Report Digit recognizer Mnist  
#1 model

Amount and size of layers: 4 layers: 1) input layer, size: (784, 256)  
 2) hidden layer, size: (256, 256)  
 3) hidden layer, size: (256, 64)  
 4) output layer, size: (64, 10)

Learning rate: 0.01375

Optimization algorithm: Adam

Loss function: softmax cross entropy

Batch size: 300

Amount of epochs: 250

Activation functions of layers: 1,2,3) Sigmoid  
 4) Softmax

Plots of training and validation accuracy and loss:  
