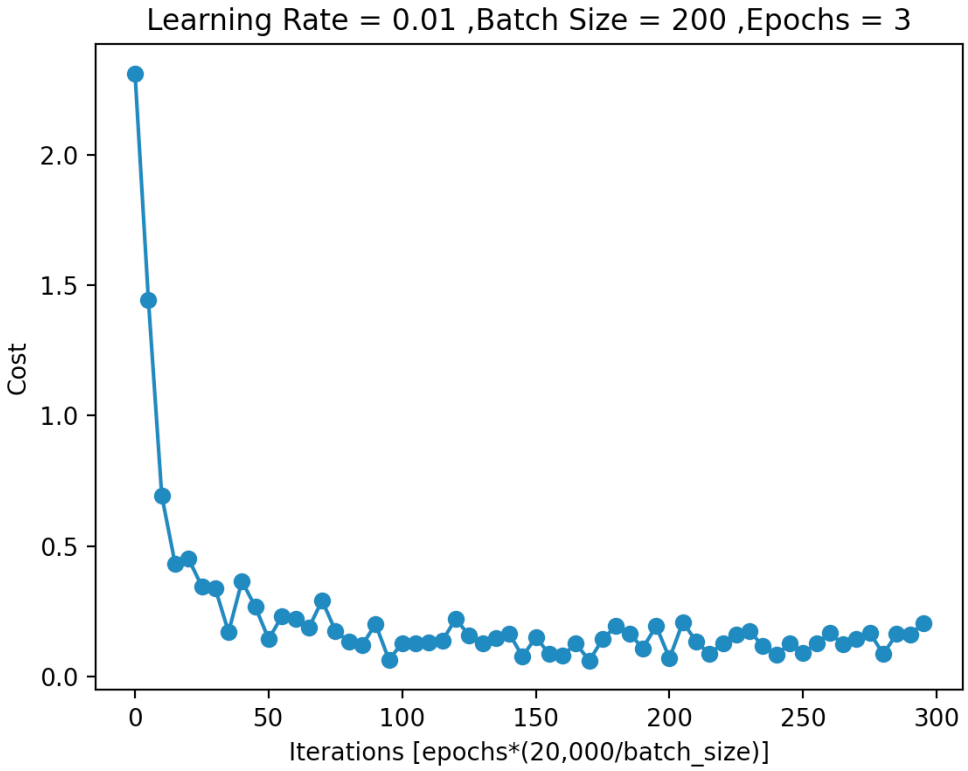
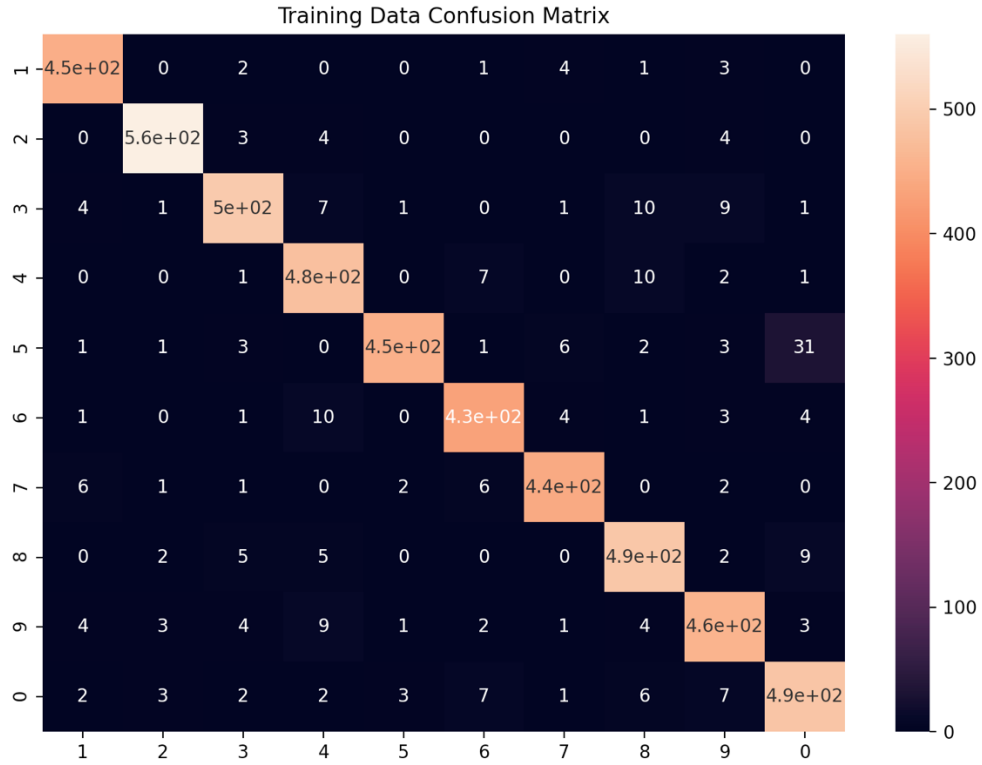
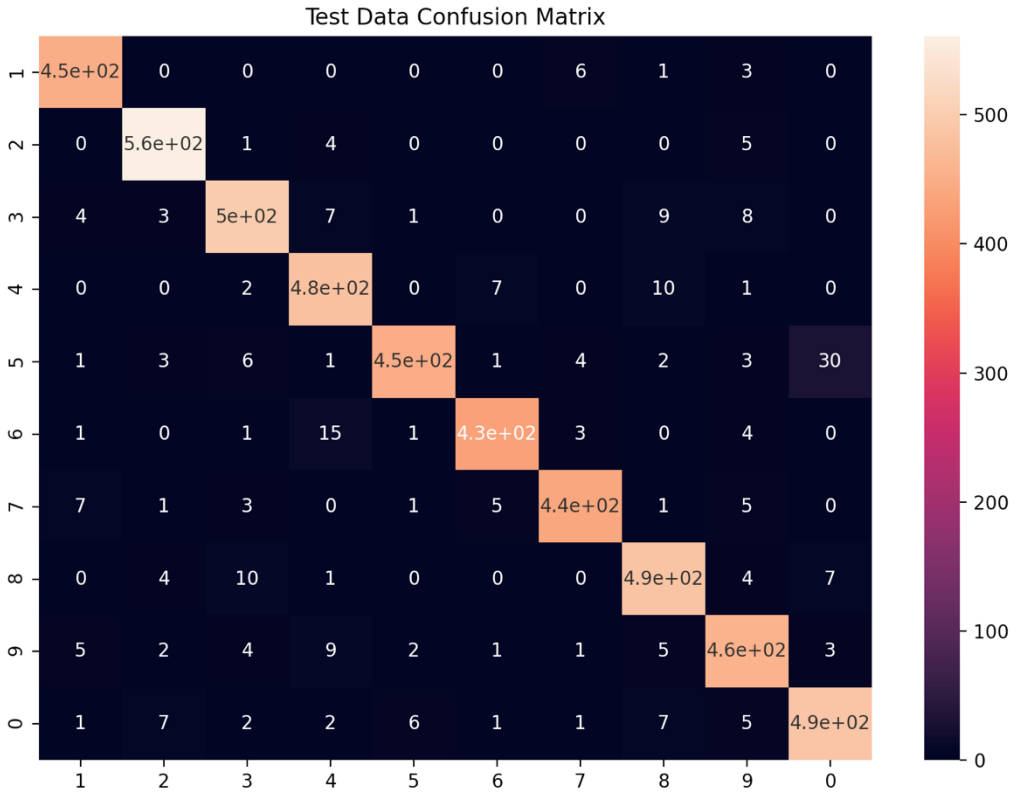
Brian Matejevich
HW 2
April 4th, 2020

Problem 1







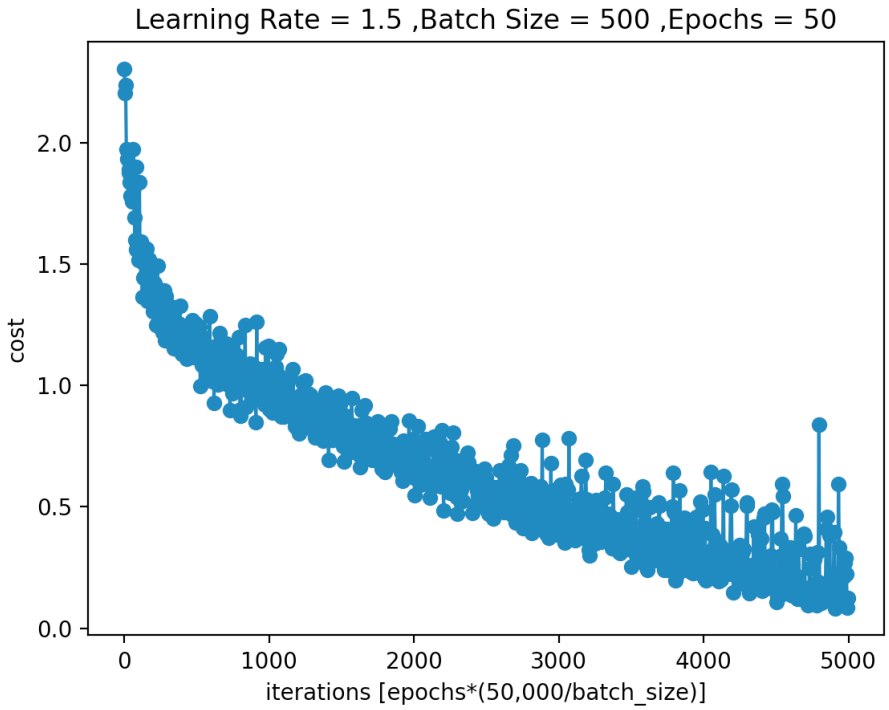
Top 1 error rate

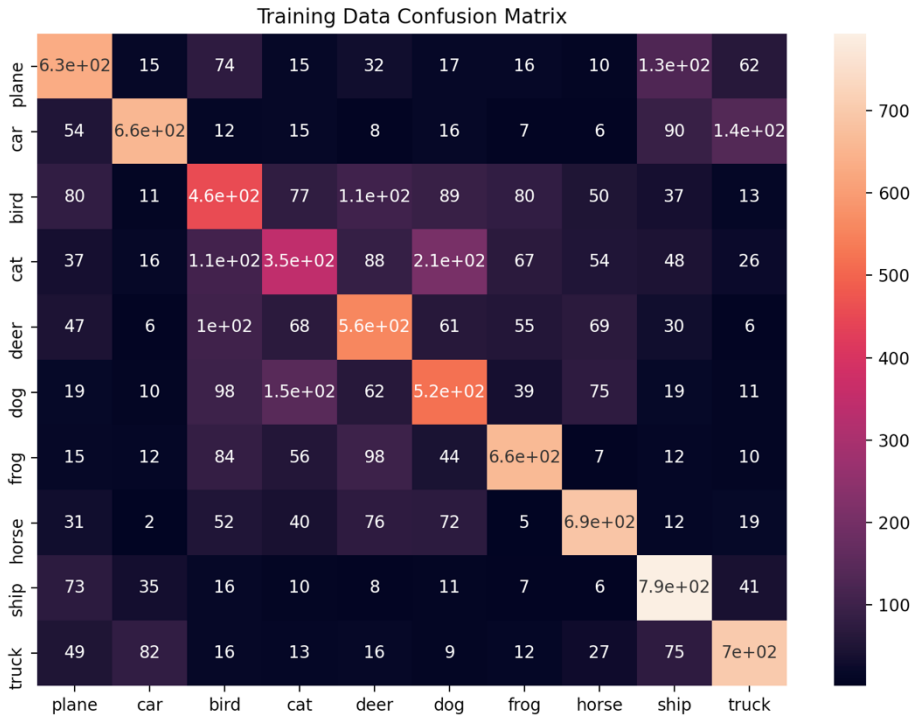
CIFAR											(%)
Average Precision	1	2	3	4	5	6	7	8	9	0	Mean
Training	0.36	0.22	0.44	0.74	0.14	0.48	0.34	0.68	0.7	0.98	0.51
Testing	0.38	0.4	0.58	0.78	0.22	0.3	0.3	0.7	0.76	0.8	0.52

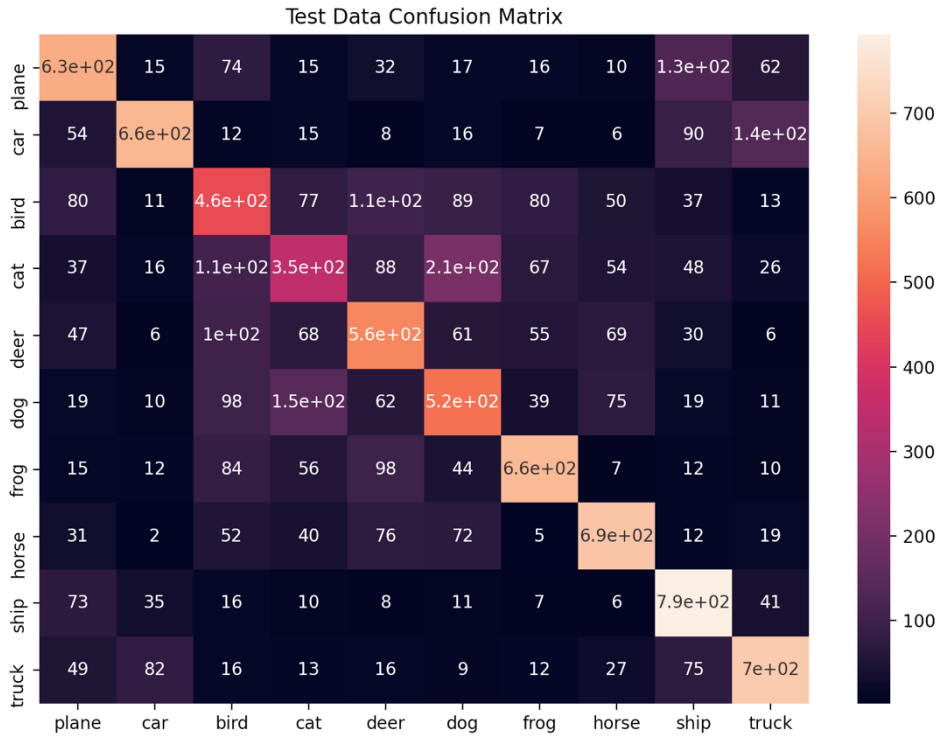
Top 3 error rate

CIFAR											(%)
Average Precision	1	2	3	4	5	6	7	8	9	0	Mean
Training	0.02	0.06	0.02	0.04	0.16	0.12	0.04	0.02	0.12	0.12	0.07
Testing	0.02	0.06	0.06	0.0	0.14	0.04	0.1	0.08	0.1	0.16	0.08

Problem 2


Learning Rate = 1.5 ,Batch Size = 500 ,Epochs = 50
cost
2.0
1.5
1.0
0.5
0.0
0
1000
2000
3000
4000
5000
iterations [epochs*(50,000/batch_size)]


Training Data Confusion Matrix
plane 6.3e+02 15 74 15 32 17 16 10 1.3e+02 62
car 54 6.6e+02 12 15 8 16 7 6 90 1.4e+02
bird 80 11 4.6e+02 77 1.1e+02 89 80 50 37 13
cat 37 16 1.1e+02 3.5e+02 88 2.1e+02 67 54 48 26
deer 47 6 1e+02 68 5.6e+02 61 55 69 30 6
dog 19 10 98 1.5e+02 62 5.2e+02 39 75 19 11
frog 15 12 84 56 98 44 6.6e+02 7 12 10
horse 31 2 52 40 76 72 5 6.9e+02 12 19
ship 73 35 16 10 8 11 7 6 7.9e+02 41
truck 49 82 16 13 16 9 12 27 75 7e+02
plane car bird cat deer dog frog horse ship truck
700
600
500
400
300
200
100



Top 1 error rate

MNIST											(%)
Average Precision	Airplane	Car	Bird	Cat	Deer	Dog	Frog	Horse	Ship	Truck	Mean
Training	0.33	0.01	0.36	0.17	0.54	0.27	0.05	0.14	0.63	0.08	0.26
Testing	4.05	1.89	5.61	4.41	4.95	5.27	2.88	3.04	4.52	3.24	3.99

Top 3 error rate

MNIST											(%)
Average Precision	Airplane	Car	Bird	Cat	Deer	Dog	Frog	Horse	Ship	Truck	Mean
Training	0.0	0.02	0.02	0.04	0.02	0.03	0.03	0.03	0.01	0.01	0.02
Testing	0.89	0.9	2.0	2.17	1.25	1.78	1.21	1.4	0.59	0.98	1.32

● What you learned through this assignment.

The combination of choice of batch size and epoch # has a lot to do with training time

● Implementation of your convolutional neural network architecture.

Problem 1 had an architecture using 2 convolution layers, three relu layers, 2 sets of dropout and finally a softmax for the output layer

Problem 2 used an architecture of a convolution, 2 relu layers, one pooling layer and 2 fully connected layers

● Confusion matrix – See above

● Analyze and report hyperparameters obtained. (optimizer, batch size, epochs, layers etc)

Problem 1- adam optimizer, batch size 200, 3 epochs, learning rate of 0.01, used 20,000 random samples from the training set

Problem 2 – adadelta optimizer, batch size of 500, 50 epochs, learning rate of 1.5 and 50,000 random training samples

● Training, validation, test curves and classification result. – See above

● Best hyperparameters and train-test split obtained and your interpretation of these results.

Problem 1- adam optimizer, batch size 200, 3 epochs, learning rate of 0.01, used 20,000 random samples from the training set. Looking at the loss curve the plot levels of and more epochs will have little affect. The accuracy is already really good so we don't want to overtrain.

Problem 2 – adadelta optimizer, batch size of 500, 50 epochs, learning rate of 1.5 and 50,000 random training samples. With 50 epochs the training takes about 25 minutes but it is very accurate. These images have more details and features so it makes sense it takes longer to train effectively.