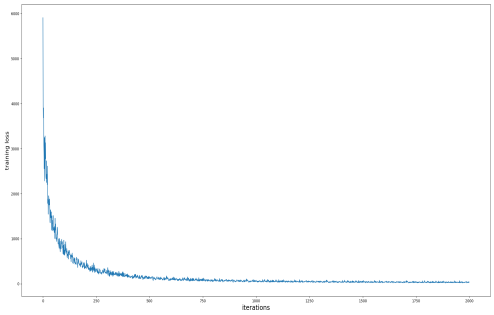
1 SLP Linear



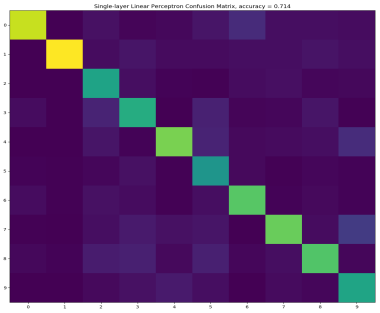
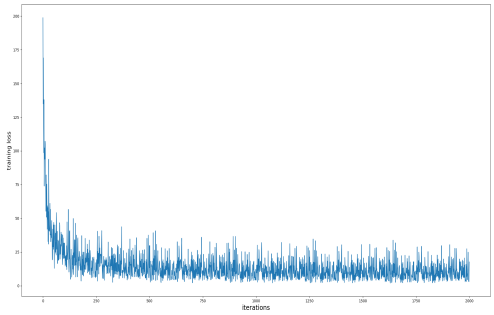


Figure 1: SLP Linear training loss and confusion matrix

Figure 1 shows the training loss and confusion matrix of SLP linear model. Intial weights are taken from a normal distributio with mean 0 and standard deviation 1. Parameters used in this are LEARNING_RATE = 0.001, DECAY_RATE = 0.9, NUM_ITERATIONS = 2000, BATCH_SIZE = 32 It achieves about 71% accuracy.

2 SLP



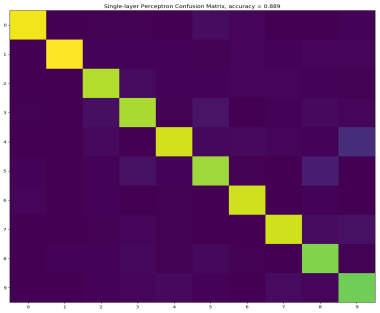
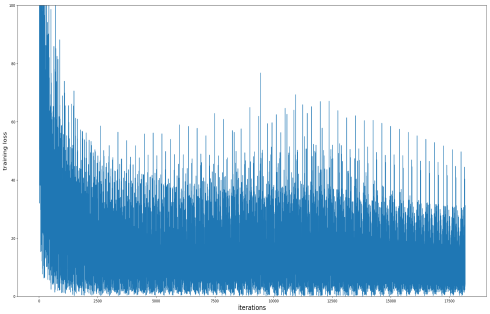


Figure 2: SLP training loss and confusion matrix

Figure 2 shows the training loss and confusion matrix of SLP model. Intial weights are taken from a normal distributio with mean 0 and standard deviation 1. Parameters used in this are LEARNING_RATE = 0.04, DECAY_RATE = 0.9, NUM_ITERATIONS = 2000, BATCH_SIZE = 32 It achieves about 88.9% accuracy.

3 MLP

Figure 3 shows the training loss and confusion matrix of MLP model. The pipeline architecture is as described in the problem statement. Intial weights are taken from a normal



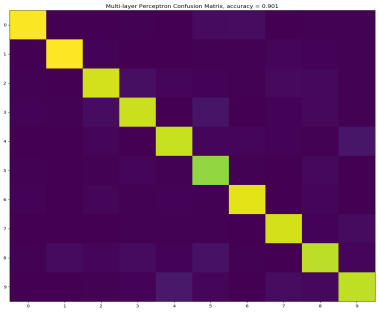
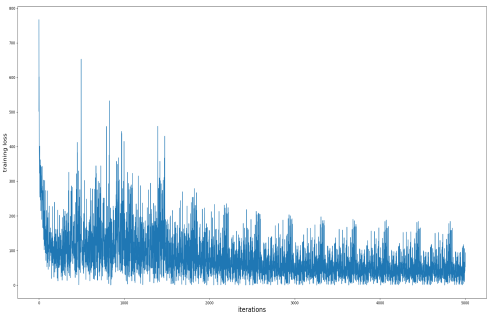


Figure 3: MLP training loss and confusion matrix

distributio with mean 0 and standard deviation 1. Parameters used in this are LEARNING_RATE = 0.02, DECAY_RATE = 0.9, NUM_ITERATIONS = 18170, BATCH_SIZE = 32. A leaky relu with epsilon = 0.9 is used instead of relu. And learning rate decay is done per 500 iterations rather than 1000 iterations. It achieves about 90.1% accuracy.

4 CNN



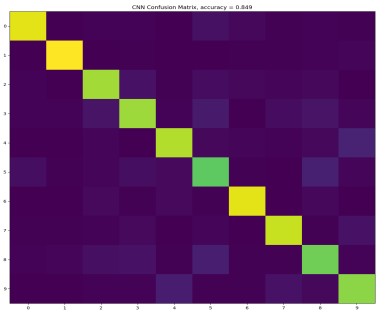


Figure 4: CNN training loss and confusion matrix

Figure 4 shows the training loss and confusion matrix of CNN model. The pipeline architecture is as described in the problem statement. Intial weights are taken from a normal distributio with mean 0 and standard deviation 1. Parameters used in this are LEARNING_RATE = 0.002, DECAY_RATE = 0.6, NUM_ITERATIONS = 7500, BATCH_SIZE = 32. A leaky relu with epsilon = 0.5 is used instead of relu. And learning rate decay is done per 500 iterations rather than 1000 iterations. It achieves about 84.9% accuracy.