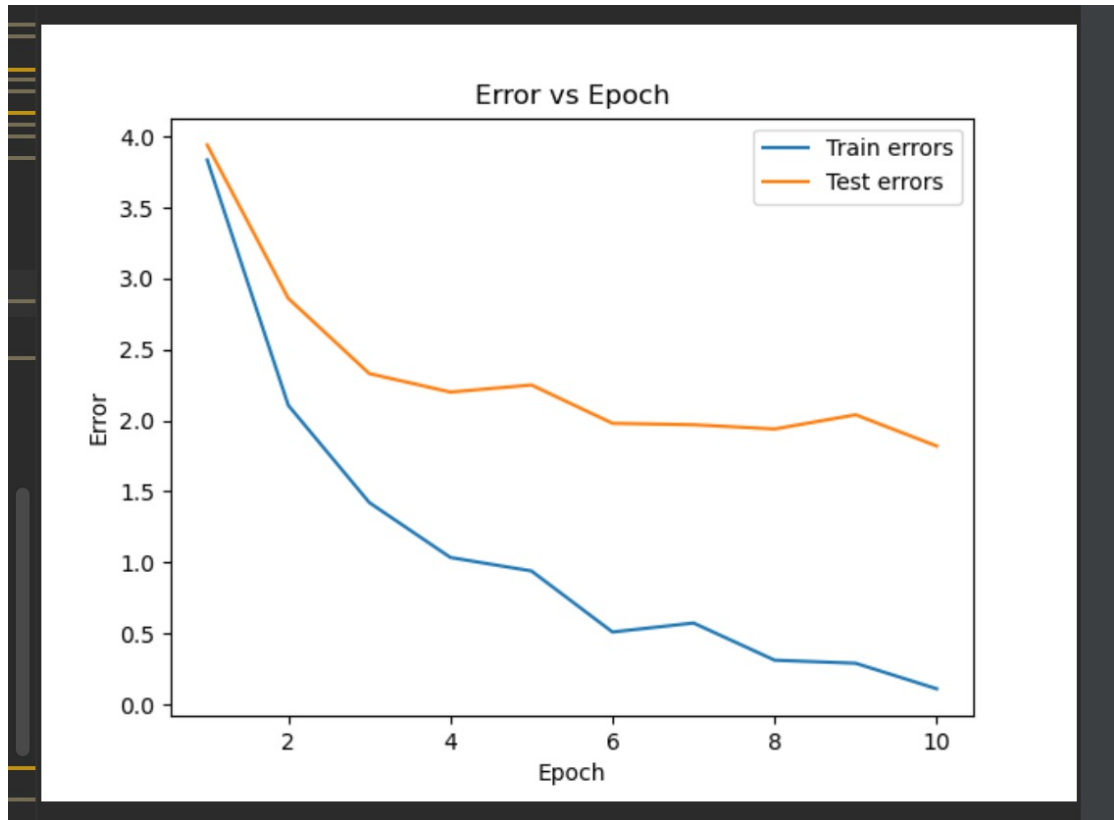


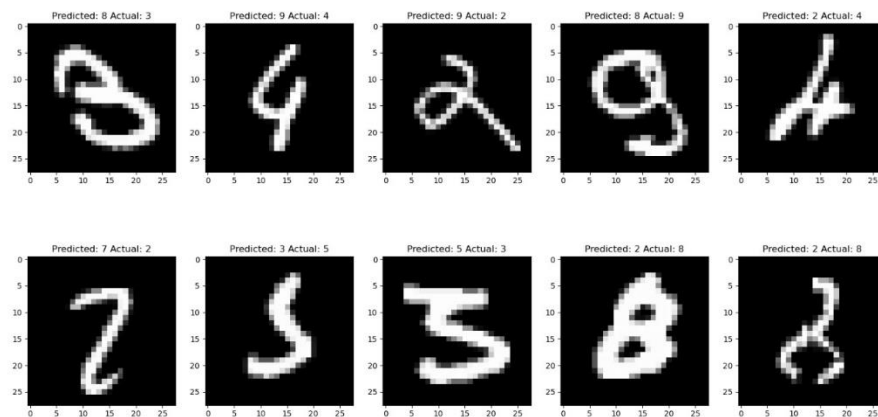
Assignment 3 – Applied Deep Learning

1.

Plot of Error vs Epoch

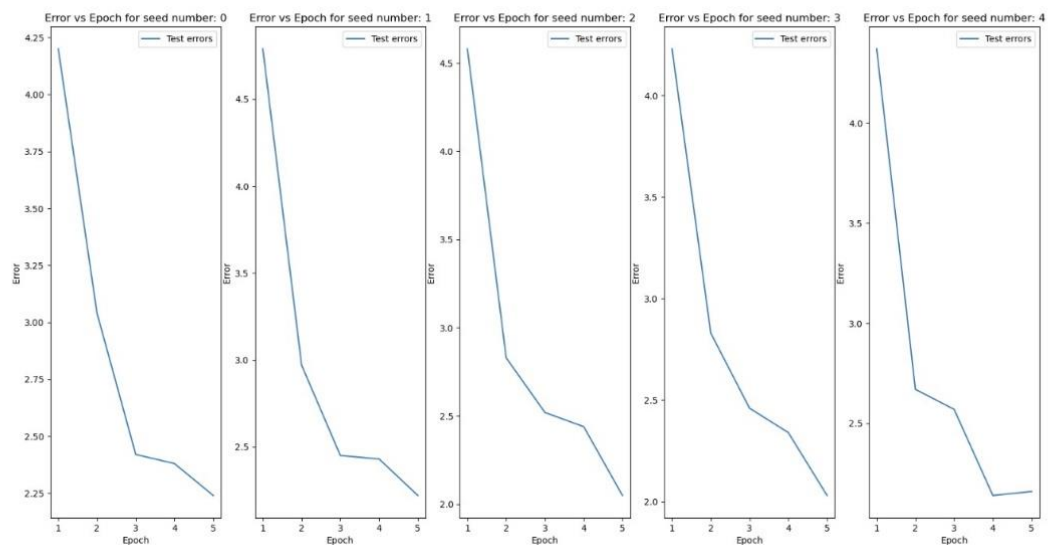


plot of 10 of the misclassified images:



2.

plot of the test errors during training from the separate runs:



the mean of the final test errors is: 2.14

the standard deviation of the final test errors is: 0.08602325267042644

The model is robust to the choice of the seed because the standard deviation of the final test error is small and the model's performance remains relatively stable across different seed numbers.

3.

the best validation error achieved for seed 0 is: 2.23
the test error achieved is: 2.1

the best validation error achieved for seed 1 is: 2.19
the test error achieved is: 2.13

the best validation error achieved for seed 2 is: 2.26
the test error achieved is: 2.23

the best validation error achieved for seed 3 is: 2.37
the test error achieved is: 2.25

the best validation error achieved for seed 4 is: 2.16
the test error achieved is: 2.01

4.

Batch size	Learning rate	Hidden size	Best validation error achieved	Test error achieved
100	0.001	50	4.74	4.57
100	0.001	500	2.27	2.32
100	0.01	50	3.87	3.86
100	0.01	500	3.19	3.24
100	0.1	50	20.97	20.97
100	0.1	500	11.59	11.43
1000	0.001	50	7.88	7.79
1000	0.001	500	4.57	4.52
1000	0.01	50	4.05	3.87
1000	0.01	500	2.46	2.54
1000	0.1	50	7.71	7.94
1000	0.1	500	5.81	5.49

As we can see the best results achieved by batch size = 100, learning rate = 0.001, hidden size = 500.

```
for batch size = 100, learning rate = 0.001 and hidden size = 50 the best validation error achieved is: 4.74
the test error achieved is: 4.57

for batch size = 100, learning rate = 0.001 and hidden size = 500 the best validation error achieved is: 2.27
the test error achieved is: 2.32

for batch size = 100, learning rate = 0.01 and hidden size = 50 the best validation error achieved is: 3.87
the test error achieved is: 3.86

for batch size = 100, learning rate = 0.01 and hidden size = 500 the best validation error achieved is: 3.19
the test error achieved is: 3.24

for batch size = 100, learning rate = 0.1 and hidden size = 50 the best validation error achieved is: 20.97
the test error achieved is: 20.97

for batch size = 100, learning rate = 0.1 and hidden size = 500 the best validation error achieved is: 11.59
the test error achieved is: 11.43

for batch size = 1000, learning rate = 0.001 and hidden size = 50 the best validation error achieved is: 7.88
the test error achieved is: 7.79

for batch size = 1000, learning rate = 0.001 and hidden size = 500 the best validation error achieved is: 4.57
the test error achieved is: 4.52

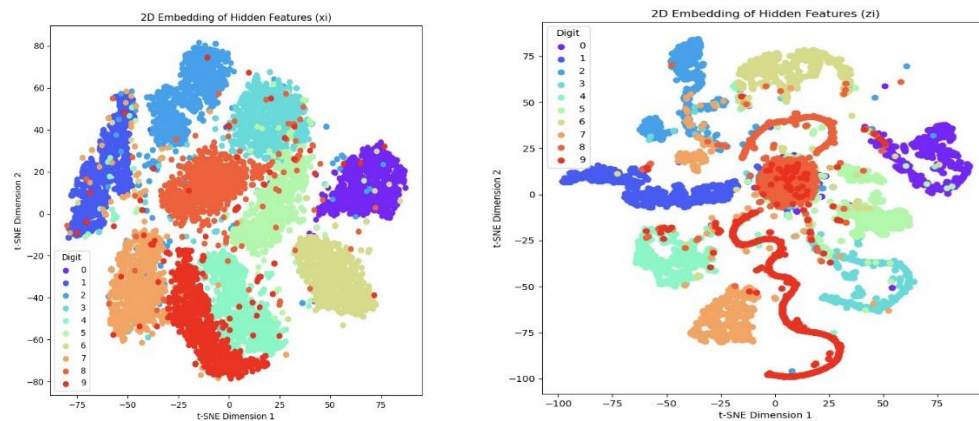
for batch size = 1000, learning rate = 0.01 and hidden size = 50 the best validation error achieved is: 4.05
the test error achieved is: 3.87

for batch size = 1000, learning rate = 0.01 and hidden size = 500 the best validation error achieved is: 2.46
the test error achieved is: 2.54

for batch size = 1000, learning rate = 0.1 and hidden size = 50 the best validation error achieved is: 7.71
the test error achieved is: 7.94

for batch size = 1000, learning rate = 0.1 and hidden size = 500 the best validation error achieved is: 5.81
the test error achieved is: 5.49
```

5.



As we can see after training the model and applying the first layer, the clusters that represent the labels become more separable.

The learned model makes it easier to distinguish between the label's clusters, and easier to decide the right predictions given the new space (which in this case seen by the plot).