Performance Report

Ruhul Azgor

January 3, 2024

1 Learning Curve of Different Model

For all the models, the given parameter was set.

- Mini-batch size: 1024
- Learning Rate decay: 0.5
- Beta1: 0.9
- Beta2: 0.99
- Epsilon: 1e-8
- Xavier Initialization

1.1 Model 1

1.1.1 Architecture

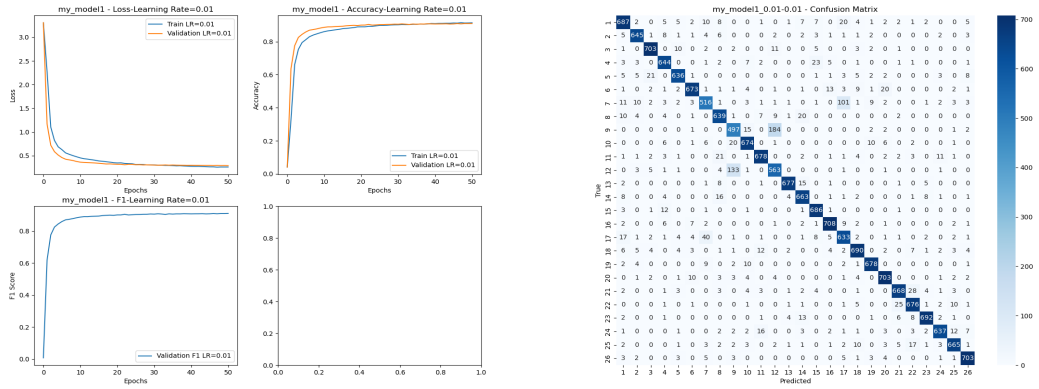
```
Dense(784, 512),
Relu(),
Dropout(probability=.15),
Dense(512, 256),
Relu(),
Dropout(probability=.15),
Dense(256, 128),
Relu(),
Dropout(probability=.10),
Dense(128, 64),
Relu(),
Dropout(probability=.10),
Dense(64, 26),
Softmax()
```

1.1.2 Learning Rate = 0.01

Best Performance:

- Best epoch: 50

- train loss: 0.2602
- validation loss: 0.2892
- train accuracy: 0.9142
- validation accuracy: 0.9099
- validation f1 score: 0.9096



(a) Model 1 Learning Curve with Learning Rate 0.01 (b) Model 1 Confusion Matrix with Learning Rate 0.01

Figure 1: Model 1 Learning Curve and Confusion Matrix with Learning Rate 0.01

1.1.3 Learning Rate = 0.005

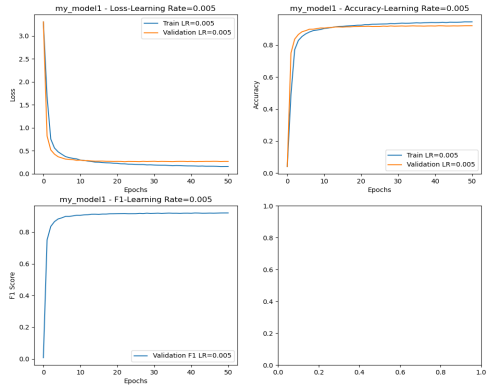
Best Performance:

- Best epoch: 50
- train loss: 0.1554
- validation loss: 0.2662
- train accuracy: 0.9448
- validation accuracy: 0.9206
- validation f1 score: 0.9204

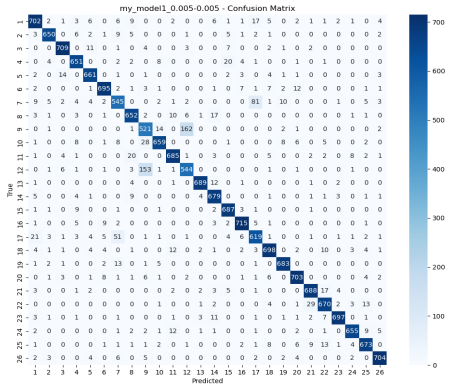
1.1.4 Learning Rate = 0.001

Best Performance:

- Best epoch: 49



(a) Model 1 Learning Curve with Learning Rate 0.005 0.005



(b) Model 1 Confusion Matrix with Learning Rate

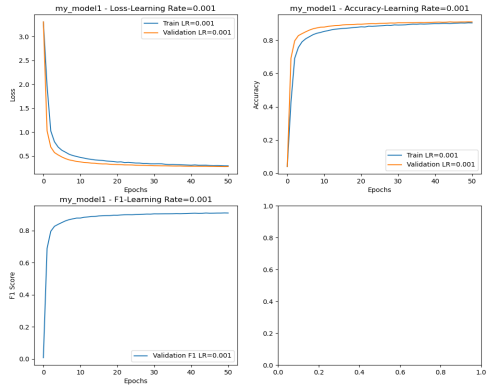
Figure 2: Model 1 Learning Curve and Confusion Matrix with Learning Rate 0.005

- train loss: 0.2959
- validation loss: 0.2793
- train accuracy: 0.9046
- validation accuracy: 0.9105
- validation f1 score: 0.9102

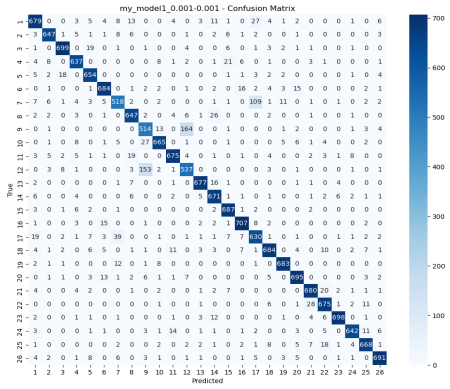
1.1.5 Learning Rate = 0.0005

Best Performance:

- Best epoch: 47
- train loss: 0.4765
- validation loss: 0.3790
- train accuracy: 0.8521
- validation accuracy: 0.8822
- validation f1 score: 0.8818



(a) Model 1 Learning Curve with Learning Rate 0.001 0.001



(b) Model 1 Confusion Matrix with Learning Rate

Figure 3: Model 1 Learning Curve and Confusion Matrix with Learning Rate 0.001

1.2 Model 2

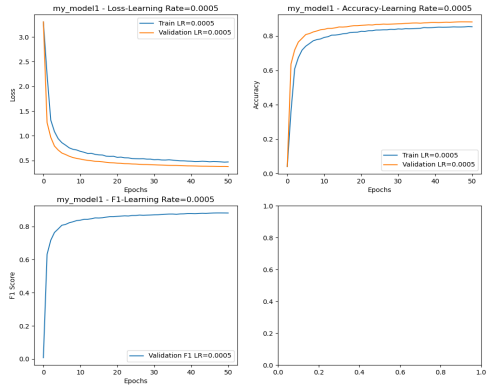
1.2.1 Architecture

```
Dense(784, 2048),
Relu(),
Dropout(probability=.4),
Dense(2048, 1024),
Relu(),
Dropout(probability=.3),
Dense(1024, 26),
Softmax()
```

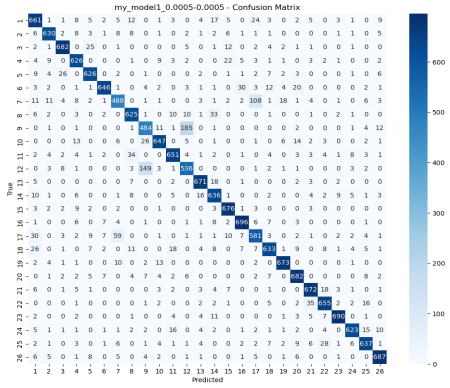
1.2.2 Learning Rate = 0.01

Best Performance:

- Best epoch: 47
- train loss: 0.3481
- validation loss: 0.2996
- train accuracy: 0.8846
- validation accuracy: 0.9074
- validation f1 score: 0.9073



(a) Model 1 Learning Curve with Learning Rate 0.0005 0.0005



(b) Model 1 Confusion Matrix with Learning Rate

Figure 4: Model 1 Learning Curve and Confusion Matrix with Learning Rate 0.0005

1.2.3 Learning Rate = 0.005

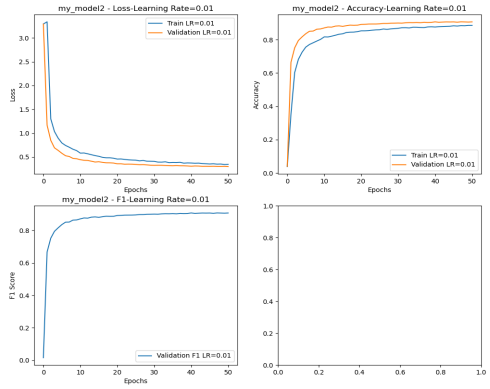
Best Performance:

- Best epoch: 49
- train loss: 0.1622
- validation loss: 0.2358
- train accuracy: 0.9413
- validation accuracy: 0.9239
- validation f1 score: 0.9237

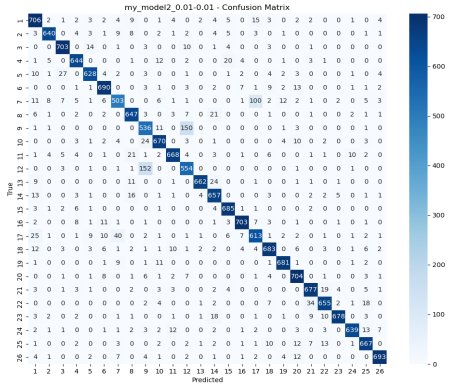
1.2.4 Learning Rate = 0.001

Best Performance:

- Best epoch: 49
- train loss: 0.1542
- validation loss: 0.2268
- train accuracy: 0.9467
- validation accuracy: 0.9248
- validation f1 score: 0.9246



(a) Model 2 Learning Curve with Learning Rate 0.01



(b) Model 2 Confusion Matrix with Learning Rate 0.01

Figure 5: Model 2 Learning Curve and Confusion Matrix with Learning Rate 0.01

1.2.5 Learning Rate = 0.0005

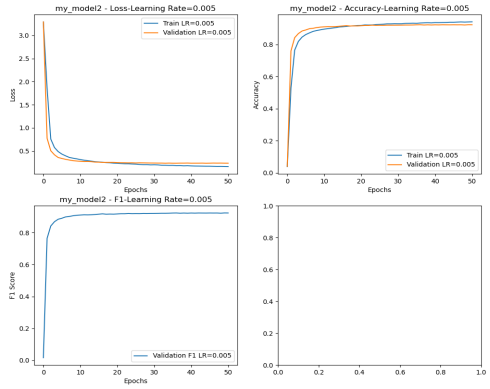
Best Performance:

- Best epoch: 50
- train loss: 0.2369
- validation loss: 0.2604
- train accuracy: 0.9225
- validation accuracy: 0.9159
- validation f1 score: 0.9157

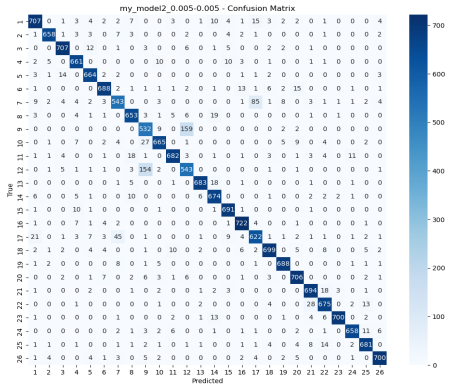
1.3 Model 3

1.3.1 Architecture

```
Dense(784, 1024),
Relu(),
Dropout(probability=.2),
Dense(1024, 512),
Relu(),
Dropout(probability=.2),
Dense(512,26),
Softmax()
```



(a) Model 2 Learning Curve with Learning Rate 0.005



(b) Model 2 Confusion Matrix with Learning Rate 0.005

Figure 6: Model 2 Learning Curve and Confusion Matrix with Learning Rate 0.005

1.3.2 Learning Rate = 0.01

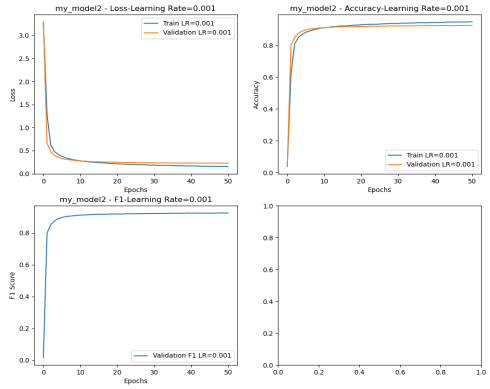
Best Performance:

- Best epoch: 49
- train loss: 0.1794
- validation loss: 0.2574
- train accuracy: 0.9346
- validation accuracy: 0.9161
- validation f1 score: 0.9156

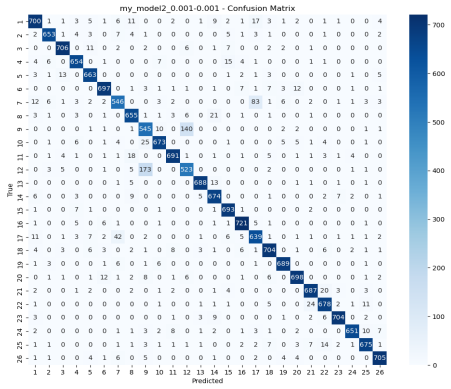
1.3.3 Learning Rate = 0.005

Best Performance:

- Best epoch: 35
- train loss: 0.1056
- validation loss: 0.2440
- train accuracy: 0.9599
- validation accuracy: 0.9260
- validation f1 score: 0.9257



(a) Model 2 Learning Curve with Learning Rate 0.001 0.001



(b) Model 2 Confusion Matrix with Learning Rate

Figure 7: Model 2 Learning Curve and Confusion Matrix with Learning Rate 0.001

1.3.4 Learning Rate = 0.001

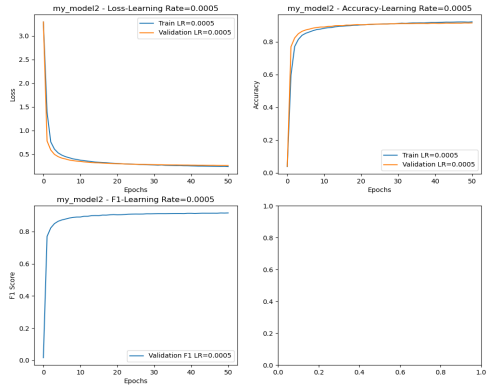
Best Performance:

- Best epoch: 41
- train loss: 0.1883
- validation loss: 0.2434
- train accuracy: 0.9366
- validation accuracy: 0.9208
- validation f1 score: 0.9205

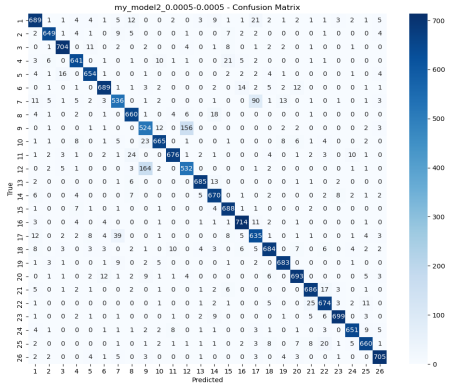
1.3.5 Learning Rate = 0.0005

Best Performance:

- Best epoch: 50
- train loss: 0.2819
- validation loss: 0.2898
- train accuracy: 0.9101
- validation accuracy: 0.9074
- validation f1 score: 0.9070



(a) Model 2 Learning Curve with Learning Rate 0.0005 0.0005



(b) Model 2 Confusion Matrix with Learning Rate

Figure 8: Model 2 Learning Curve and Confusion Matrix with Learning Rate 0.0005

2 Best Model

2.1 Model Architecture

```
Dense(784, 1024),
Relu(),
Dropout(probability=.2),
Dense(1024, 512),
Relu(),
Dropout(probability=.2),
Dense(512,26),
Softmax()
```

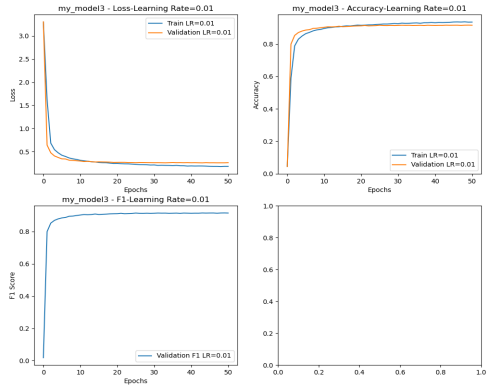
2.2 Learning Rate

Learning Rate = 0.005

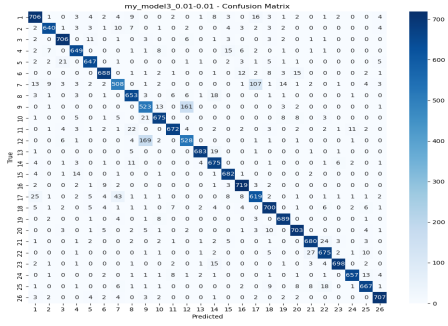
2.3 Performance on validation set

Best Performance:

- Best epoch: 50
- train loss: 0.0936
- validation loss: 0.2478
- train accuracy: 0.9647
- validation accuracy: 0.9268



(a) Model 3 Learning Curve with Learning Rate 0.01



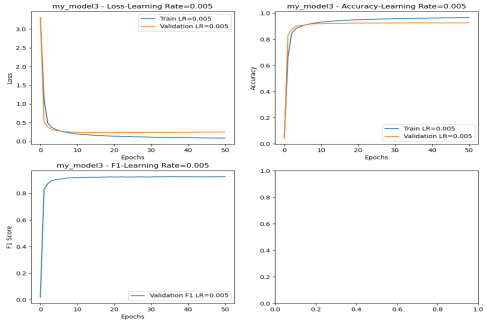
(b) Model 3 Confusion Matrix with Learning Rate 0.01

Figure 9: Model 3 Learning Curve and Confusion Matrix with Learning Rate 0.01

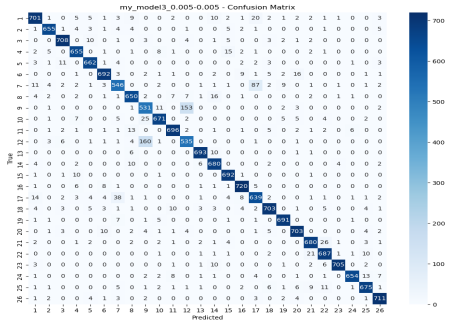
- validation f1 score: 0.9265

3 Performance on Test Set

- Test Loss: 0.26707993365492666
- Test Accuracy: 0.9237019230769231
- Test Macro F1 Score: 0.9237425799805024

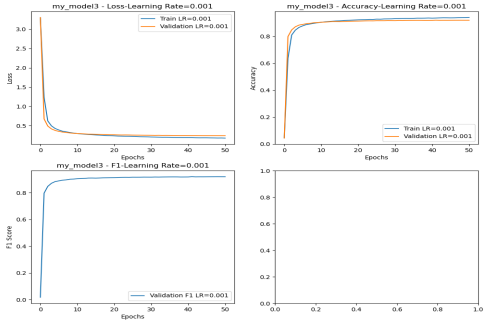


(a) Model 3 Learning Curve with Learning Rate 0.005

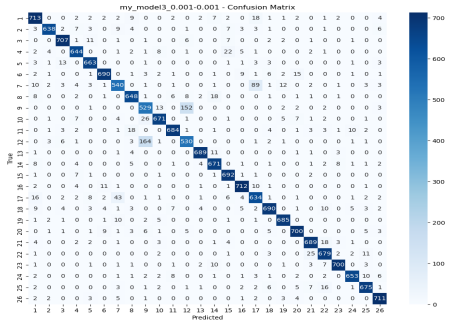


(b) Model 3 Confusion Matrix with Learning Rate 0.005

Figure 10: Model 3 Learning Curve and Confusion Matrix with Learning Rate 0.005

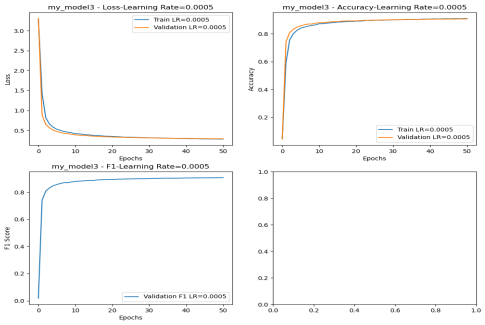


(a) Model 3 Learning Curve with Learning Rate 0.001 0.001

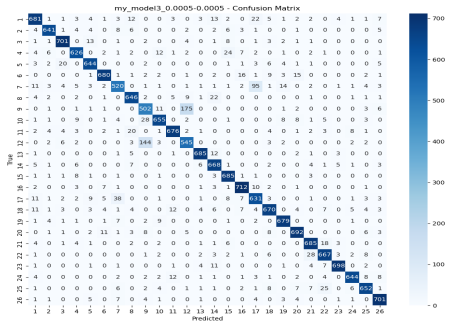


(b) Model 3 Confusion Matrix with Learning Rate 0.001

Figure 11: Model 3 Learning Curve and Confusion Matrix with Learning Rate 0.001

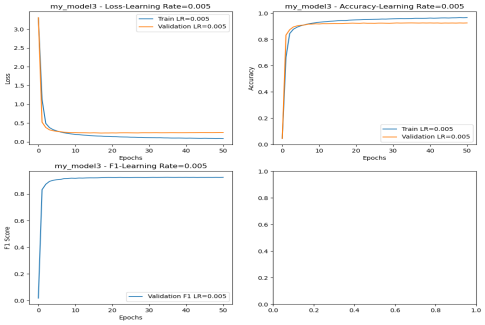


(a) Model 3 Learning Curve with Learning Rate 0.0005 0.0005

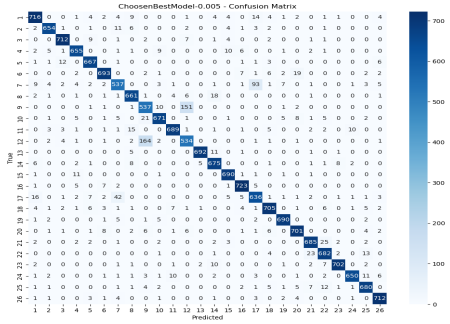


(b) Model 3 Confusion Matrix with Learning Rate

Figure 12: Model 3 Learning Curve and Confusion Matrix with Learning Rate 0.0005



(a) Choosen Best Model Learning Curve



(b) Choosen Best Model Confusion Matrix

Figure 13: Choosen Best Model Learning Curve and Confusion Matrix

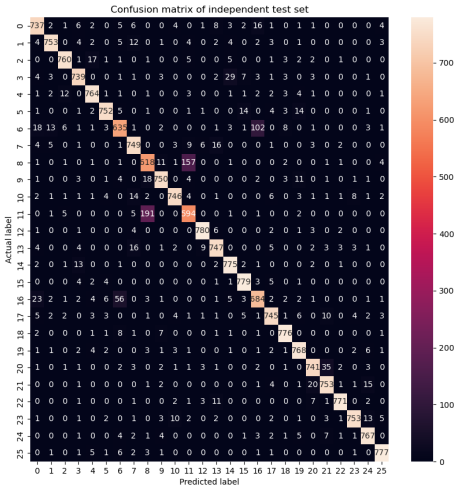


Figure 14: Choosen Best Model Confusion Matrix on Test Set