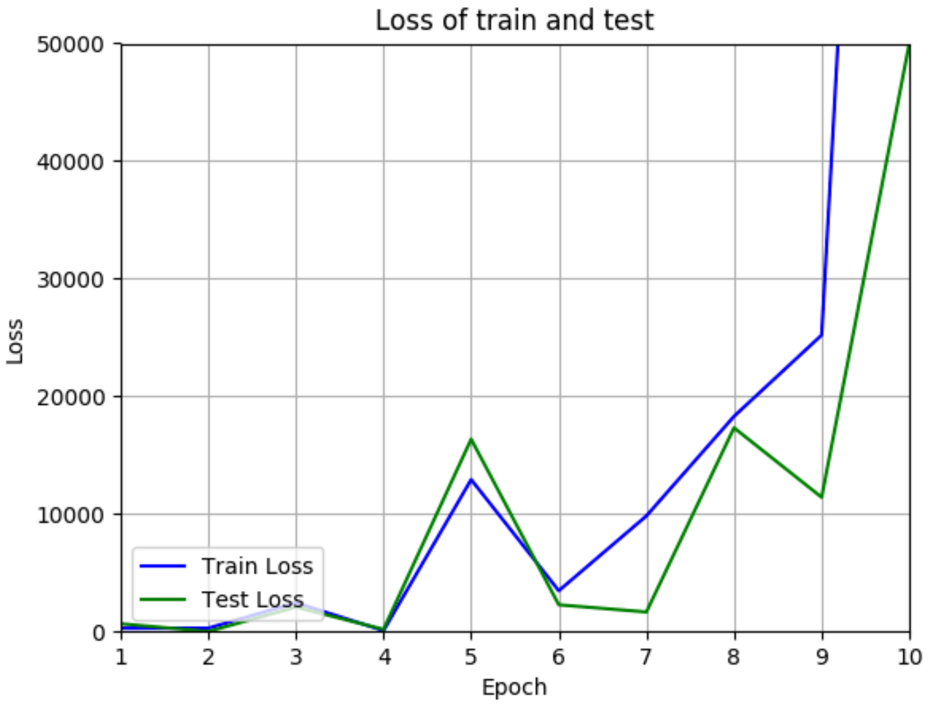
Adagrad: learning rate = 0.01



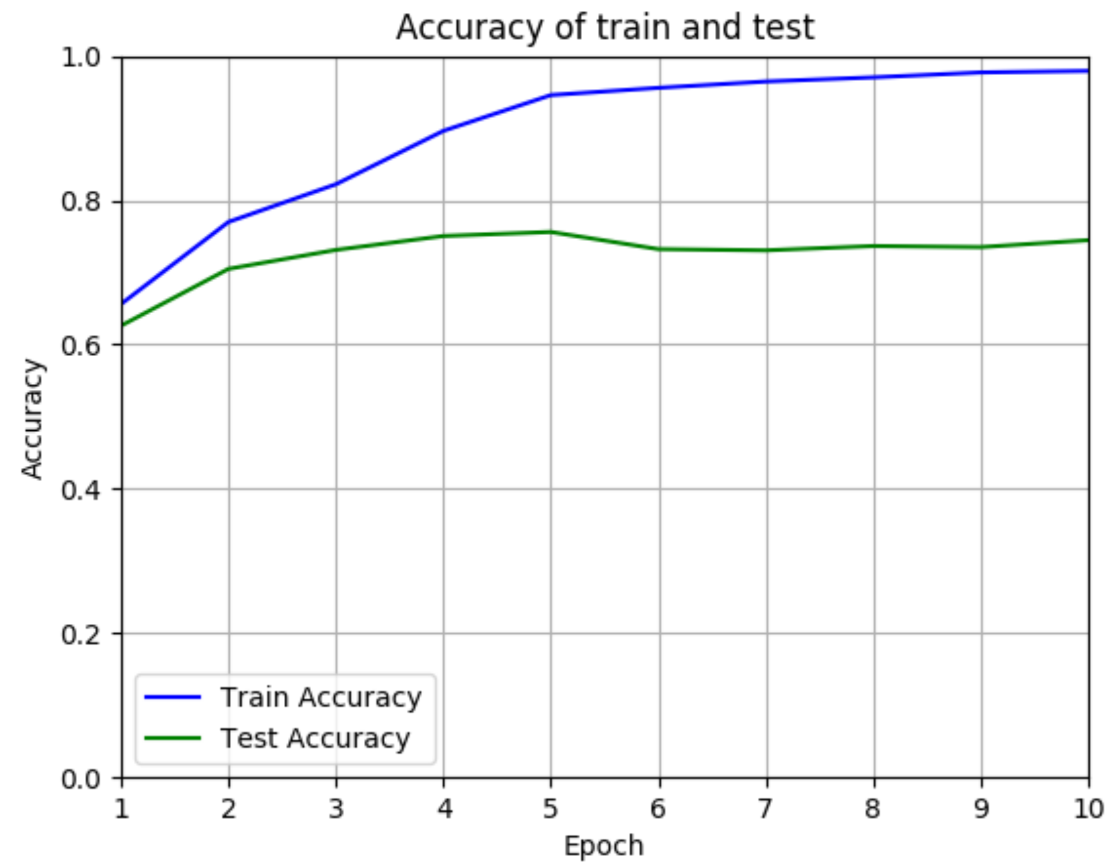


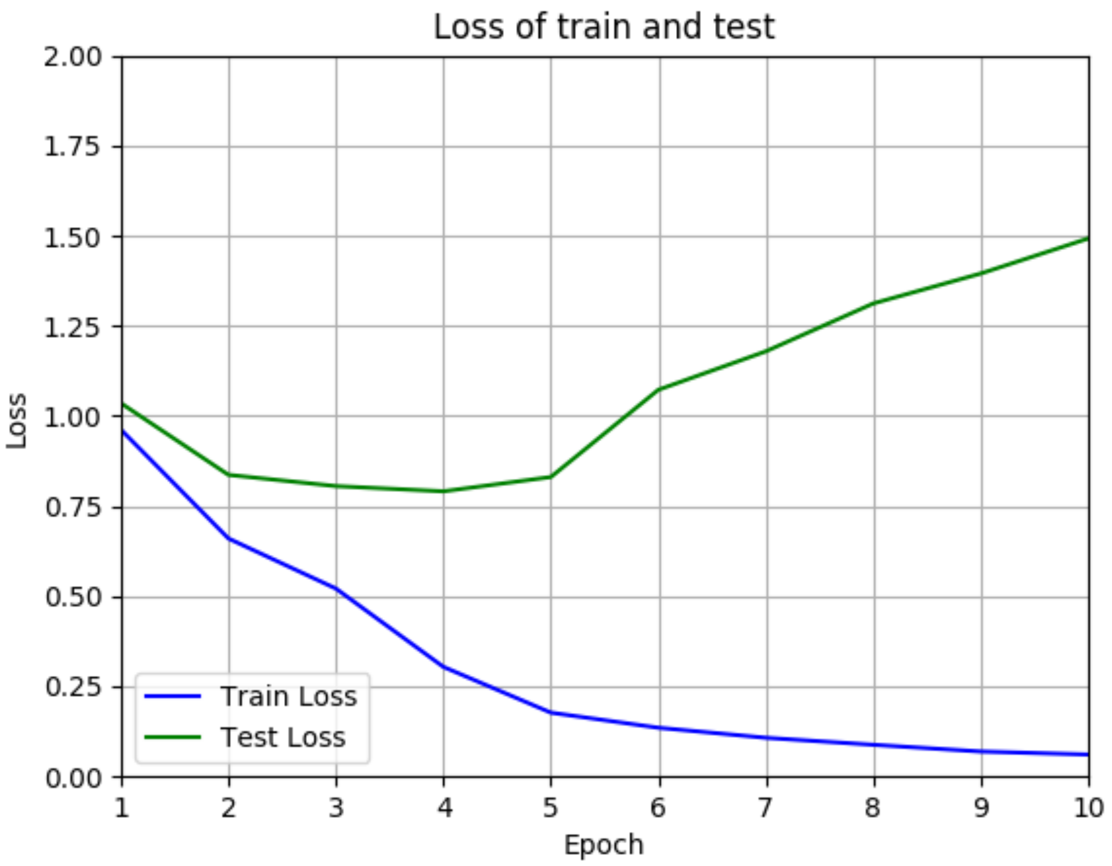
Adam: learning rate = 0.01. With this learning rate, training and testing losses seem awful. By reducing the learning rate, the network performed better and converged.



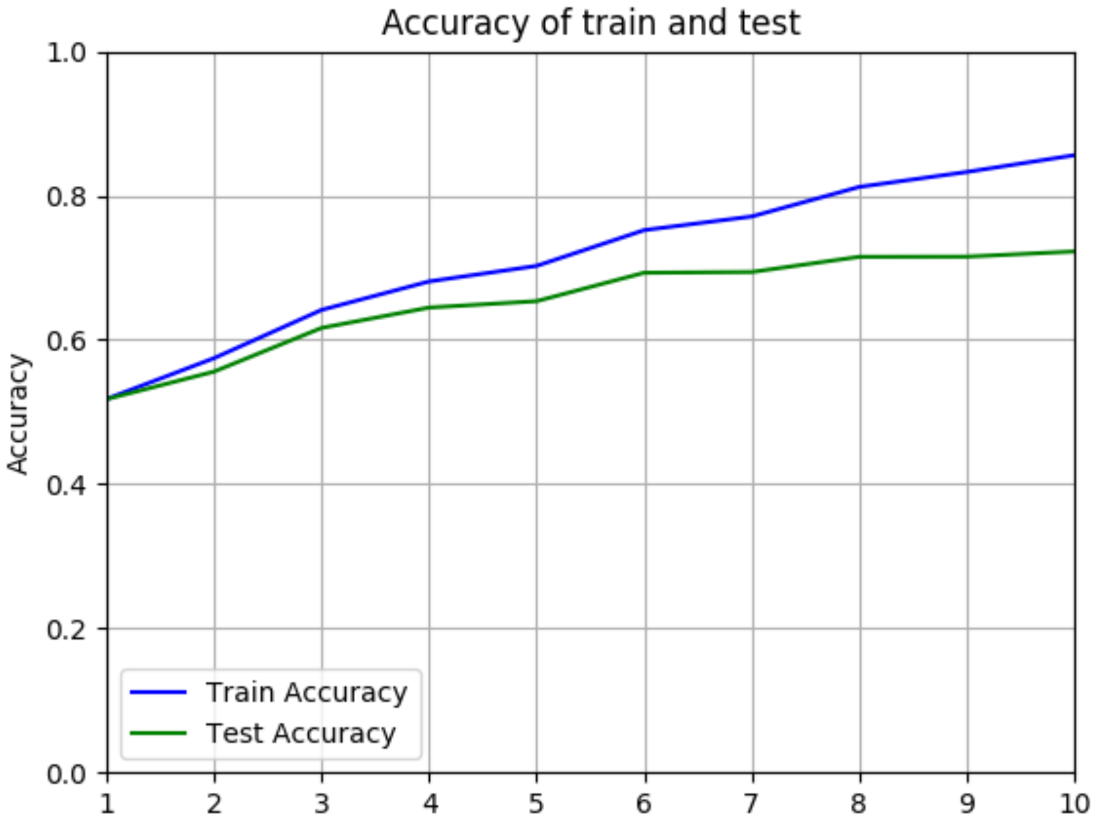


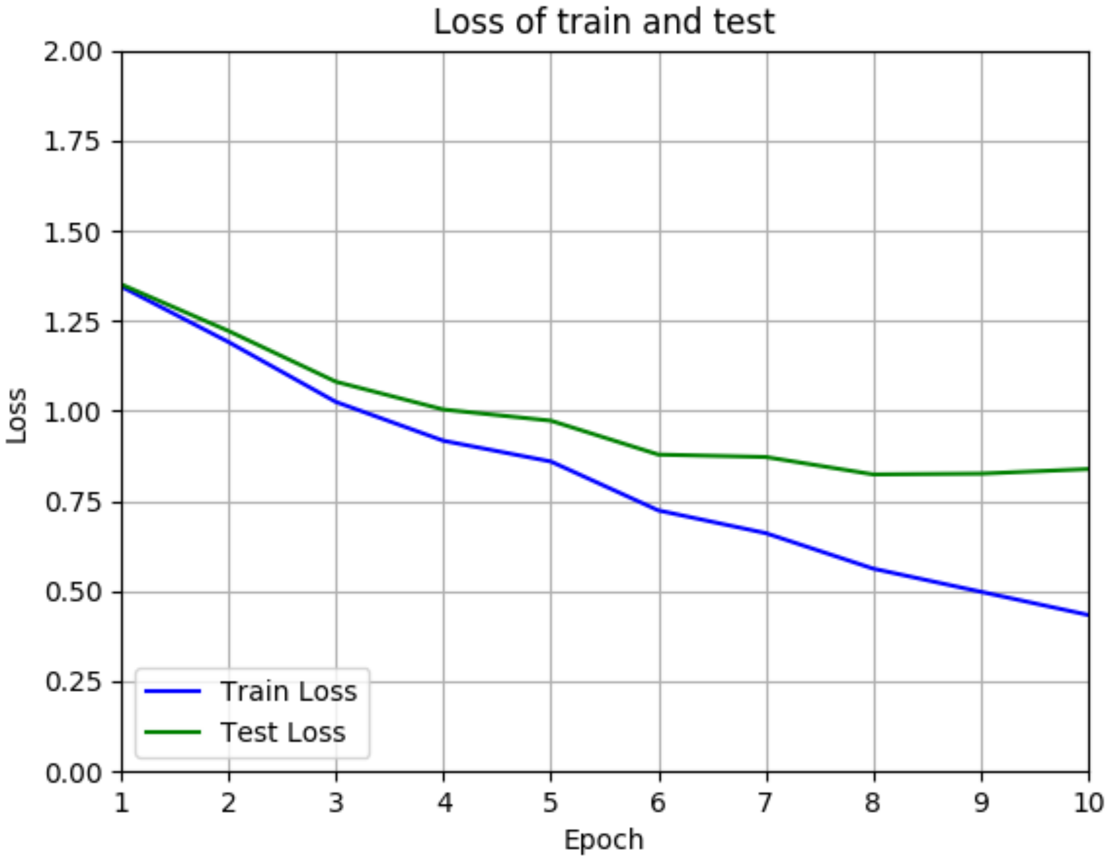
Adam: learning rate = 0.001. With this learning rate, I got the best results with Adam.



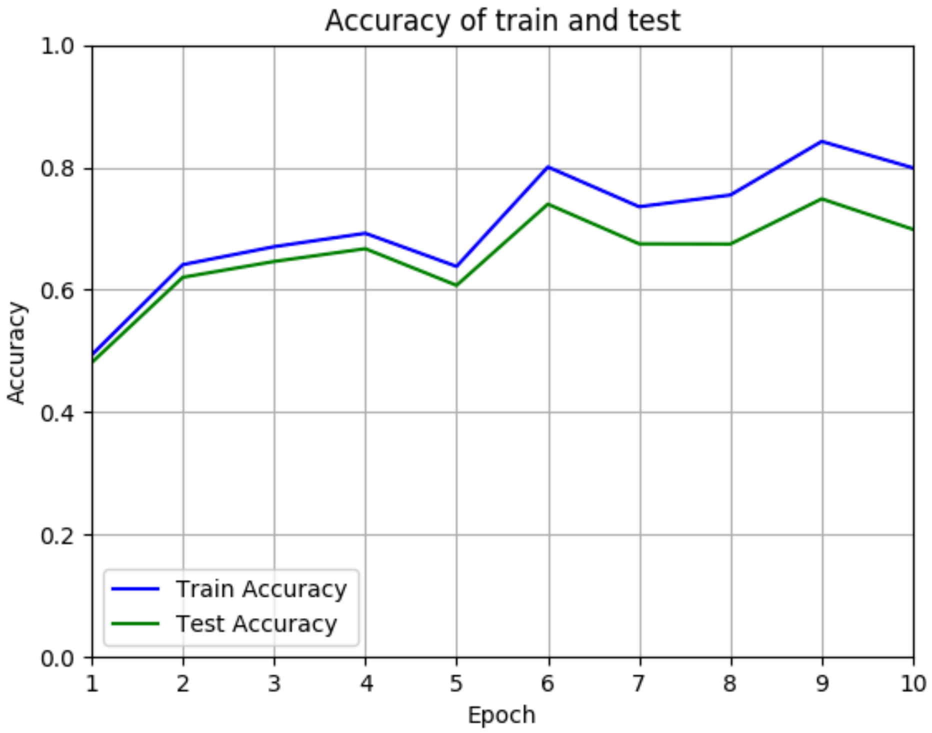


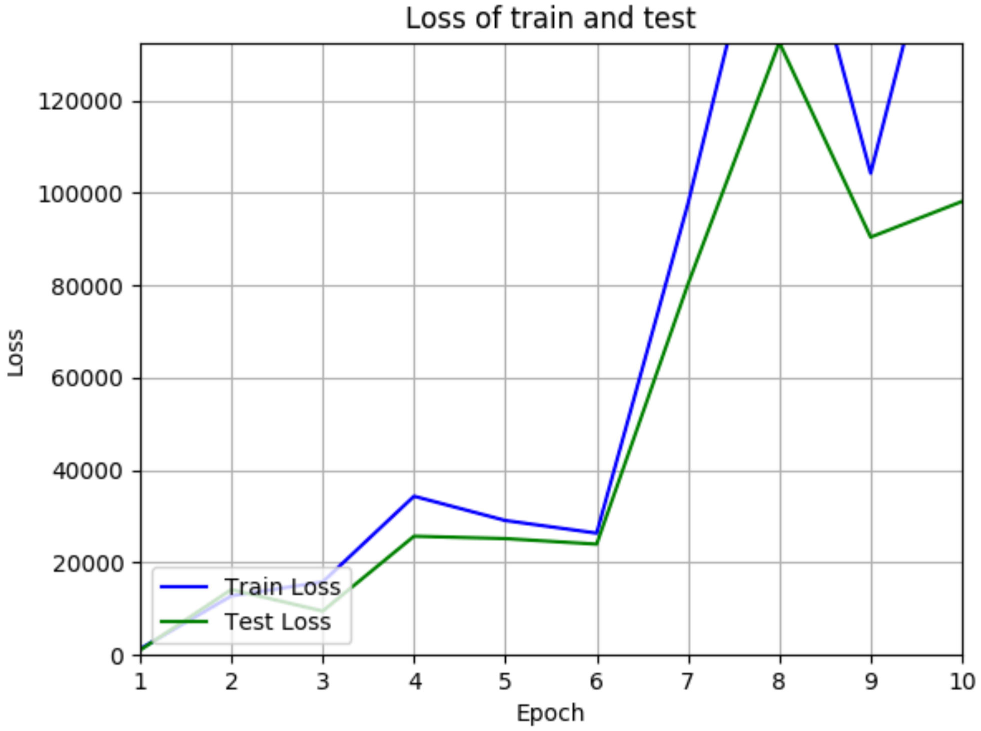
Adam: learning rate = 0.0001. With this learning rate, I got better results in loss of testing set comparing to other learning rates of Adam.



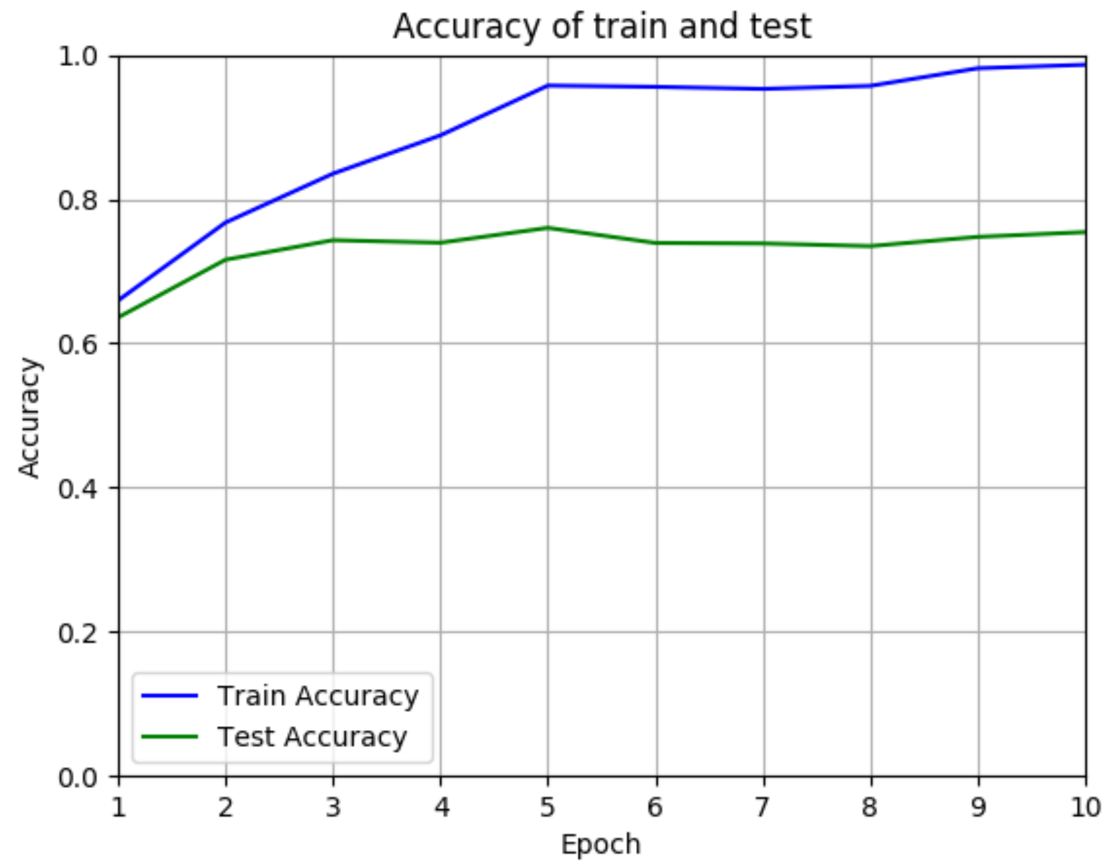


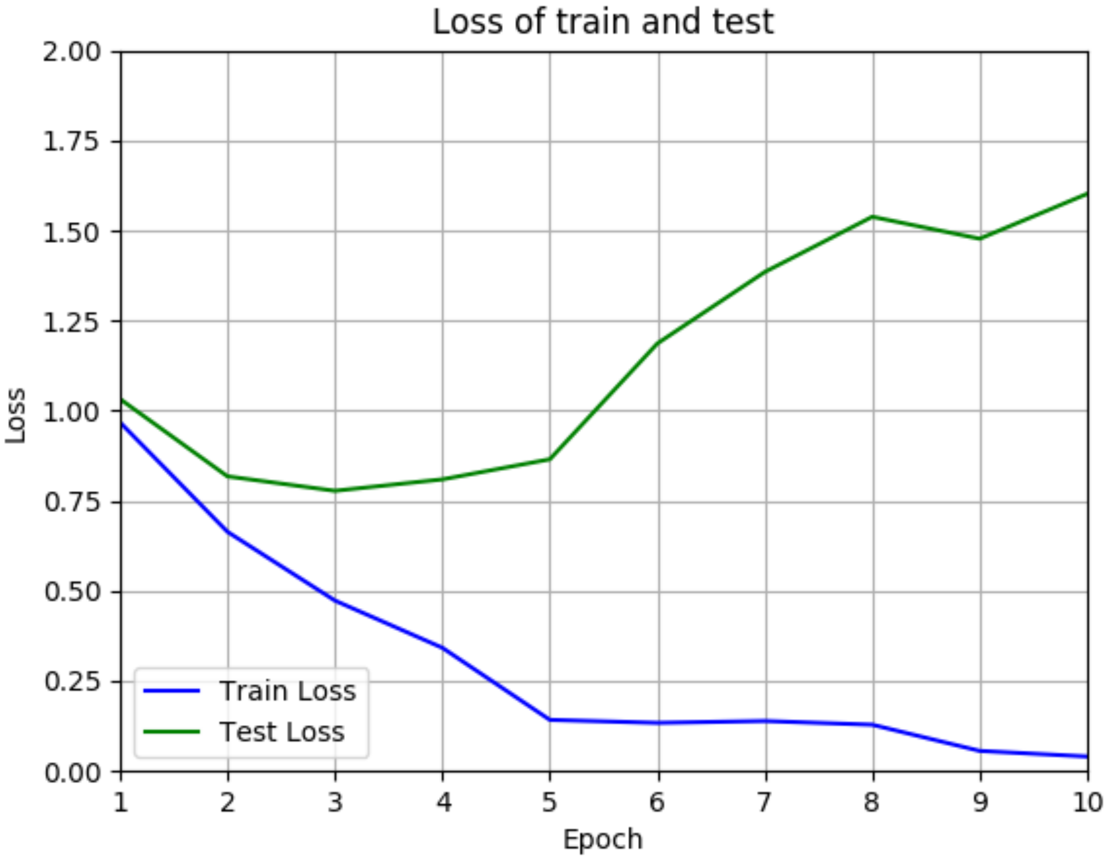
RMSprop: learning rate = 0.01. The same as Adam, with this learning rate, training and testing losses seem awful. By reducing the learning rate, the network performed better and converged finally.



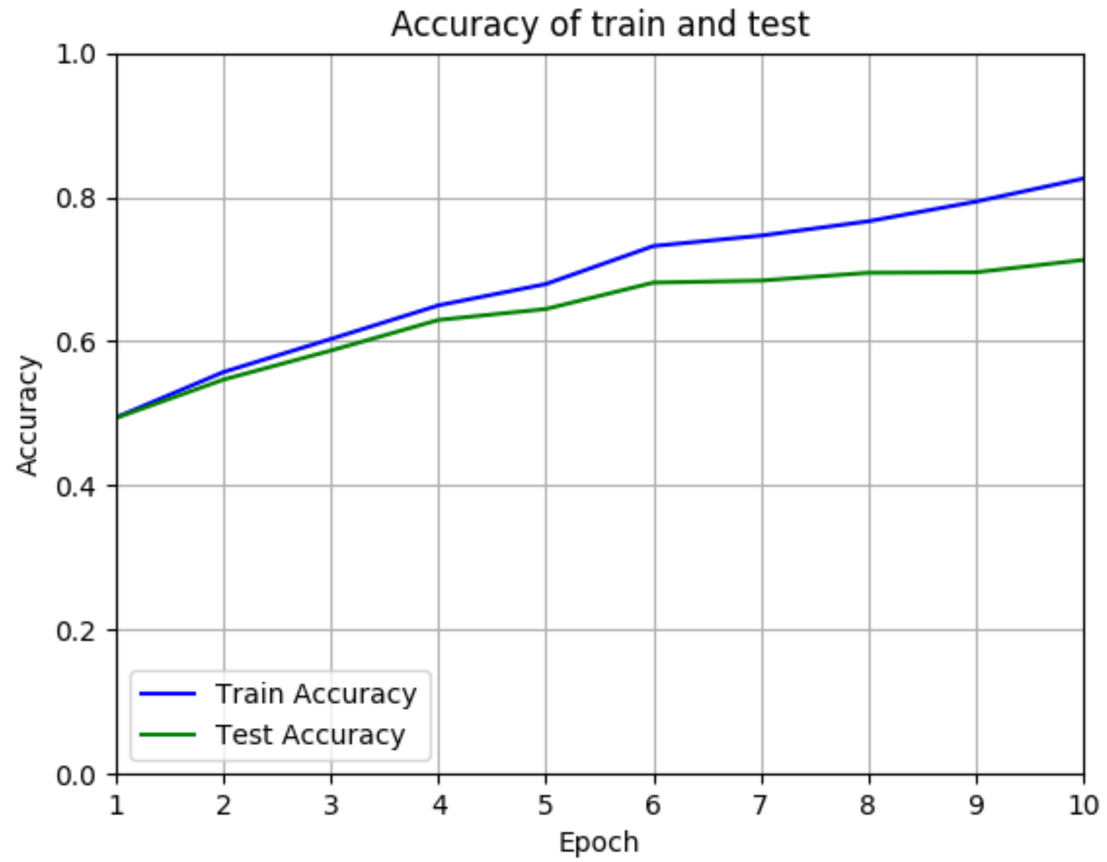


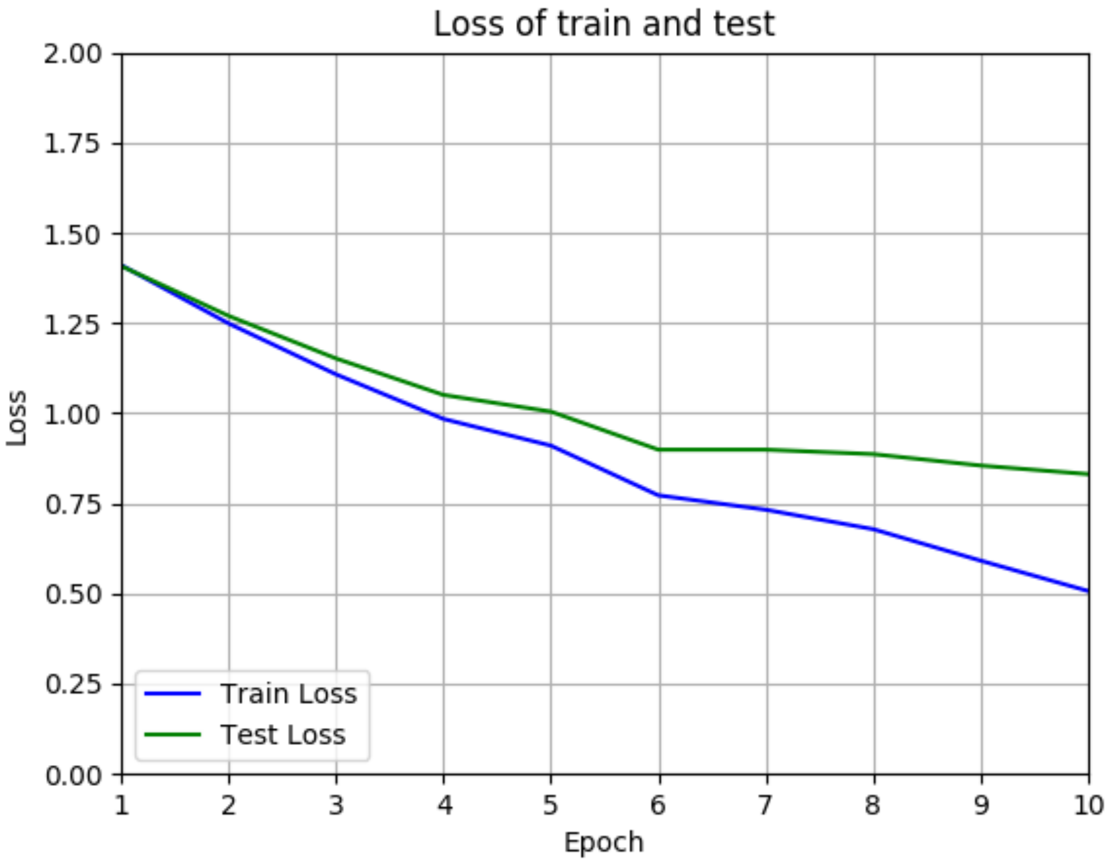
RMSprop: learning rate = 0.001. With this learning rate, I got the best results with RMSprop.



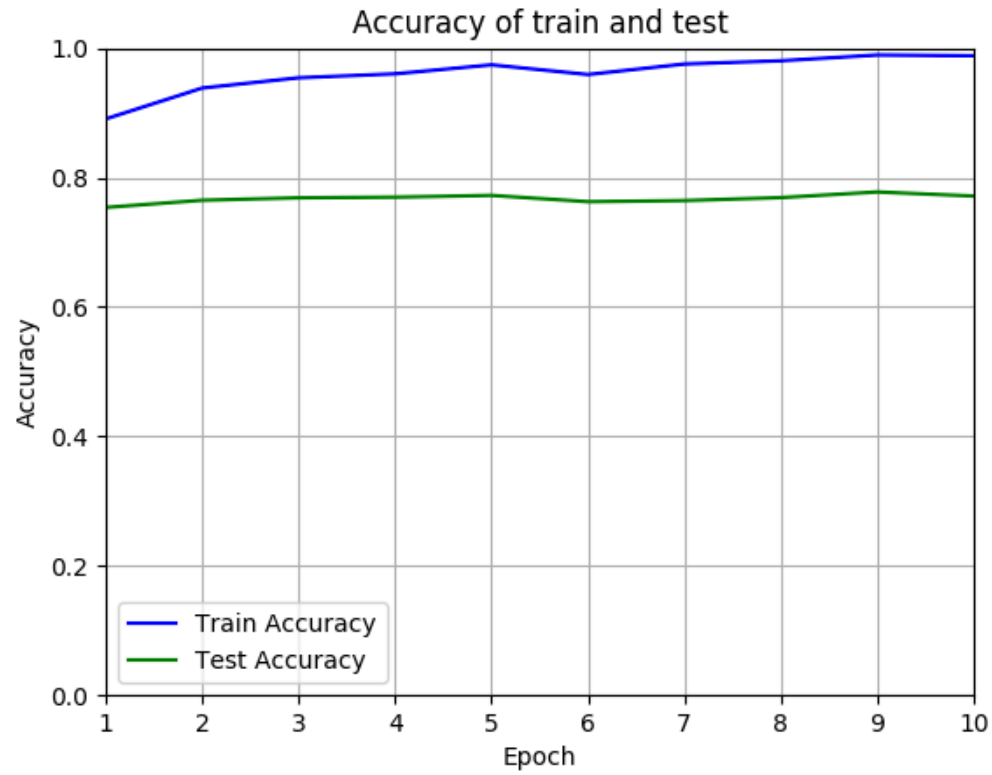


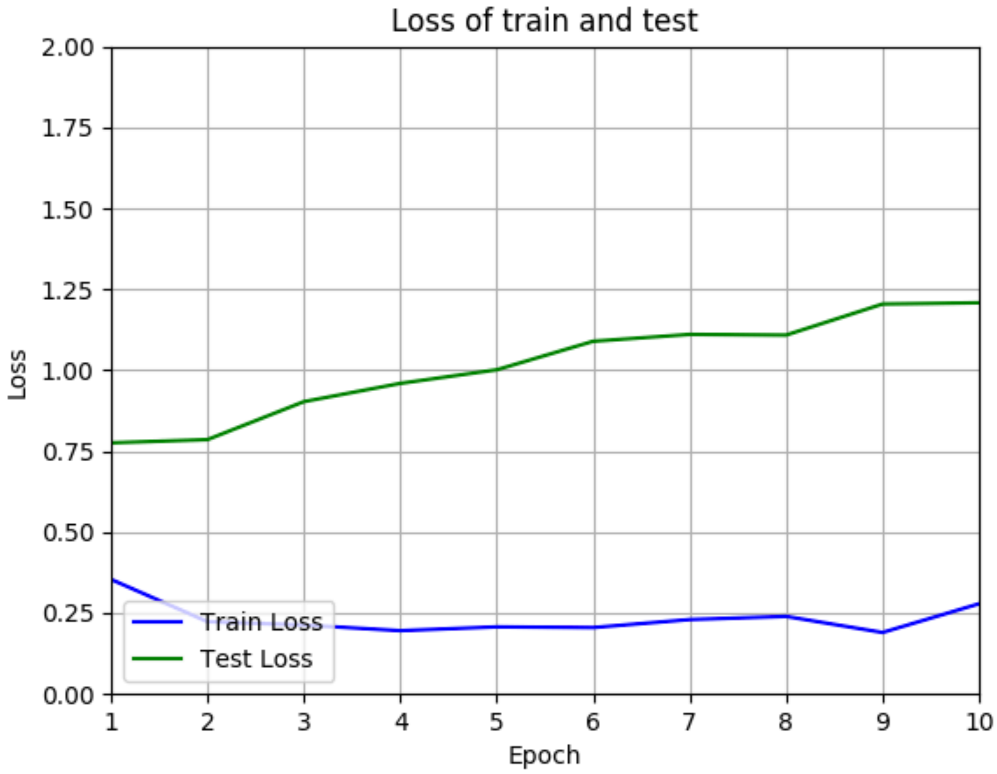
RMSprop: learning rate = 0.0001. With this learning rate, I got better results in loss of testing set comparing to other learning rates of RMSprop.



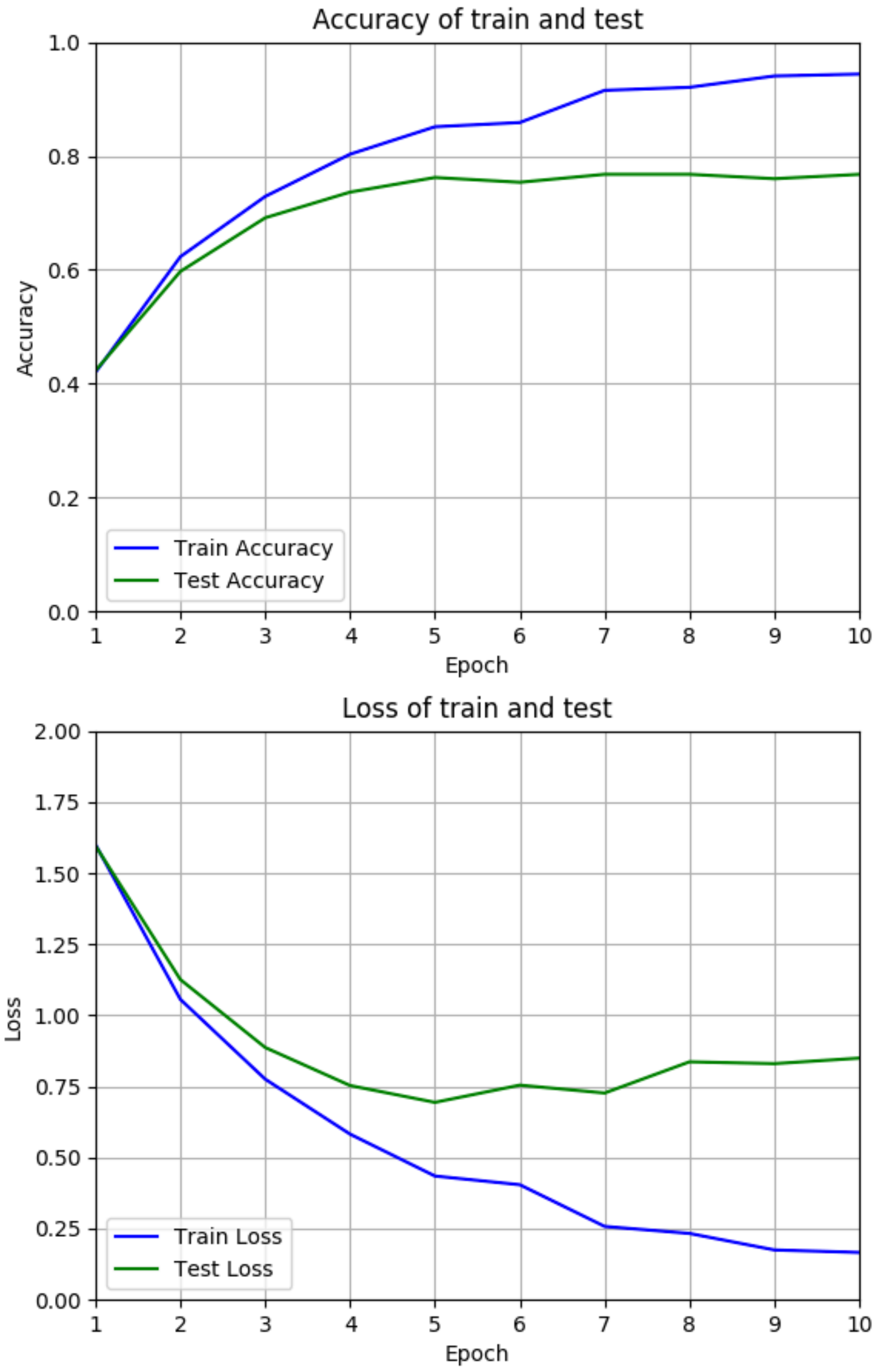


SGD:

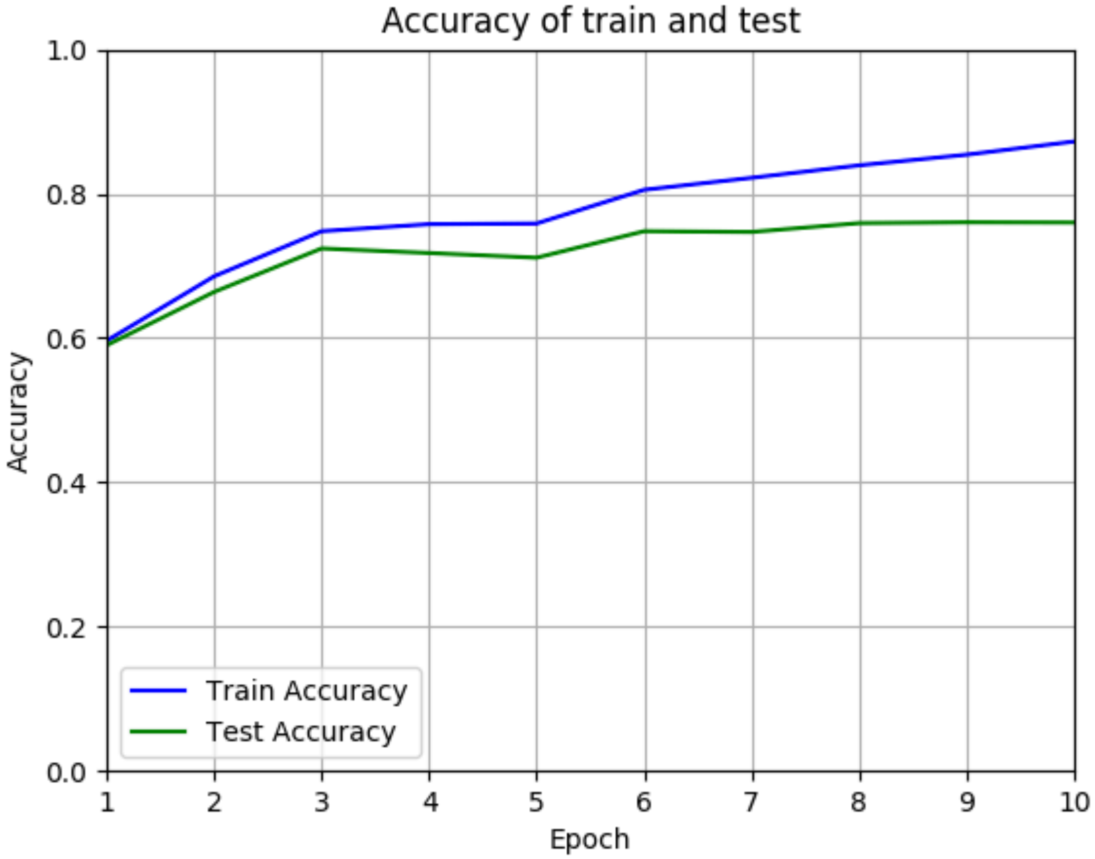

Accuracy of train and test
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9
10
Epoch
Train Accuracy
Test Accuracy

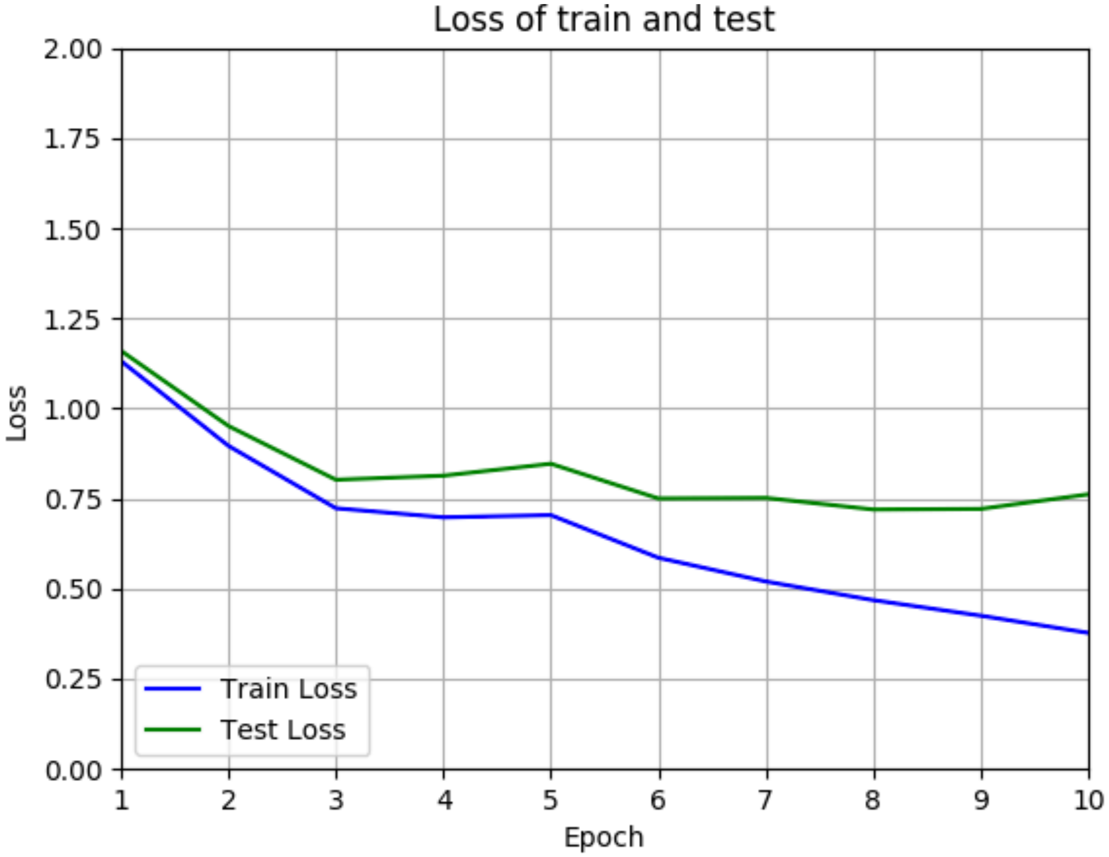

Loss of train and test
Loss
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
6
7
8
9
10
Epoch
Train Loss
Test Loss

Setting 4) Tuning network.
Adding 2 convolutional hidden layers:



Dropout:


Accuracy of train and test
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
1
2
3
4
5
6
7
8
9
10
Epoch
Train Accuracy
Test Accuracy


Loss of train and test
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Loss
1
2
3
4
5
6
7
8
9
10
Epoch
Train Loss
Test Loss