Assignment 3 CS519F16
Xi Yu

Setting (1) - No dropout layer in FCNN

	Best loss	Best error rate
Train	0.7768	0.2657
Validation	0.8817	0.2937

Setting (2) - Add a 512-unit fully connected layer on model (1)

	Best loss	Best error rate
Train	0.6163	0.2031
Validation	0.8357	0.2683

Setting (3) - Add dropout layers on both fully connected layers

	Best loss	Best error rate
Train	1.2101	0.3731
Validation	1.0442	0.3228

Setting (4) - Use Adagrad to turn the learning rate

	Best loss	Best error rate
Train	0.8736	0.2861
Validation	0.8707	0.2871

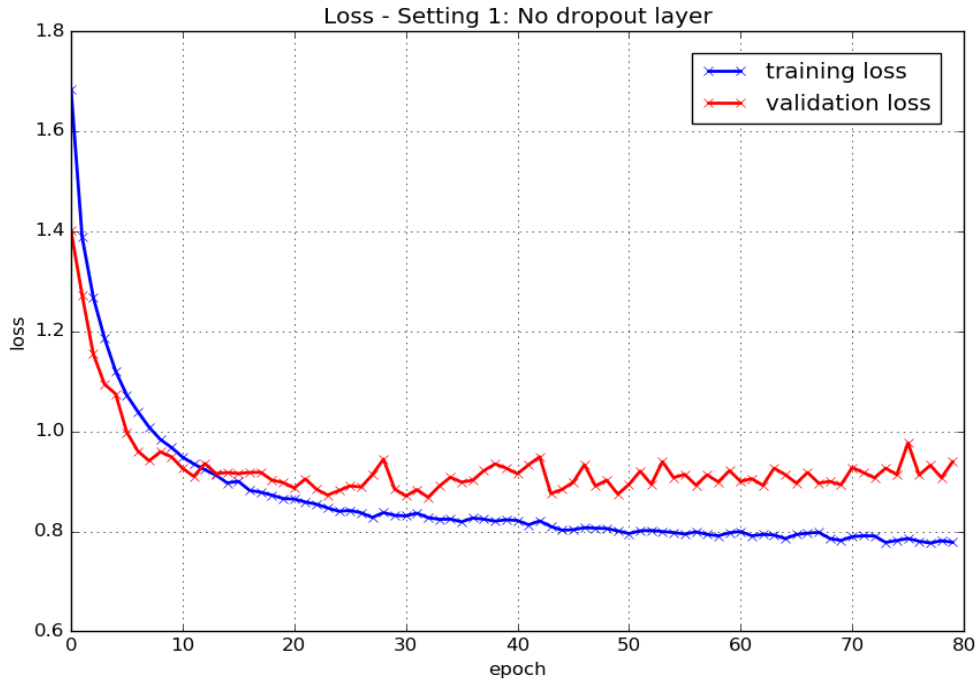
Setting (5.1) - From the original setting, reduce the number of hidden unit of FCNN to 64.

	Best loss	Best error rate
Train	0.9331	0.3238
Validation	0.8656	0.3032

Setting (5.2) - Based on the original setting, change the activation function of FCNN to Sigmoid.

	Best loss	Best error rate
Train	0.8518	0.3006
Validation	0.8247	0.2891

* All record history is stored in “record_*.txt” files.
* Plots are shown as follow.



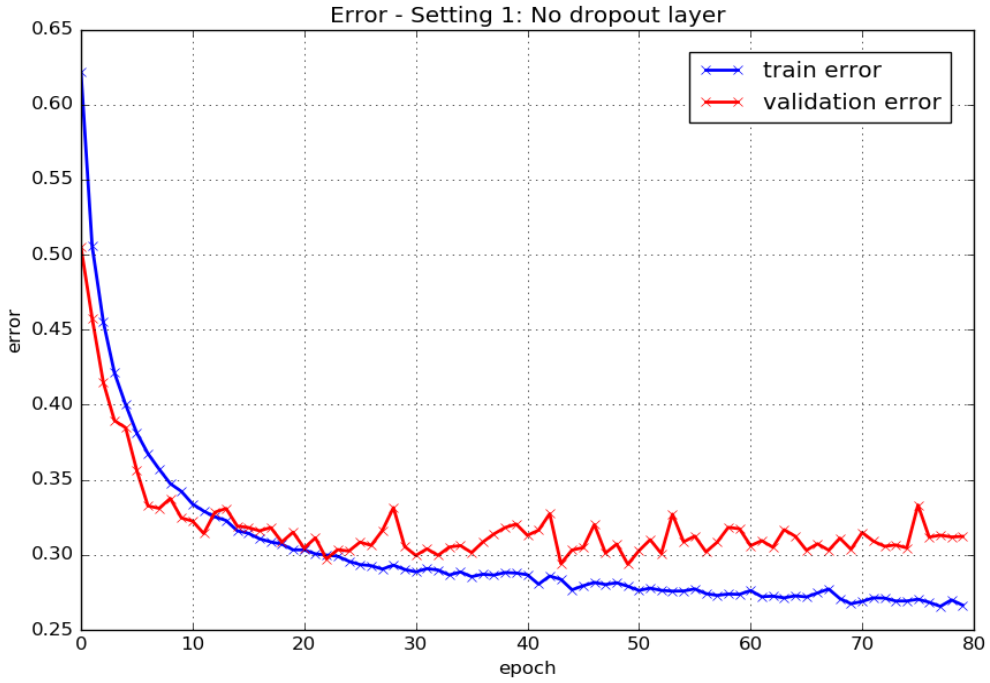
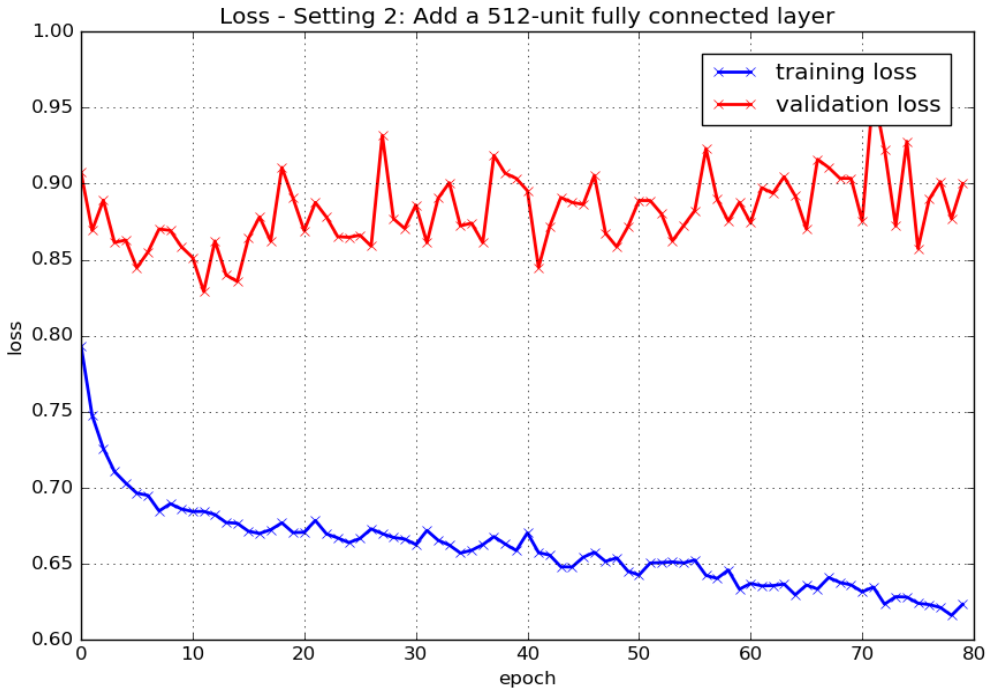


Figure 1. Setting 1 - No dropout layer in FCNN



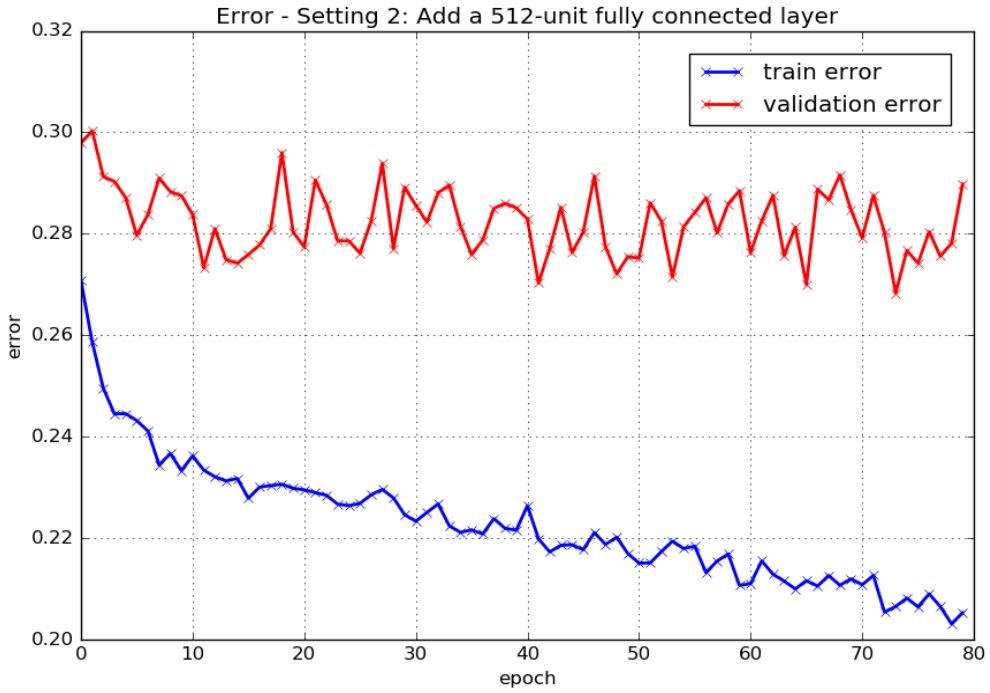
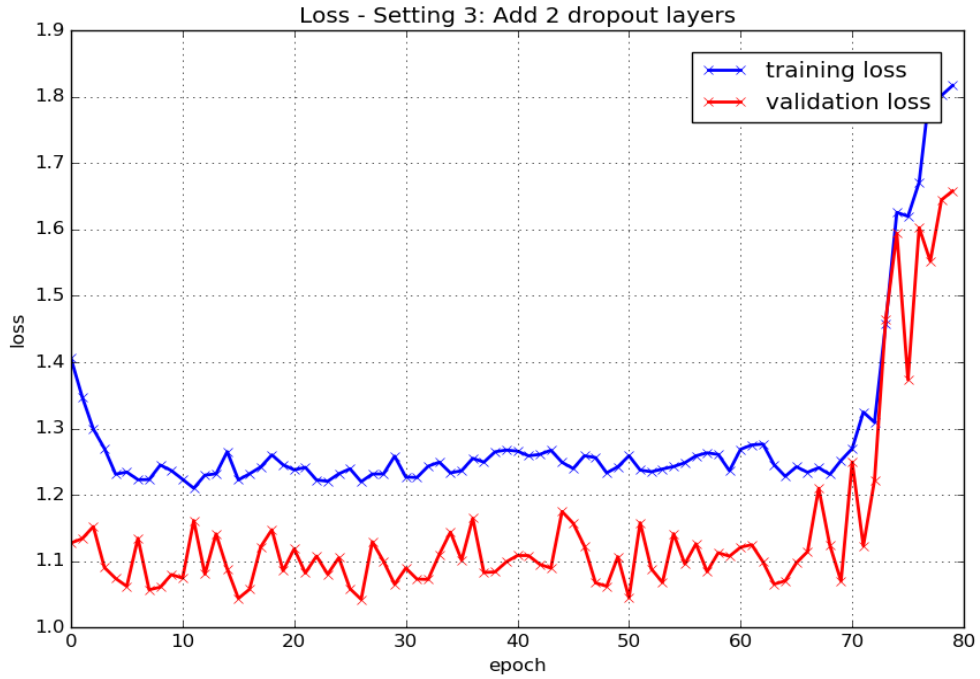


Figure 2. Add a 512-unit fully connected layer



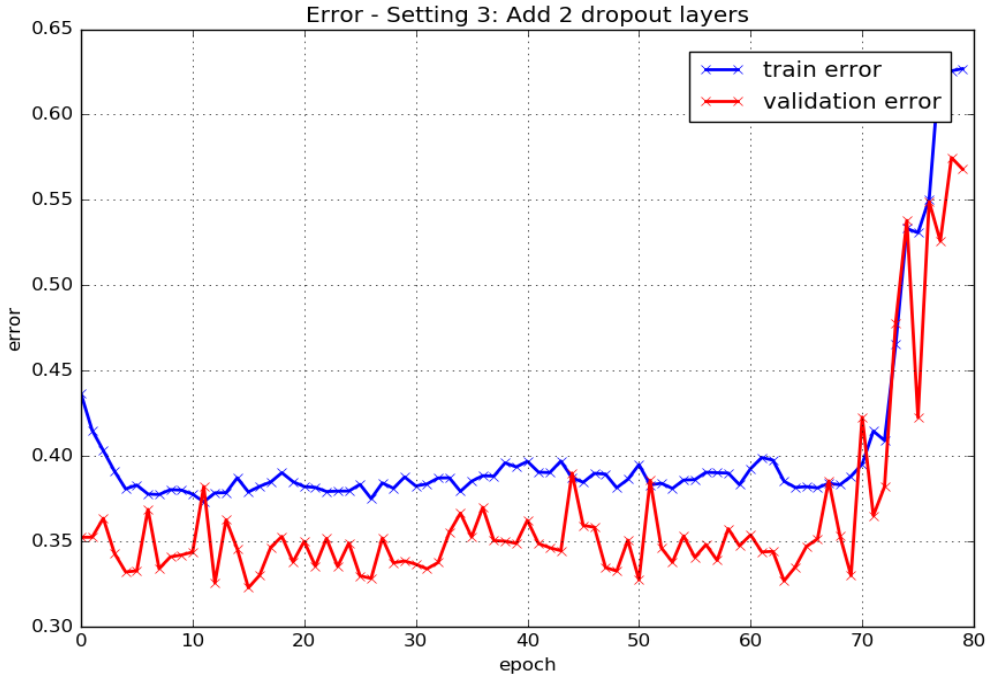
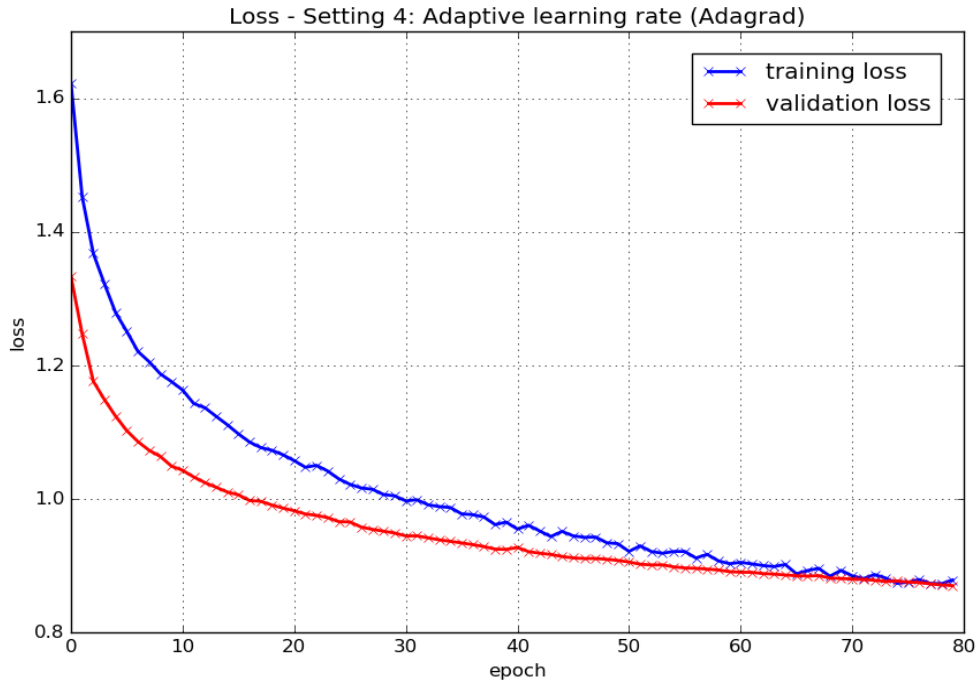


Figure 3. Add 2 dropout layers



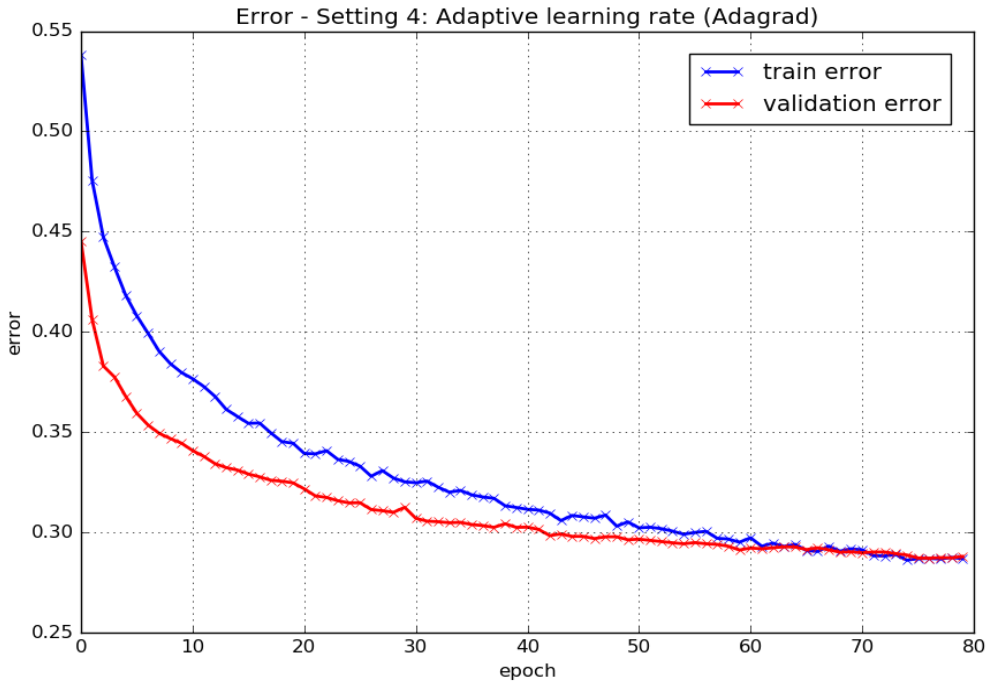
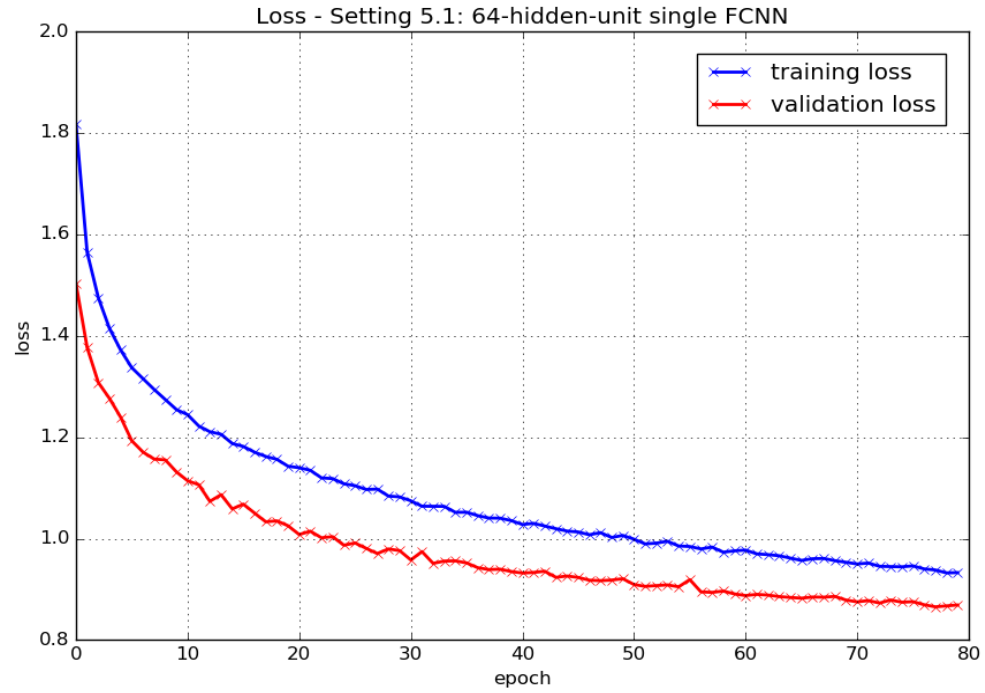


Figure 4. Adaptive learning rate (Adagrad)



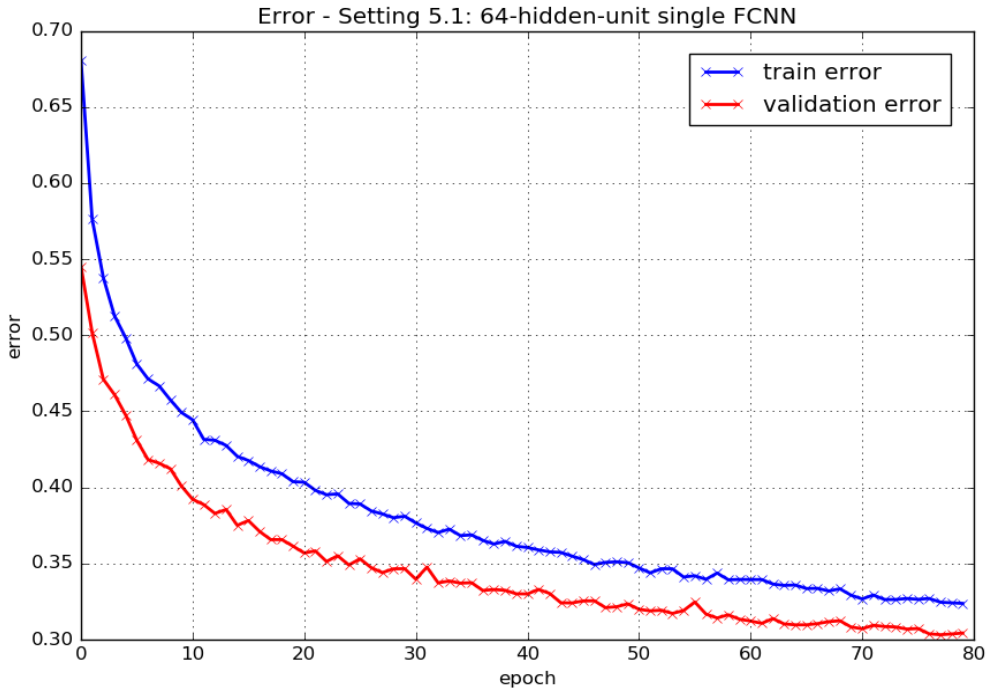
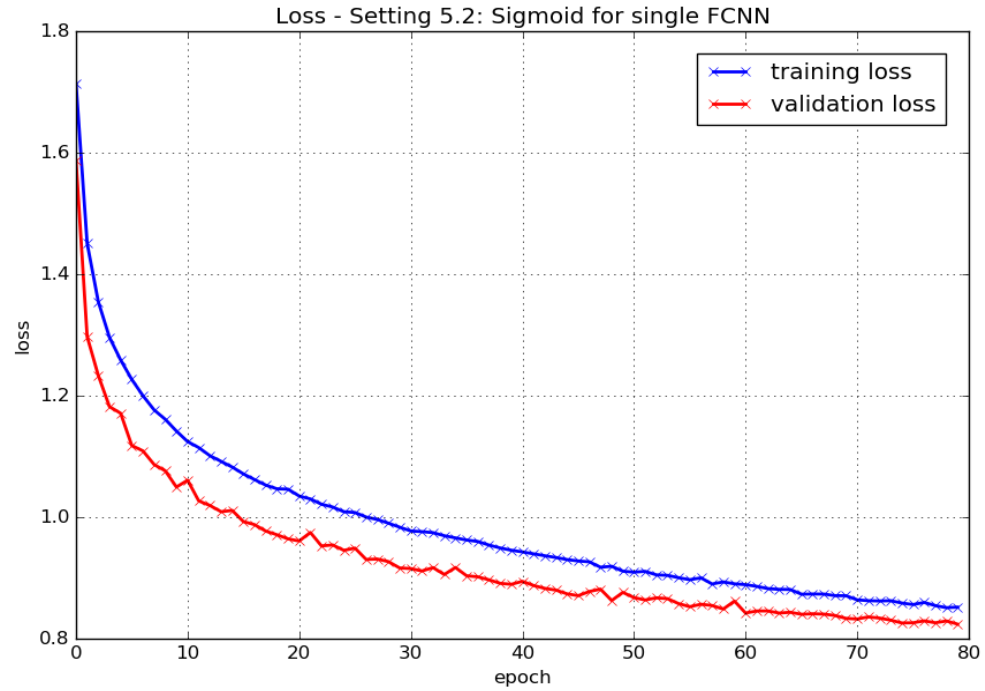


Figure 5. 64-hidden-unit in fully connected layer



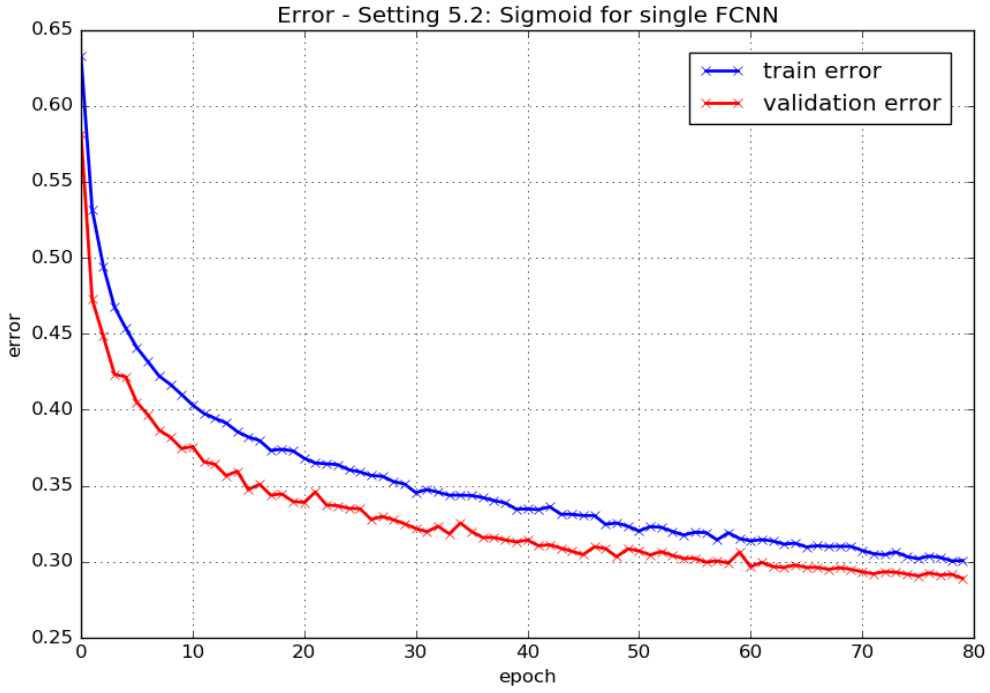


Figure 6. Sigmoid activation function for fully connected layer