Report Digit recognizer Mnist  
#3 model

Amount and size of layers: 3 layers: 1) input layer, size: (784, 512)  
 2) hidden layer, size: (512, 64)  
 3) output layer, size: (64, 10)

Learning rate: 0.005

Optimization algorithm: GradientDescent

Loss function: softmax cross entropy

Batch size: 120

Amount of epochs: 100

Activation functions of layers: 1) Relu  
 2) Sigmoid  
 3) Softmax

Plots of training and validation accuracy and loss:  
