CS 547 MP2 Report

Hongpeng Guo hg5@illinois.edu

In this MP, I implemented a convolution neural network with one hidden layer. The input layer is of dimension 28 * 28 and the output layer is of dimension 10. The convolution kernel size is 4 * 4. To achieve more powerful CNN model. We set 8 channels for the kernel.

I chose the activation function to be the sigmoid function.

To achieve better results, we normalize the matrix W and K following $N(0, \frac{1}{d_H})$.

The training process contains 6 epochs, with the learning rates to be 0.1 as below:

The final accuracy of our model on the test data set is: 97.04%

The training accuracies and training time of each epochs are listed in the table and figure below:

Epoch (#)	Training time (s)	Accuracy
0	354.12	0.8689
1	365.88	0.9377
2	390.26	0.9496
3	431.13	0.9575
4	429.87	0.9625
5	410.76	0.9659