Assignment 3: Feed Forward Neural Networks

CSE472 (Machine Learning Sessional)

Bijoy Ahmed Saiem - 1905052

Instructions to run the code

Upload notebook in Kaggle: Log into Kaggle account and go to the code section. Click on new notebook and then select the upload notebook option to upload the 1905052.ipynb file.

Run the notebook: Once uploaded, we can run the notebook by clicking on the “Run All” button which will execute all the cells in sequence and generate the corresponding graphs and confusion matrix and other metrics for each model and finally the chosen best model is reported.

Model:

We have run three different models with 10 epochs, learning rates of 0.005, 0.0025, 0.00125, 0.000625 and batch sizes of 2000 . The model specifications are as follows:

1. Dense layer (784, 512), ReLU, Dropout(0.5), Batch Normalization Layer(512), Dense Layer(512, 10), Cross Entropy Loss with Softmax

2. Dense layer (784, 1024), ReLU, Dropout(0.5), Batch Normalization Layer(1024), Dense Layer(1024, 10), Cross Entropy Loss with Softmax

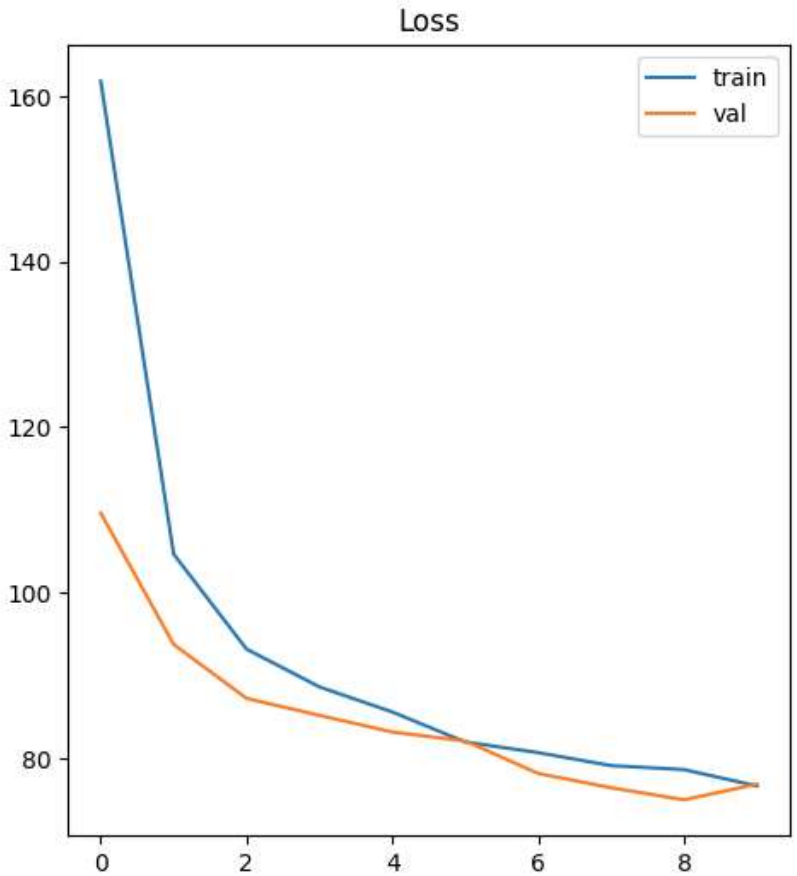
3. Dense layer (784, 256), Sigmoid, Dropout(0.5), Batch Normalization Layer(256), Dense Layer(256, 10), Cross Entropy Loss with Softmax

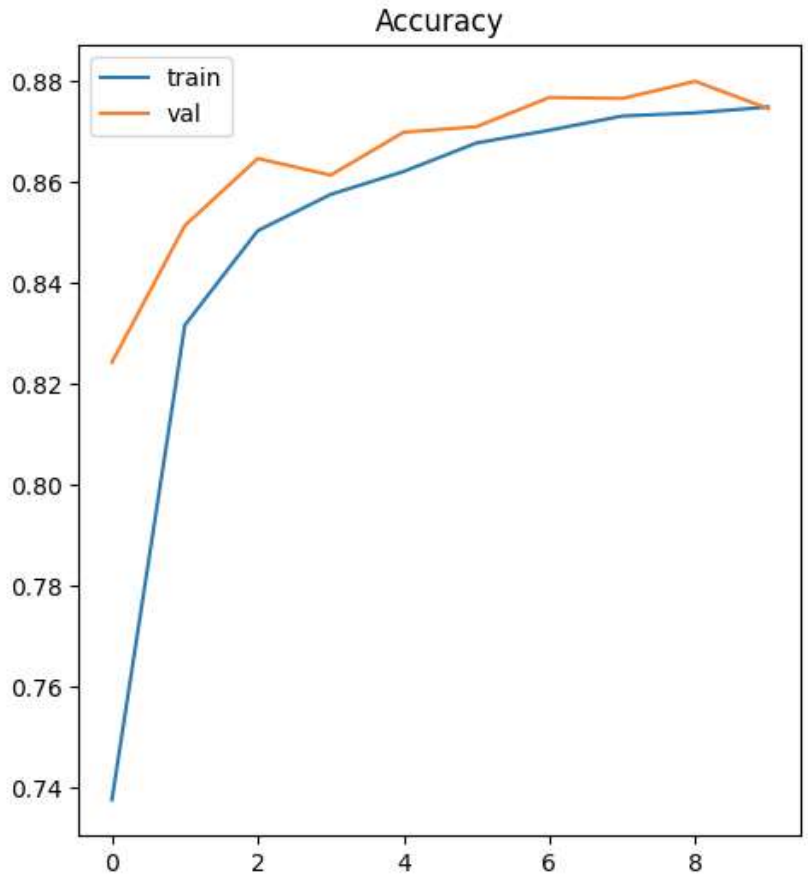
Model 1(LR = 0.005)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

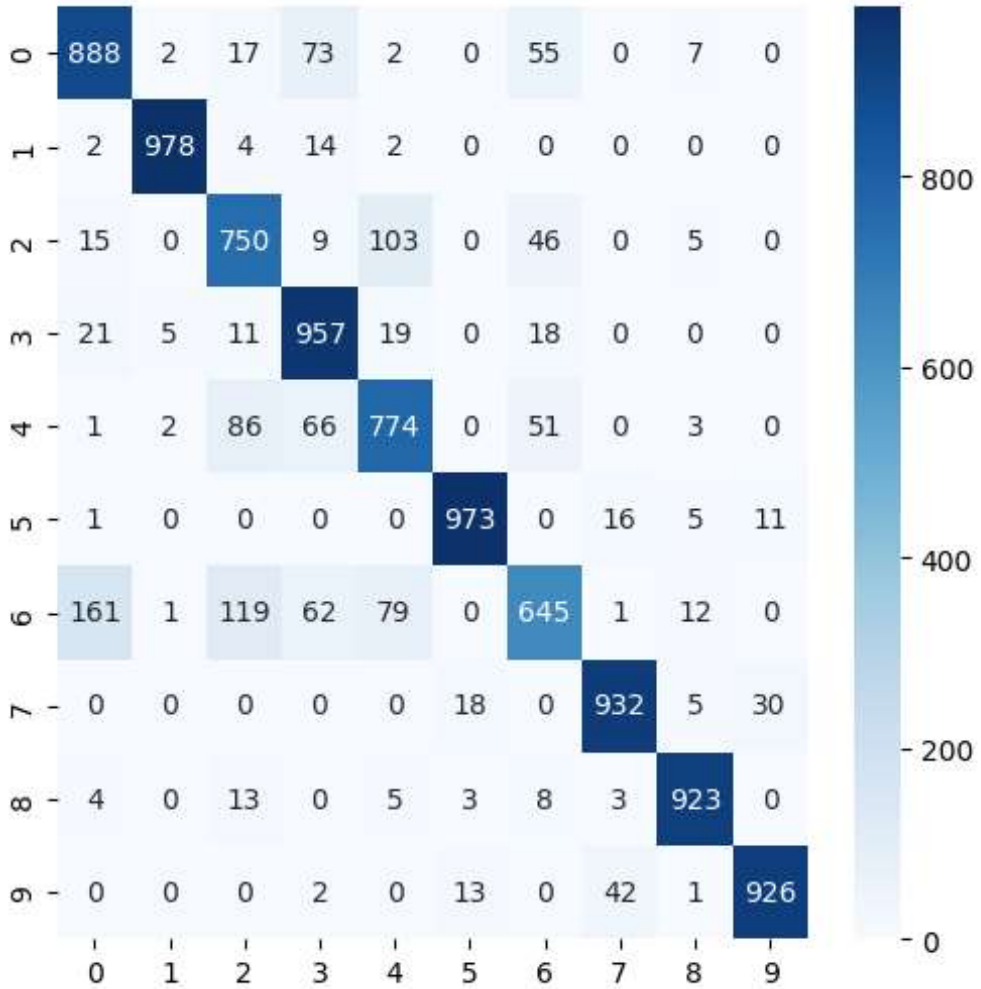
```
Epoch 1/10 - Train Loss: 161.7971, Acc: 0.7376 | Val Loss: 109.6232, Acc: 0.8243, F1: 0.8218
Epoch 2/10 - Train Loss: 104.6542, Acc: 0.8317 | Val Loss: 93.7579, Acc: 0.8514, F1: 0.8522
Epoch 3/10 - Train Loss: 93.1811, Acc: 0.8504 | Val Loss: 87.2169, Acc: 0.8647, F1: 0.8657
Epoch 4/10 - Train Loss: 88.6103, Acc: 0.8576 | Val Loss: 85.1769, Acc: 0.8614, F1: 0.8629
Epoch 5/10 - Train Loss: 85.5803, Acc: 0.8621 | Val Loss: 83.1632, Acc: 0.8699, F1: 0.8693
Epoch 6/10 - Train Loss: 81.9579, Acc: 0.8678 | Val Loss: 82.0762, Acc: 0.8710, F1: 0.8712
Epoch 7/10 - Train Loss: 80.6714, Acc: 0.8703 | Val Loss: 78.1531, Acc: 0.8768, F1: 0.8774
Epoch 8/10 - Train Loss: 79.1015, Acc: 0.8731 | Val Loss: 76.4279, Acc: 0.8766, F1: 0.8777
Epoch 9/10 - Train Loss: 78.6228, Acc: 0.8737 | Val Loss: 74.9725, Acc: 0.8800, F1: 0.8809
Epoch 10/10 - Train Loss: 76.6426, Acc: 0.8749 | Val Loss: 76.8708, Acc: 0.8746, F1: 0.8742
```

2. Graphs


Loss
train
val
160
140
120
100
80
0
2
4
6
8


Accuracy
train
val
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0
2
4
6
8

3. Confusion Matrix

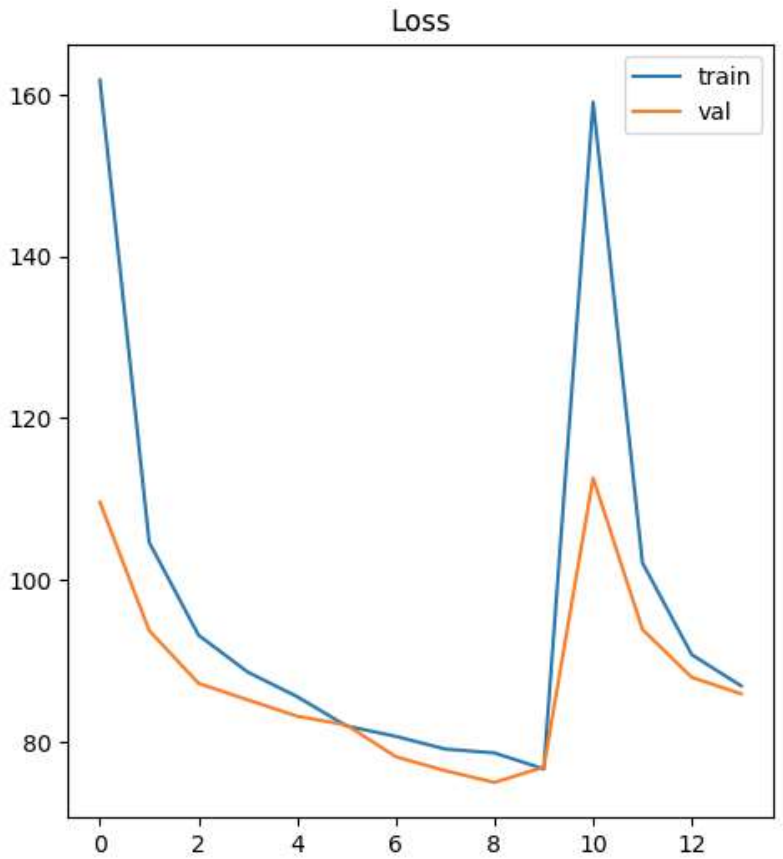


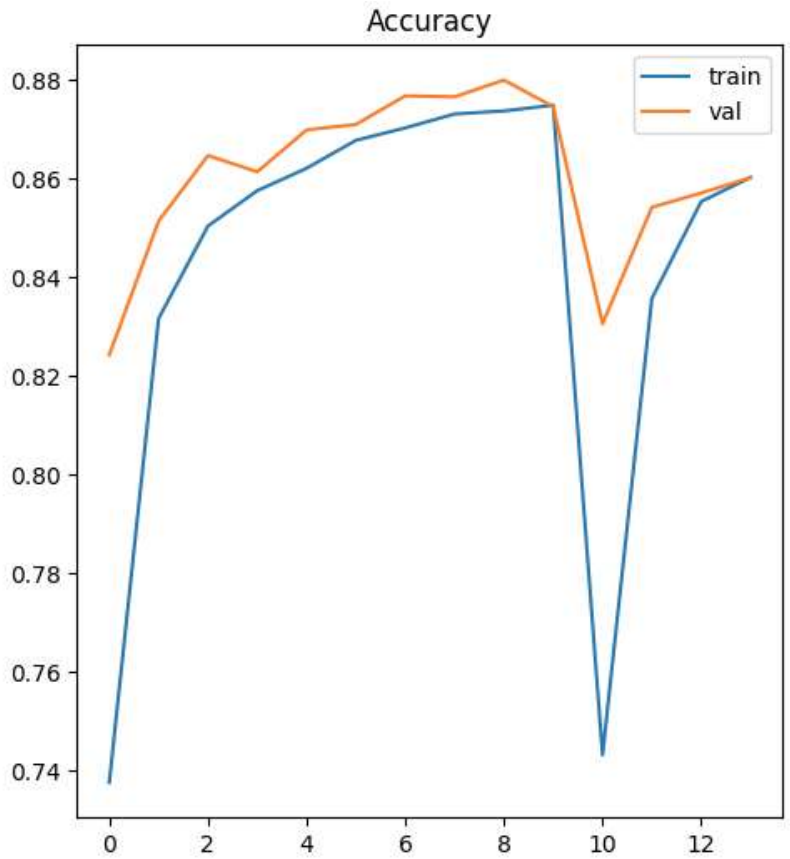
Model 1(LR = 0.0025)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

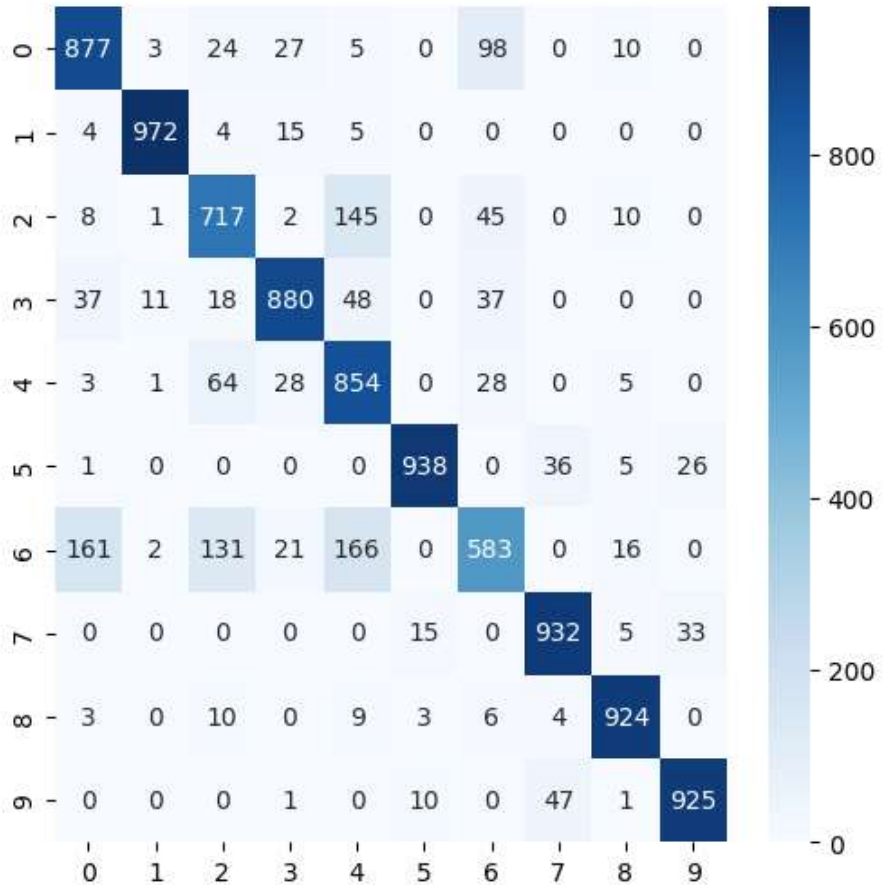
```
Epoch 1/10 - Train Loss: 159.0745, Acc: 0.7432 | Val Loss: 112.6173, Acc: 0.8306, F1: 0.8300
Epoch 2/10 - Train Loss: 102.1264, Acc: 0.8357 | Val Loss: 93.8919, Acc: 0.8542, F1: 0.8559
Epoch 3/10 - Train Loss: 90.7694, Acc: 0.8554 | Val Loss: 87.9580, Acc: 0.8571, F1: 0.8580
Epoch 4/10 - Train Loss: 86.9175, Acc: 0.8603 | Val Loss: 85.9357, Acc: 0.8602, F1: 0.8600
Early stopping triggered.
```

2. Graphs


Loss
train
val
160
140
120
100
80
0
2
4
6
8
10
12


Accuracy
train
val
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0
2
4
6
8
10
12

3. Confusion Matrix

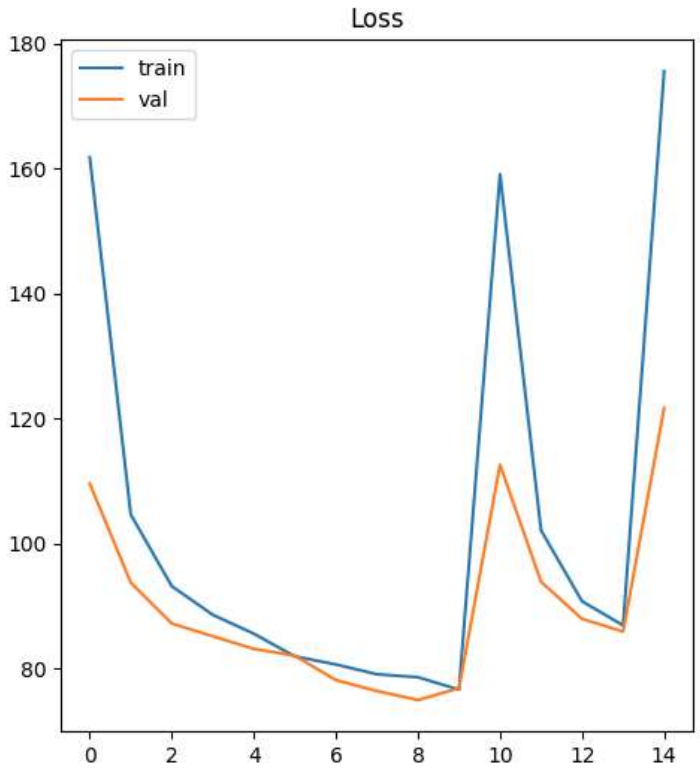


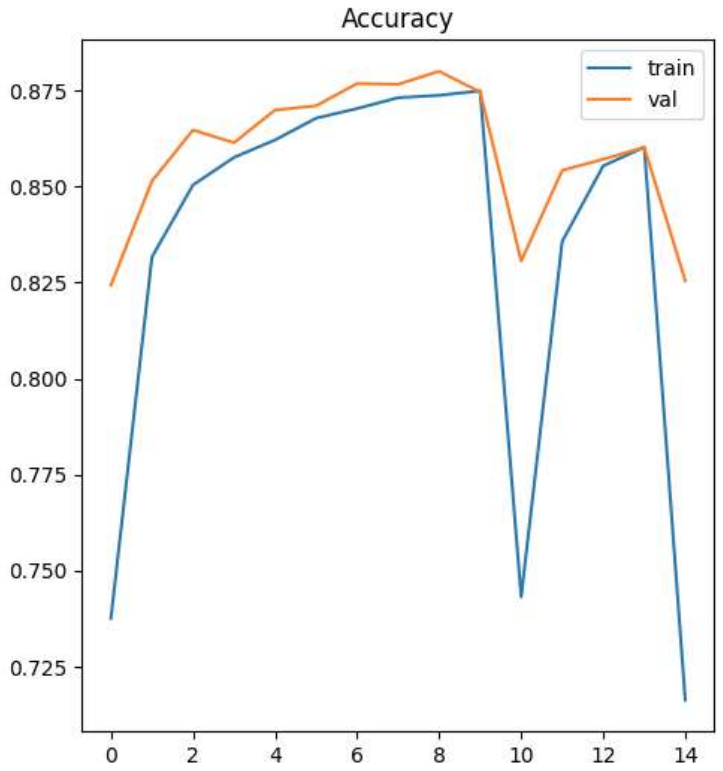
Model 1(LR = 0.00125)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

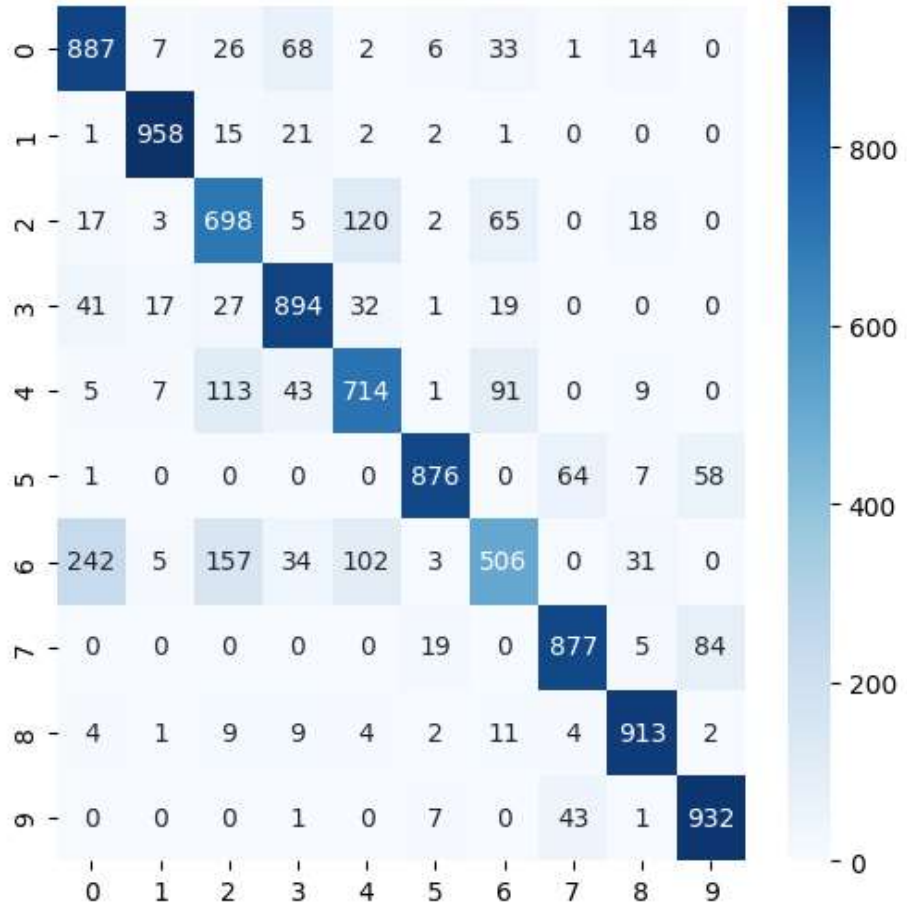
```
Epoch 1/10 - Train Loss: 175.5662, Acc: 0.7163 | Val Loss: 121.6837, Acc: 0.8255, F1: 0.8233
Early stopping triggered.
```

2. Graphs





3. Confusion Matrix

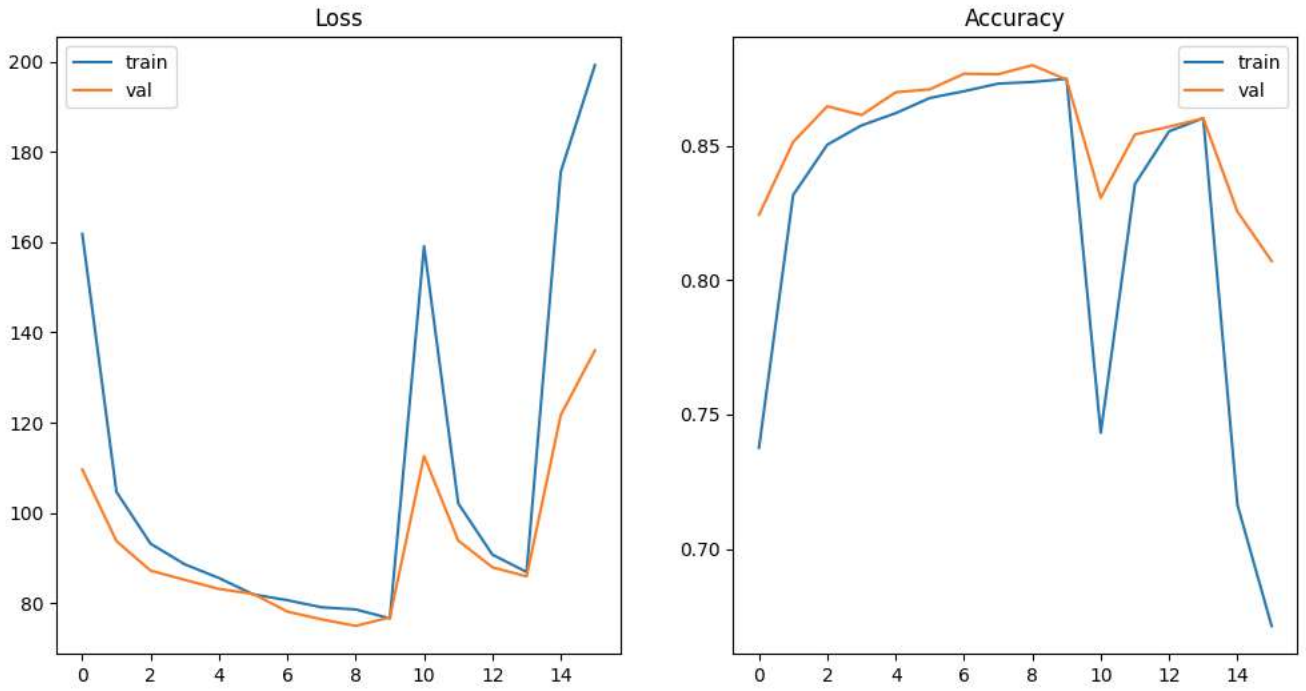


Model 1(LR = 0.000625)

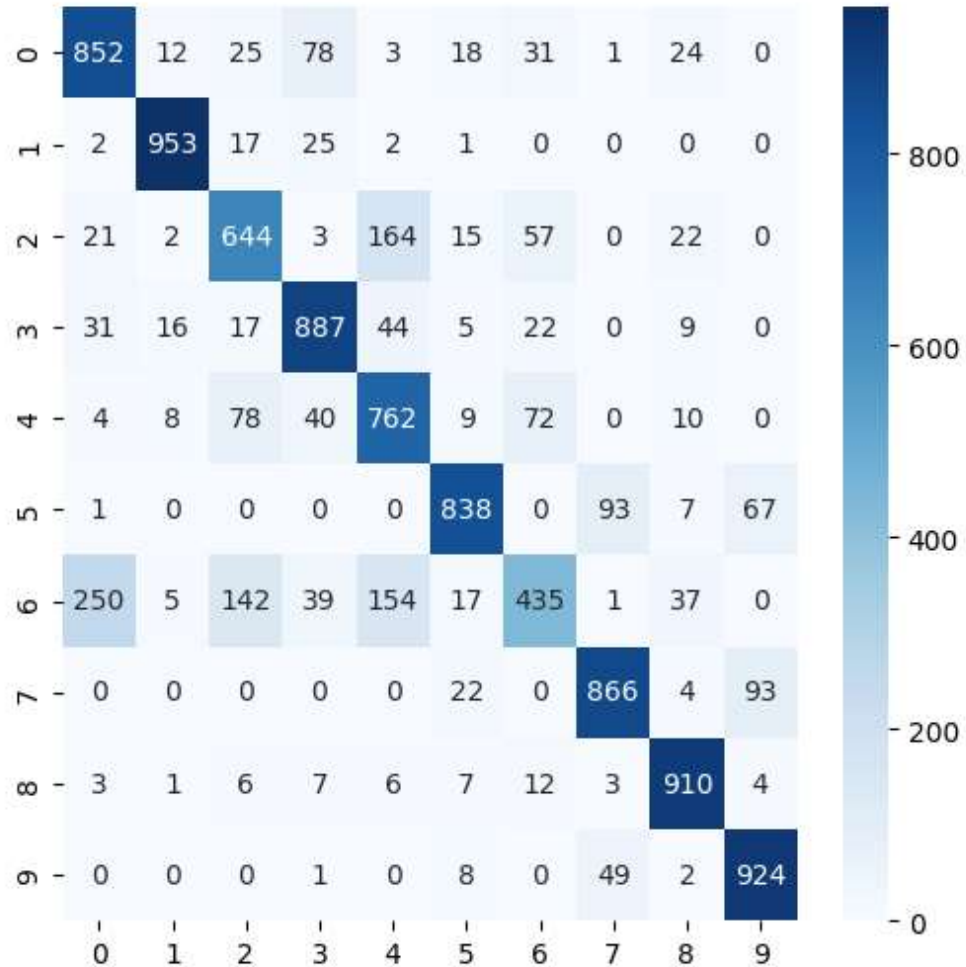
1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

```
Epoch 1/10 - Train Loss: 199.1932, Acc: 0.6713 | Val Loss: 136.0131, Acc: 0.8071, F1: 0.8026
Early stopping triggered.
```

2. Graphs



3. Confusion Matrix

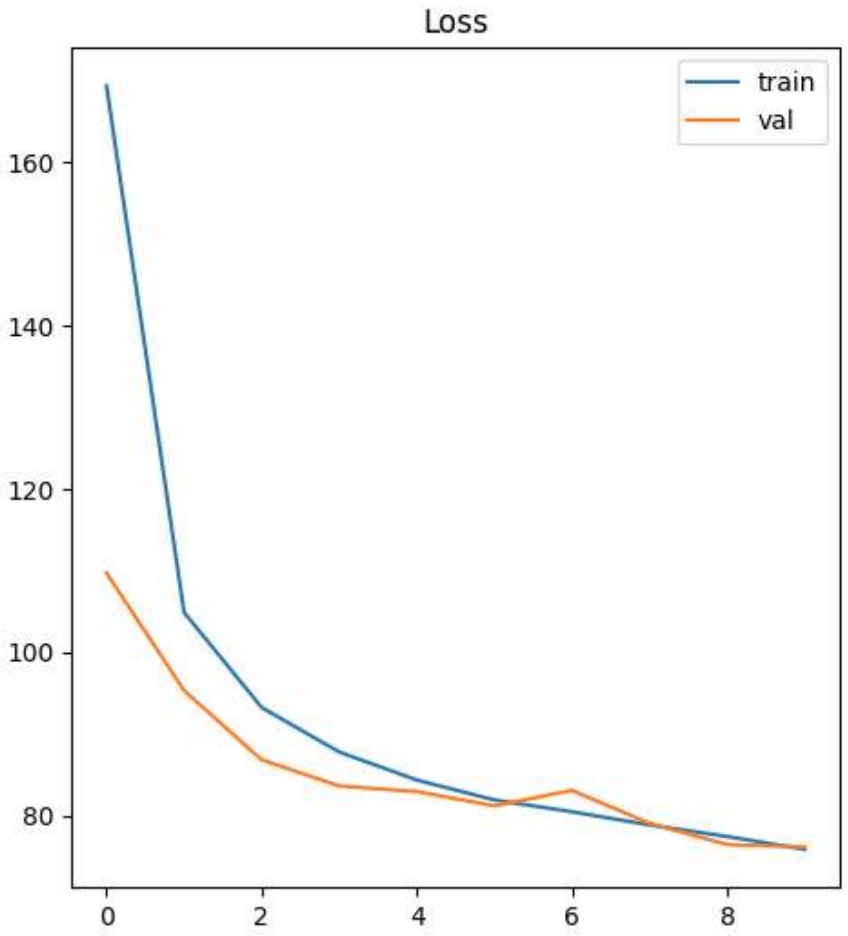


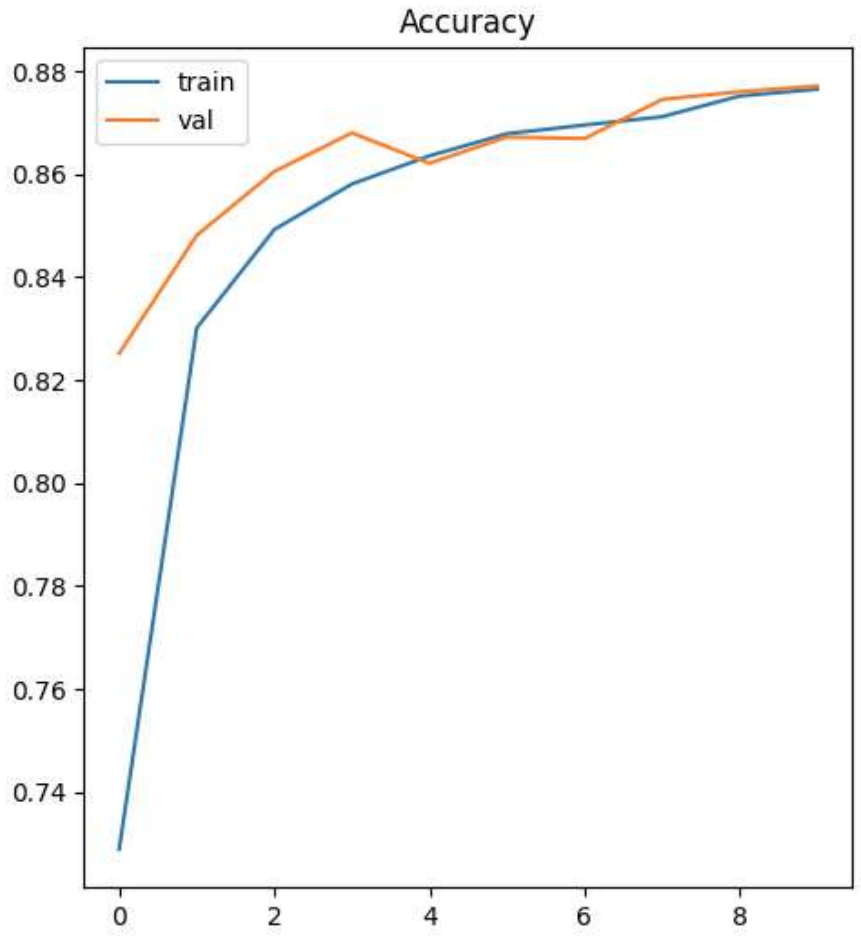
Model 2(LR = 0.005)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

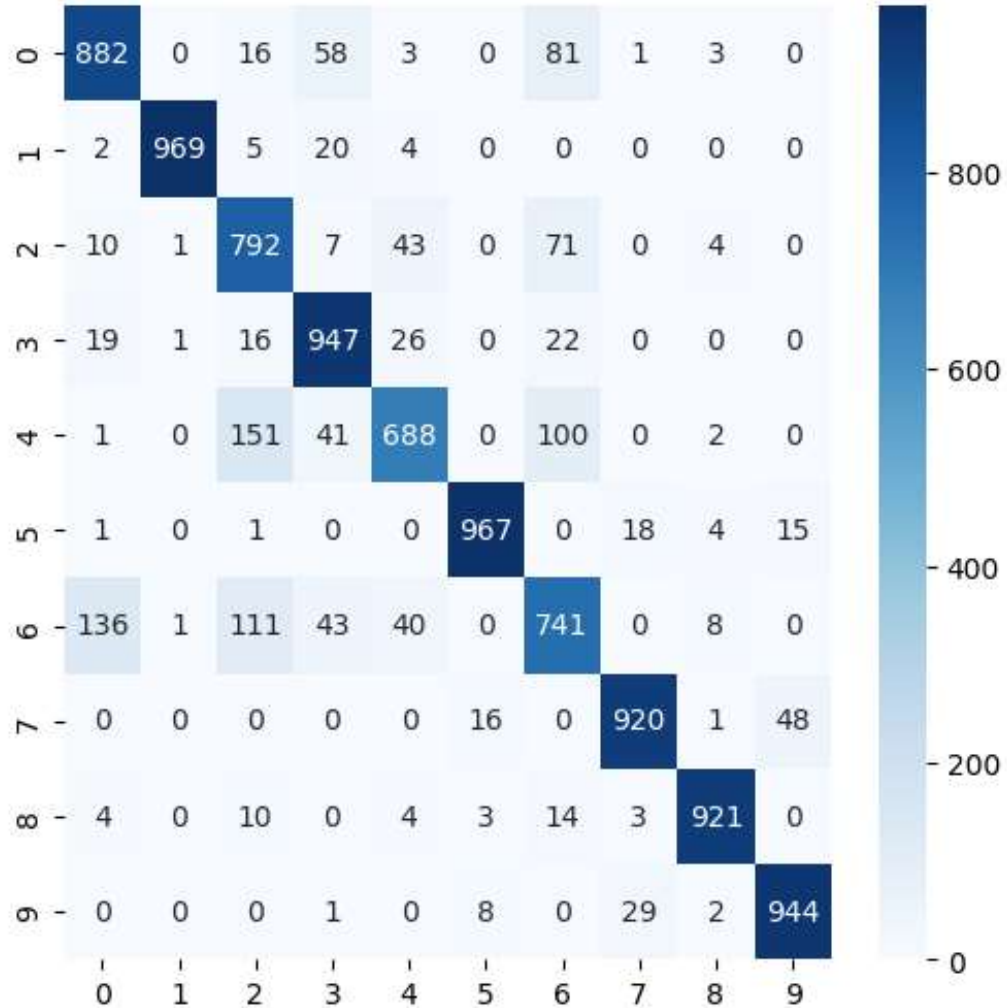
```
Epoch 1/10 - Train Loss: 169.3852, Acc: 0.7289 | Val Loss: 109.6981, Acc: 0.8252, F1: 0.8236
Epoch 2/10 - Train Loss: 104.8635, Acc: 0.8301 | Val Loss: 95.2955, Acc: 0.8481, F1: 0.8469
Epoch 3/10 - Train Loss: 93.2083, Acc: 0.8492 | Val Loss: 86.8026, Acc: 0.8605, F1: 0.8611
Epoch 4/10 - Train Loss: 87.7518, Acc: 0.8581 | Val Loss: 83.5889, Acc: 0.8680, F1: 0.8673
Epoch 5/10 - Train Loss: 84.3159, Acc: 0.8635 | Val Loss: 82.9005, Acc: 0.8621, F1: 0.8653
Epoch 6/10 - Train Loss: 81.8762, Acc: 0.8678 | Val Loss: 81.1587, Acc: 0.8672, F1: 0.8675
Epoch 7/10 - Train Loss: 80.4290, Acc: 0.8695 | Val Loss: 83.0395, Acc: 0.8669, F1: 0.8668
Epoch 8/10 - Train Loss: 78.8149, Acc: 0.8711 | Val Loss: 79.0124, Acc: 0.8745, F1: 0.8761
Epoch 9/10 - Train Loss: 77.3962, Acc: 0.8752 | Val Loss: 76.4046, Acc: 0.8760, F1: 0.8780
Epoch 10/10 - Train Loss: 75.8220, Acc: 0.8765 | Val Loss: 76.1327, Acc: 0.8771, F1: 0.8780
```

2. Graphs


Loss
train
val
160
140
120
100
80
0
2
4
6
8


Accuracy
train
val
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0
2
4
6
8

3. Confusion Matrix

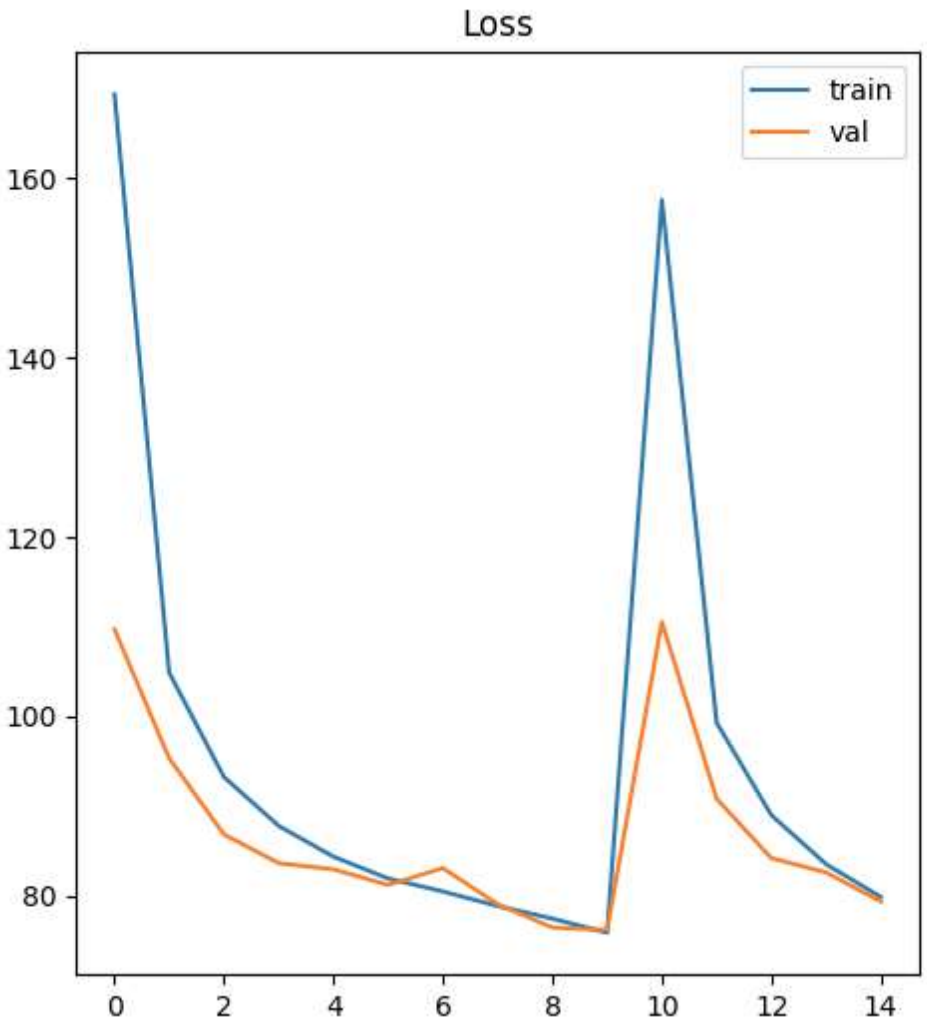


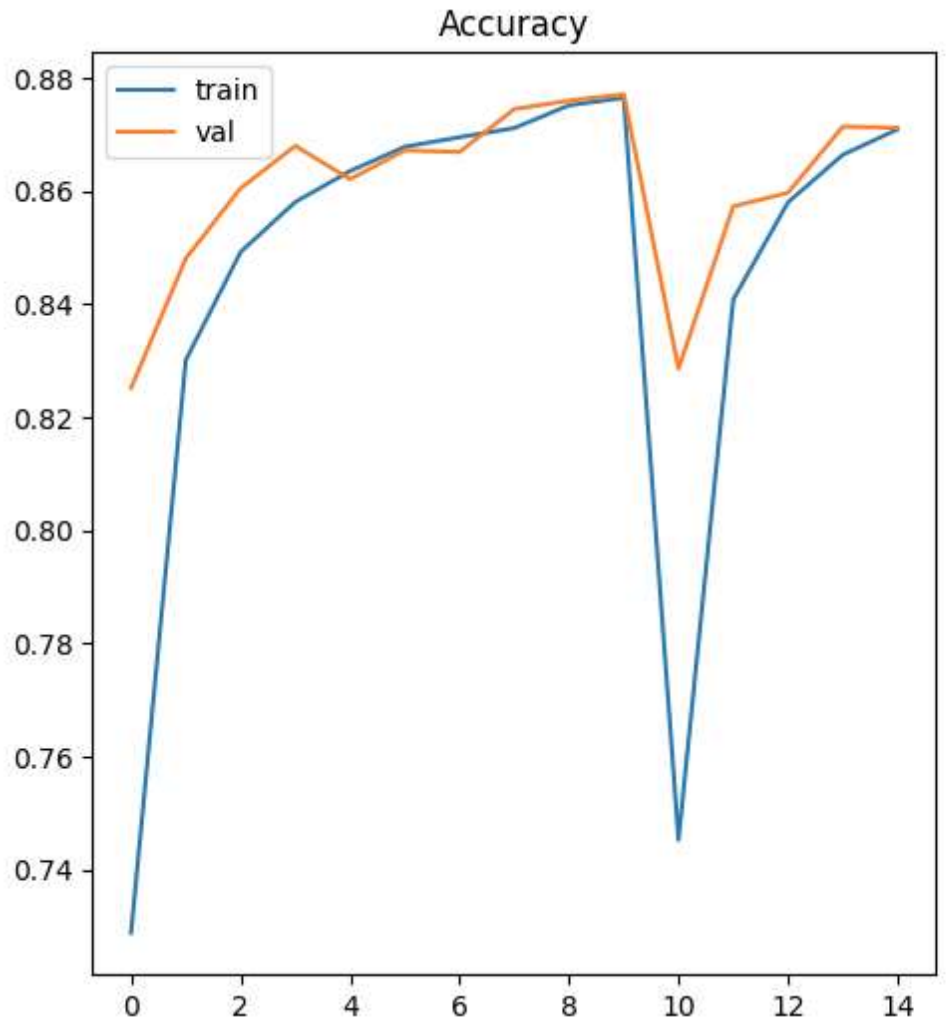
Model 2(LR = 0.0025)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

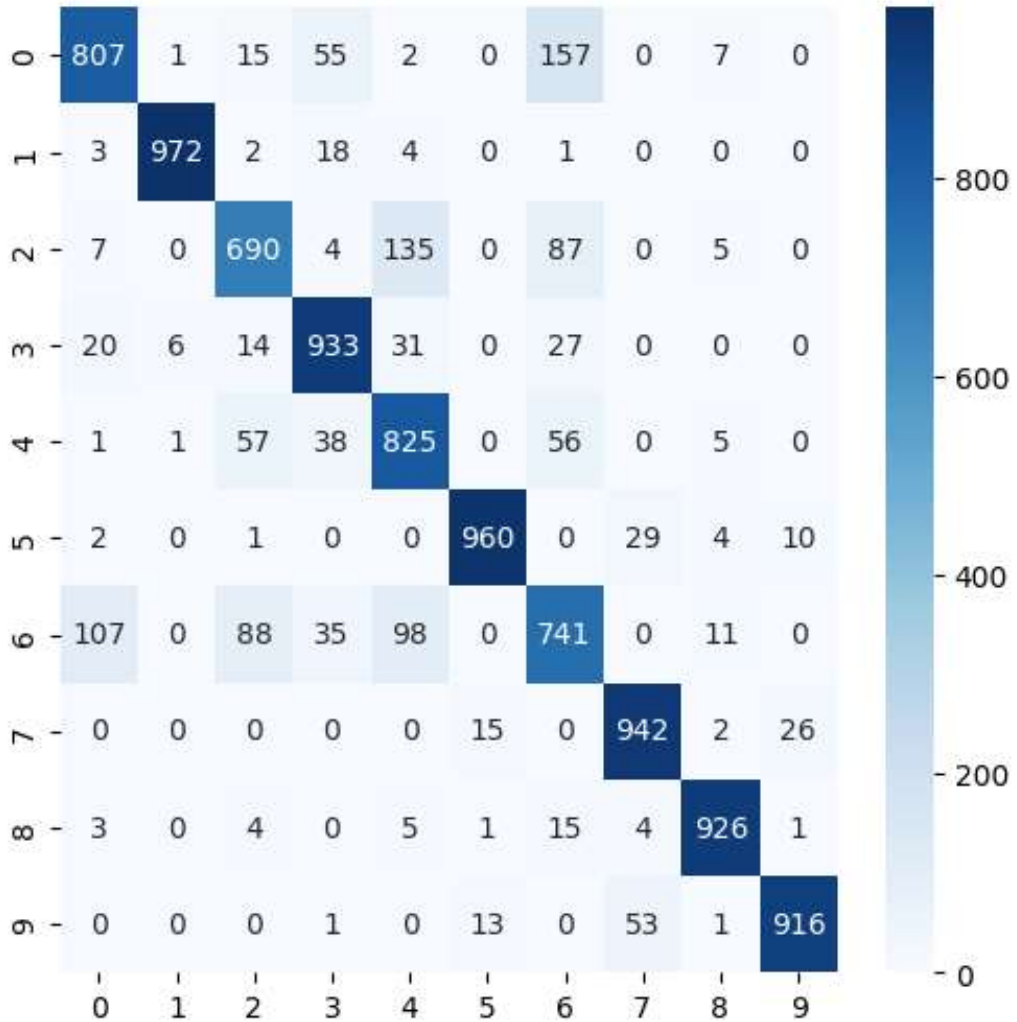
```
Epoch 1/10 - Train Loss: 157.6008, Acc: 0.7453 | Val Loss: 110.5034, Acc: 0.8286, F1: 0.8236
Epoch 2/10 - Train Loss: 99.2692, Acc: 0.8408 | Val Loss: 90.7772, Acc: 0.8573, F1: 0.8590
Epoch 3/10 - Train Loss: 88.9533, Acc: 0.8580 | Val Loss: 84.1495, Acc: 0.8597, F1: 0.8616
Epoch 4/10 - Train Loss: 83.4701, Acc: 0.8664 | Val Loss: 82.5547, Acc: 0.8714, F1: 0.8732
Epoch 5/10 - Train Loss: 79.8085, Acc: 0.8709 | Val Loss: 79.3009, Acc: 0.8712, F1: 0.8726
Early stopping triggered.
```

2. Graphs


Loss
train
val
160
140
120
100
80
0
2
4
6
8
10
12
14


Accuracy
train
val
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0
2
4
6
8
10
12
14

3. Confusion Matrix

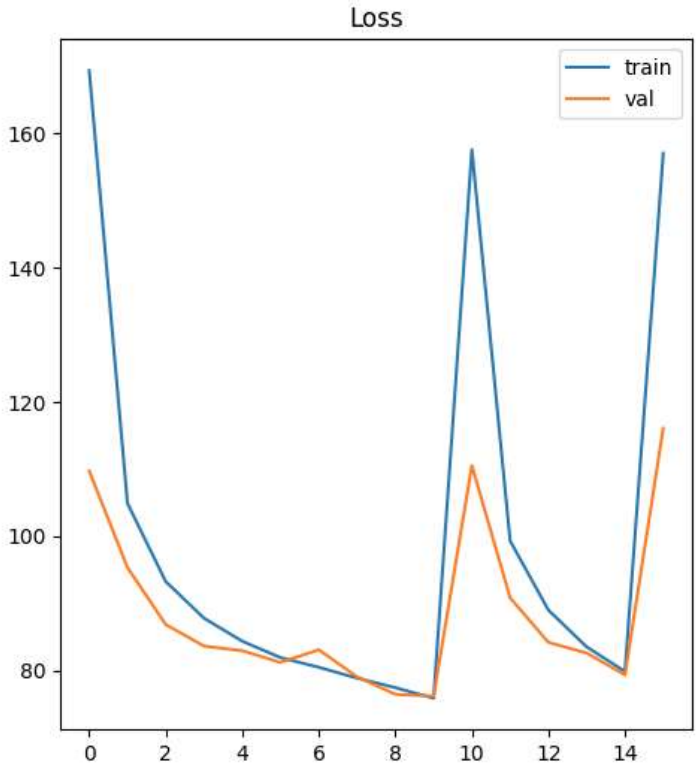


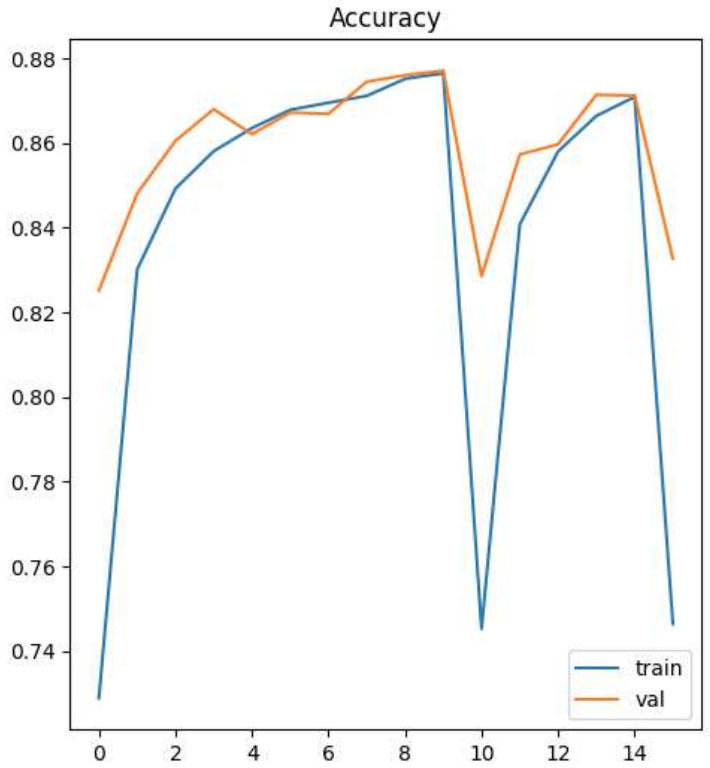
Model 2(LR = 0.00125)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

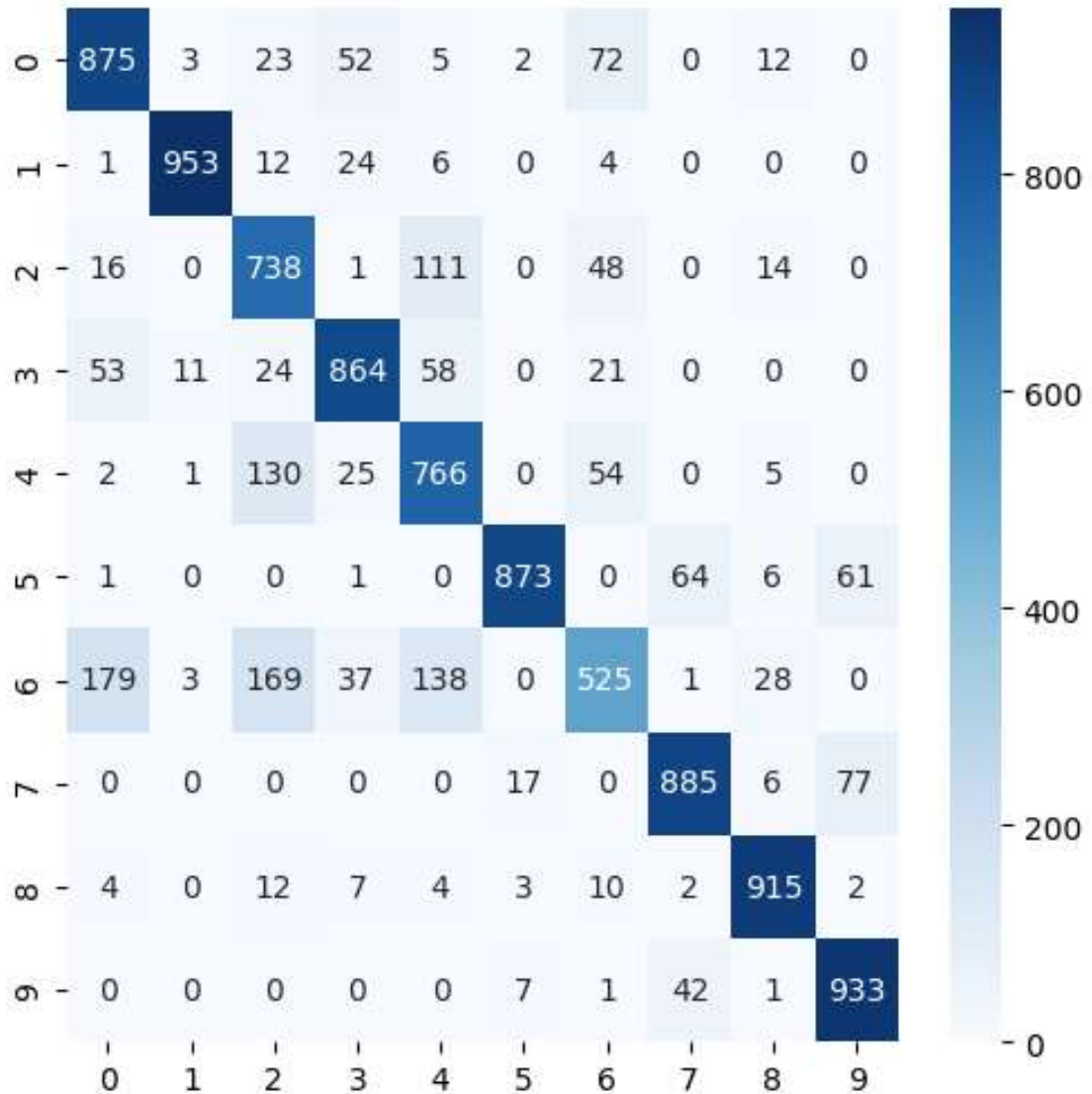
```
Epoch 1/10 - Train Loss: 157.0661, Acc: 0.7465 | Val Loss: 116.0707, Acc: 0.8327, F1: 0.8318
Early stopping triggered.
```

2. Graphs





3. Confusion Matrix

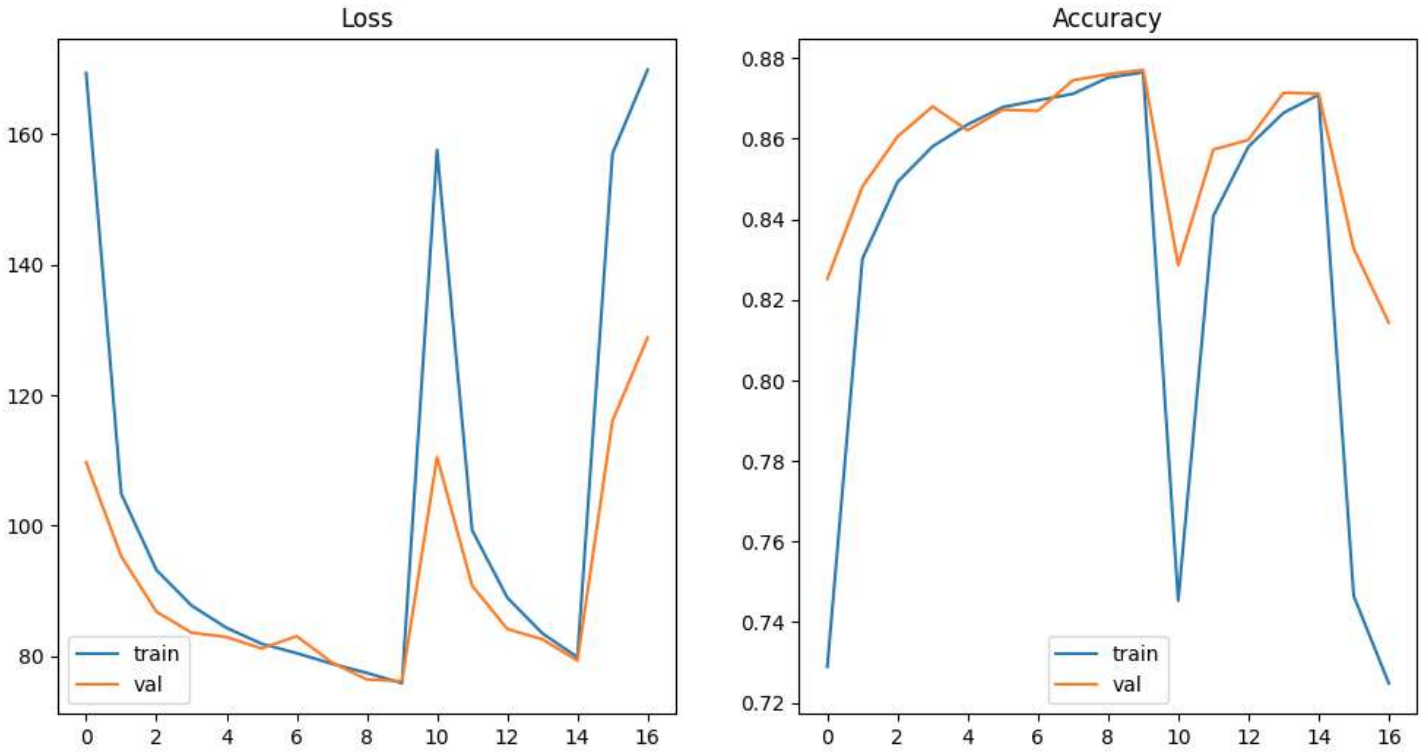


Model 2(LR = 0.000625)

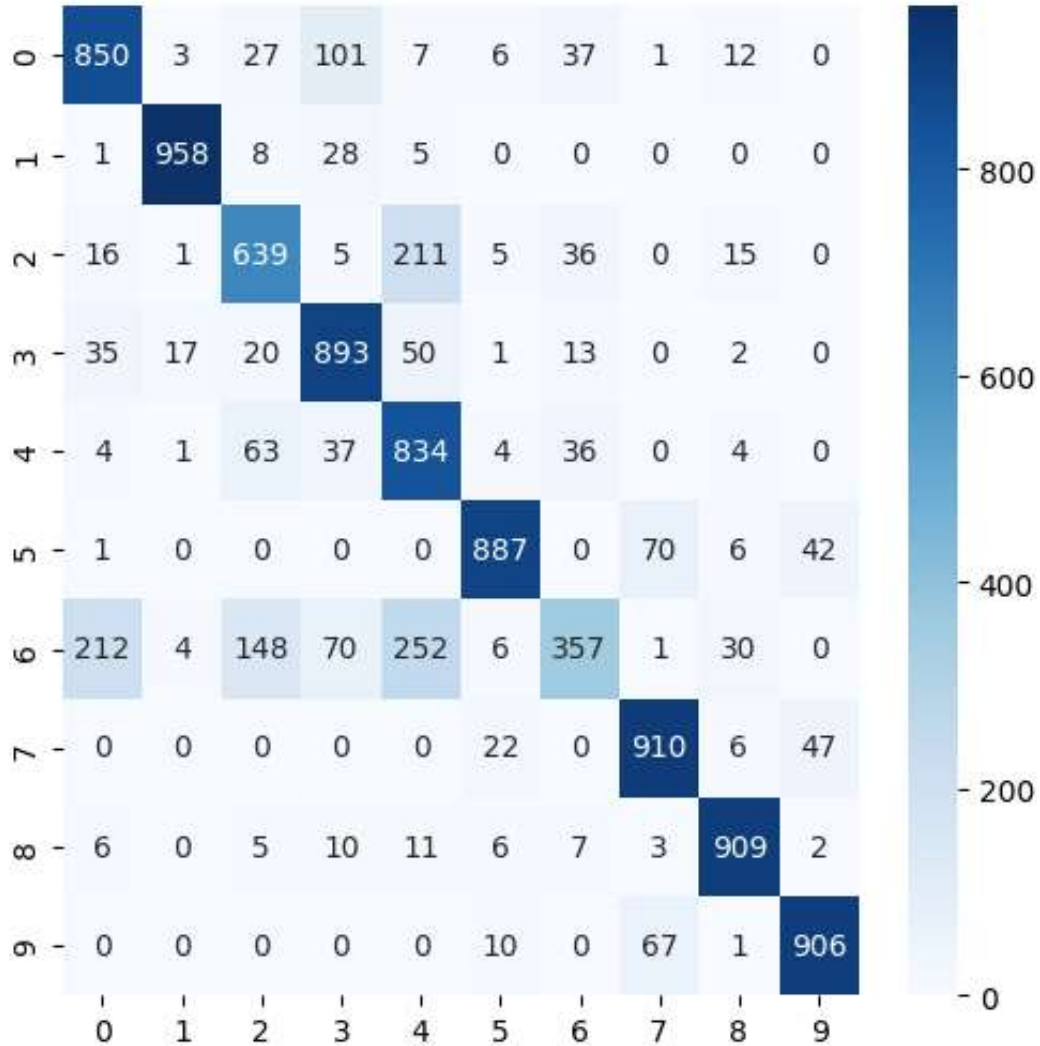
1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

```
Epoch 1/10 - Train Loss: 169.8826, Acc: 0.7248 | Val Loss: 128.8199, Acc: 0.8143, F1: 0.8082
Early stopping triggered.
```

2. Graphs



3. Confusion Matrix

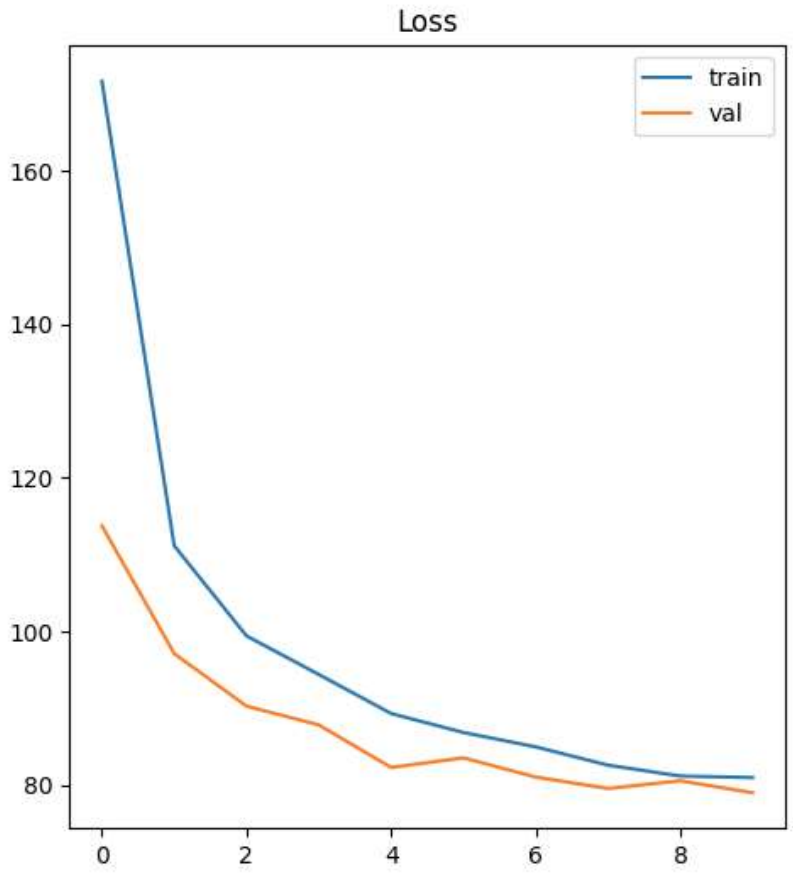


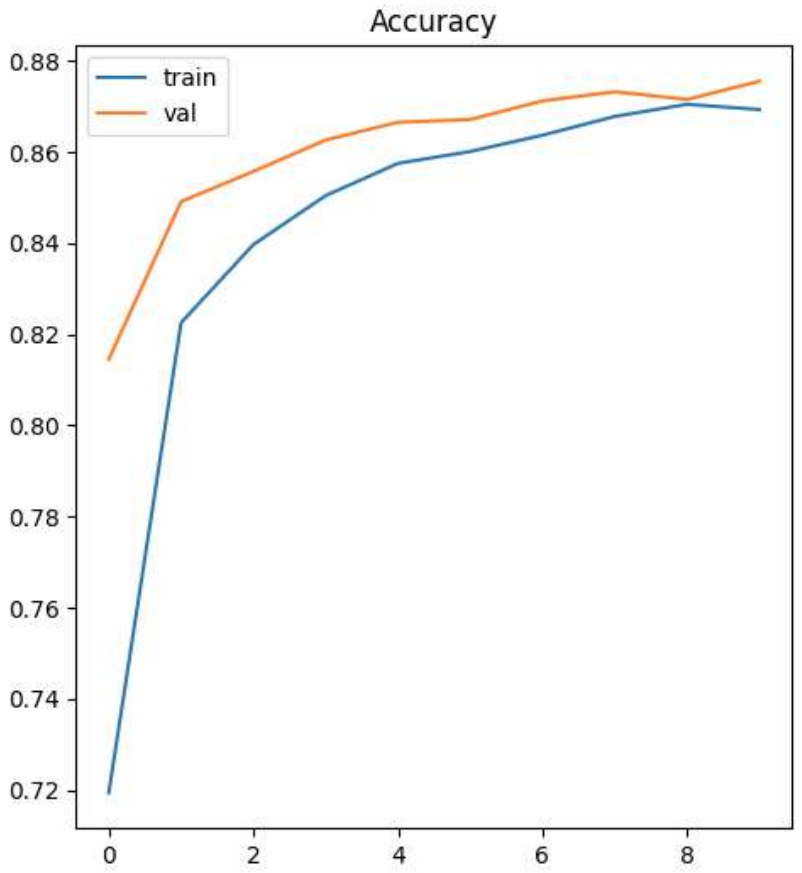
Model 3(LR = 0.005)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

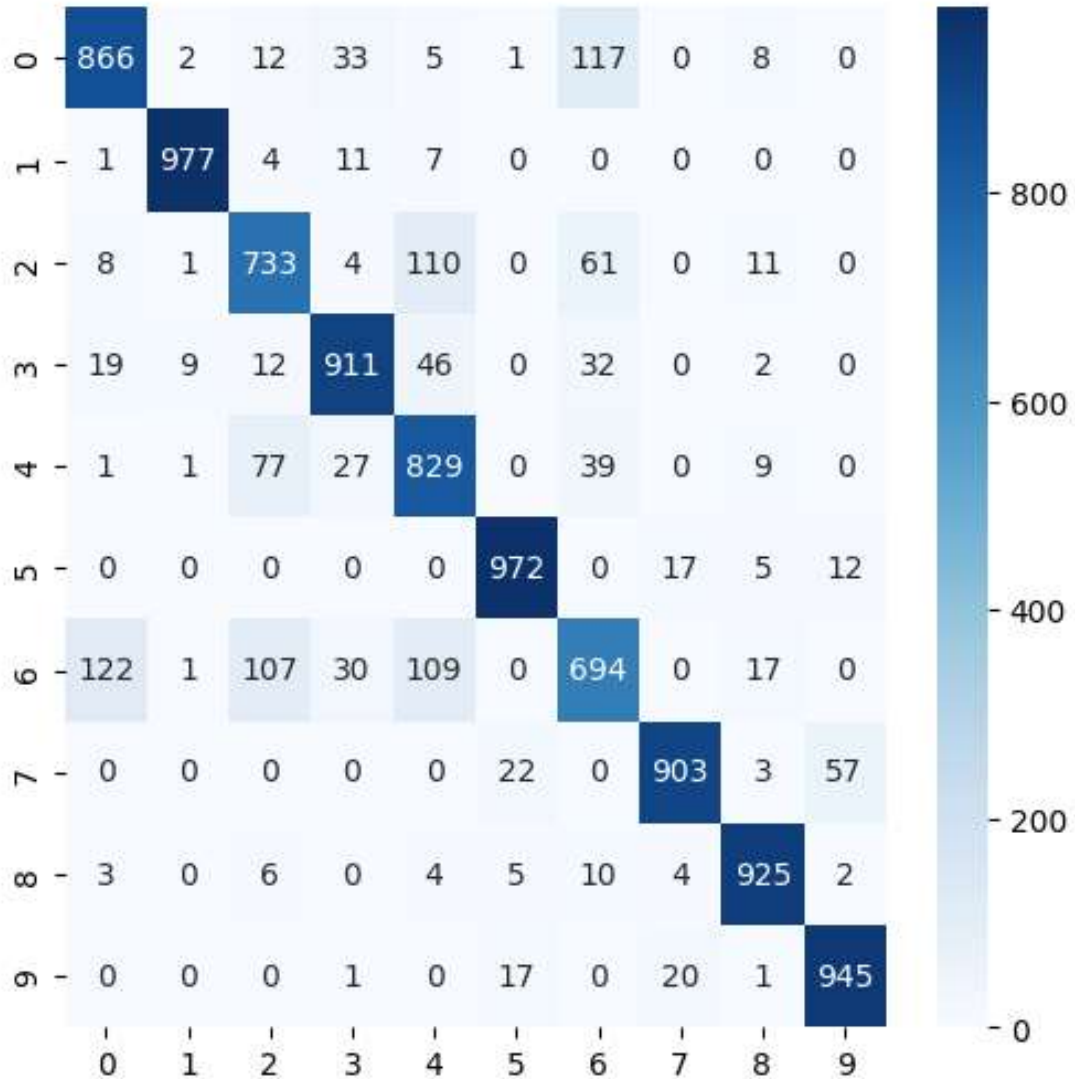
```
Epoch 1/10 - Train Loss: 171.6320, Acc: 0.7194 | Val Loss: 113.7773, Acc: 0.8145, F1: 0.8064
Epoch 2/10 - Train Loss: 111.1632, Acc: 0.8225 | Val Loss: 97.1221, Acc: 0.8491, F1: 0.8503
Epoch 3/10 - Train Loss: 99.4300, Acc: 0.8397 | Val Loss: 90.2841, Acc: 0.8557, F1: 0.8562
Epoch 4/10 - Train Loss: 94.4094, Acc: 0.8504 | Val Loss: 87.8465, Acc: 0.8626, F1: 0.8630
Epoch 5/10 - Train Loss: 89.3275, Acc: 0.8575 | Val Loss: 82.3010, Acc: 0.8665, F1: 0.8673
Epoch 6/10 - Train Loss: 86.8553, Acc: 0.8601 | Val Loss: 83.5576, Acc: 0.8671, F1: 0.8688
Epoch 7/10 - Train Loss: 84.9869, Acc: 0.8637 | Val Loss: 81.0524, Acc: 0.8712, F1: 0.8716
Epoch 8/10 - Train Loss: 82.6021, Acc: 0.8678 | Val Loss: 79.5715, Acc: 0.8732, F1: 0.8737
Epoch 9/10 - Train Loss: 81.1897, Acc: 0.8704 | Val Loss: 80.5718, Acc: 0.8715, F1: 0.8718
Epoch 10/10 - Train Loss: 80.9879, Acc: 0.8693 | Val Loss: 79.0118, Acc: 0.8755, F1: 0.8761
```

2. Graphs


Loss
train
val
160
140
120
100
80
0
2
4
6
8


Accuracy
train
val
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0.72
0
2
4
6
8

3. Confusion Matrix

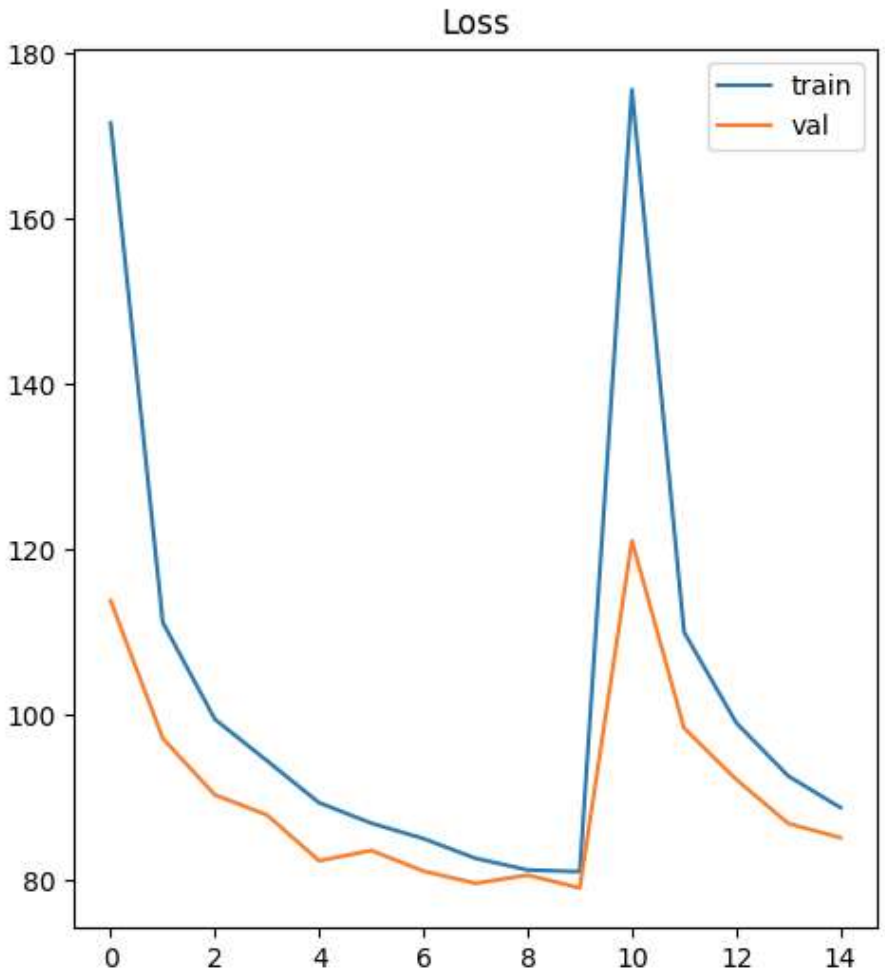


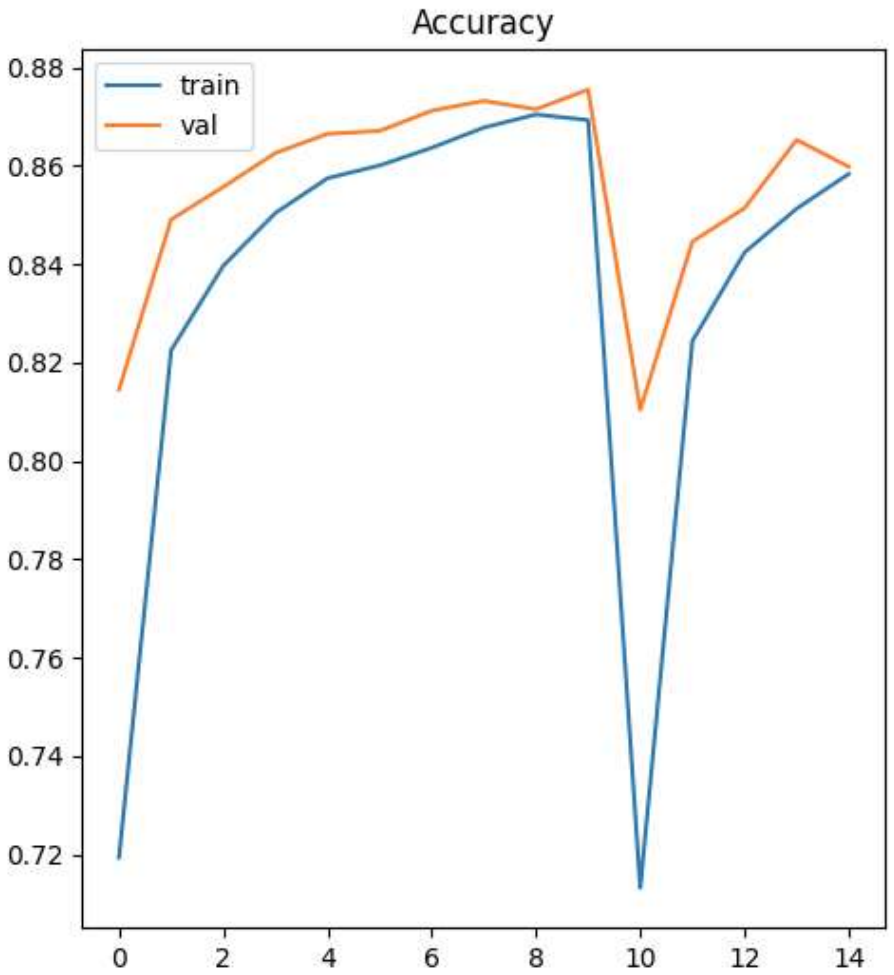
Model 3(LR = 0.0025)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

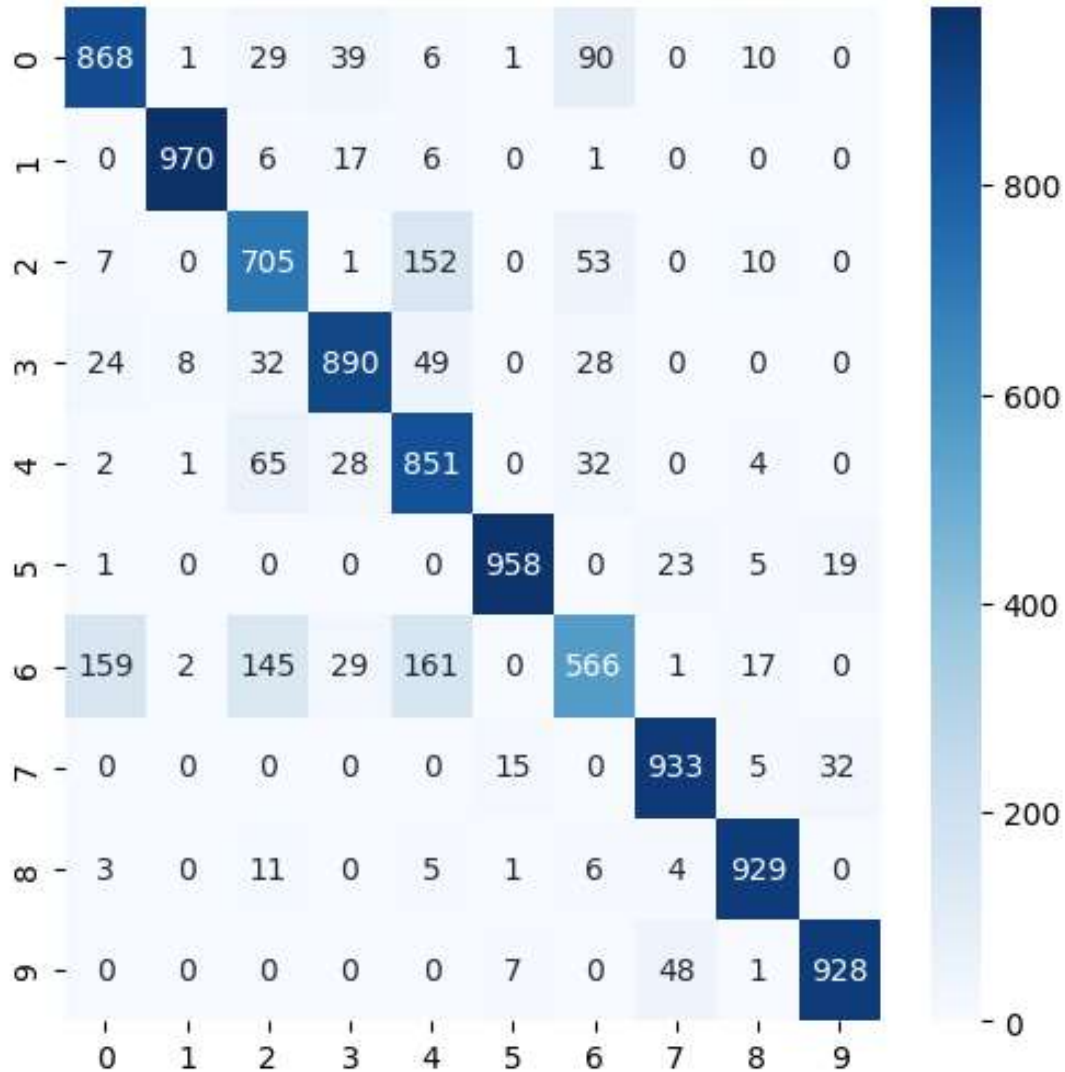
```
Epoch 1/10 - Train Loss: 175.6691, Acc: 0.7132 | Val Loss: 121.0554, Acc: 0.8104, F1: 0.8093
Epoch 2/10 - Train Loss: 109.9913, Acc: 0.8244 | Val Loss: 98.3781, Acc: 0.8446, F1: 0.8430
Epoch 3/10 - Train Loss: 99.0114, Acc: 0.8424 | Val Loss: 92.1531, Acc: 0.8514, F1: 0.8510
Epoch 4/10 - Train Loss: 92.5575, Acc: 0.8513 | Val Loss: 86.8177, Acc: 0.8653, F1: 0.8651
Epoch 5/10 - Train Loss: 88.7393, Acc: 0.8584 | Val Loss: 85.0961, Acc: 0.8598, F1: 0.8593
Early stopping triggered.
```

2. Graphs


Loss
train
val
180
160
140
120
100
80
0
2
4
6
8
10
12
14


Accuracy
train
val
0.88
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0.72
0
2
4
6
8
10
12
14

3. Confusion Matrix

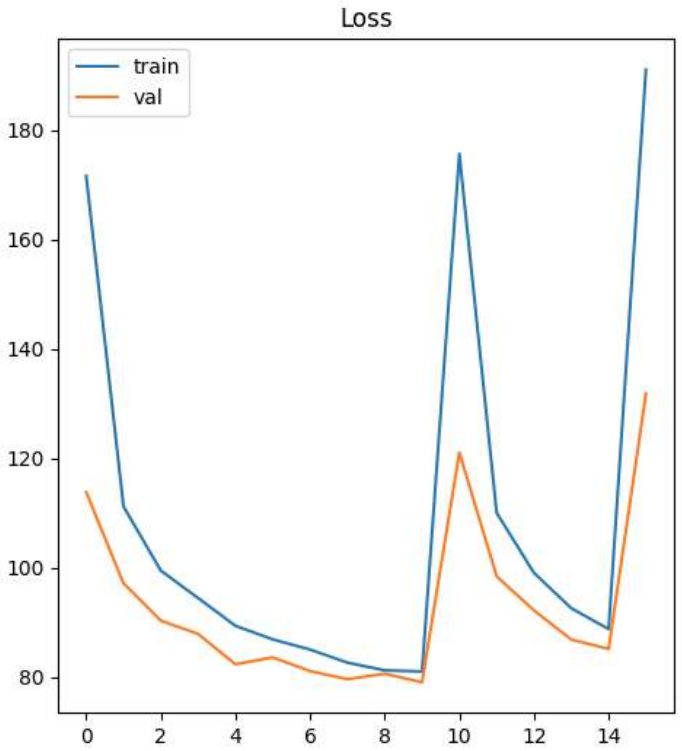


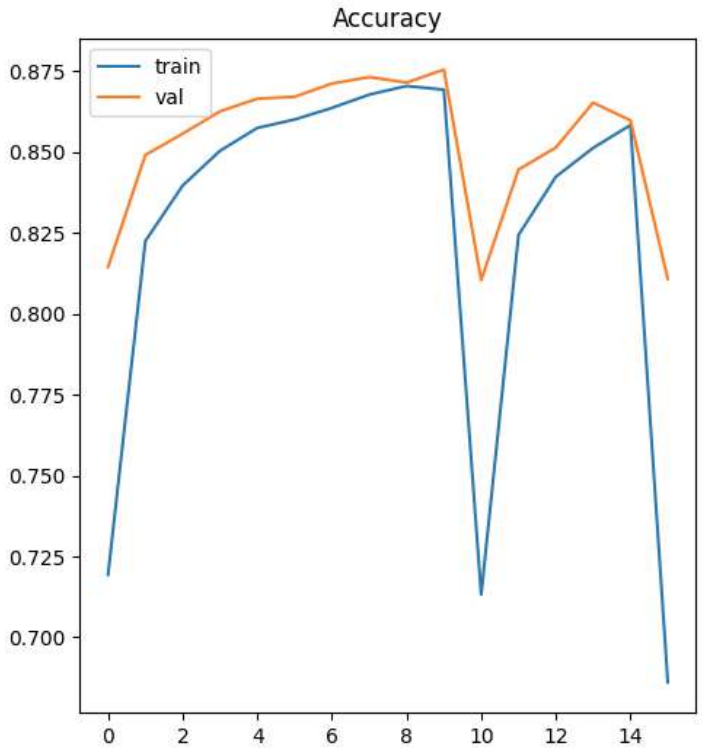
Model 3(LR = 0.00125)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

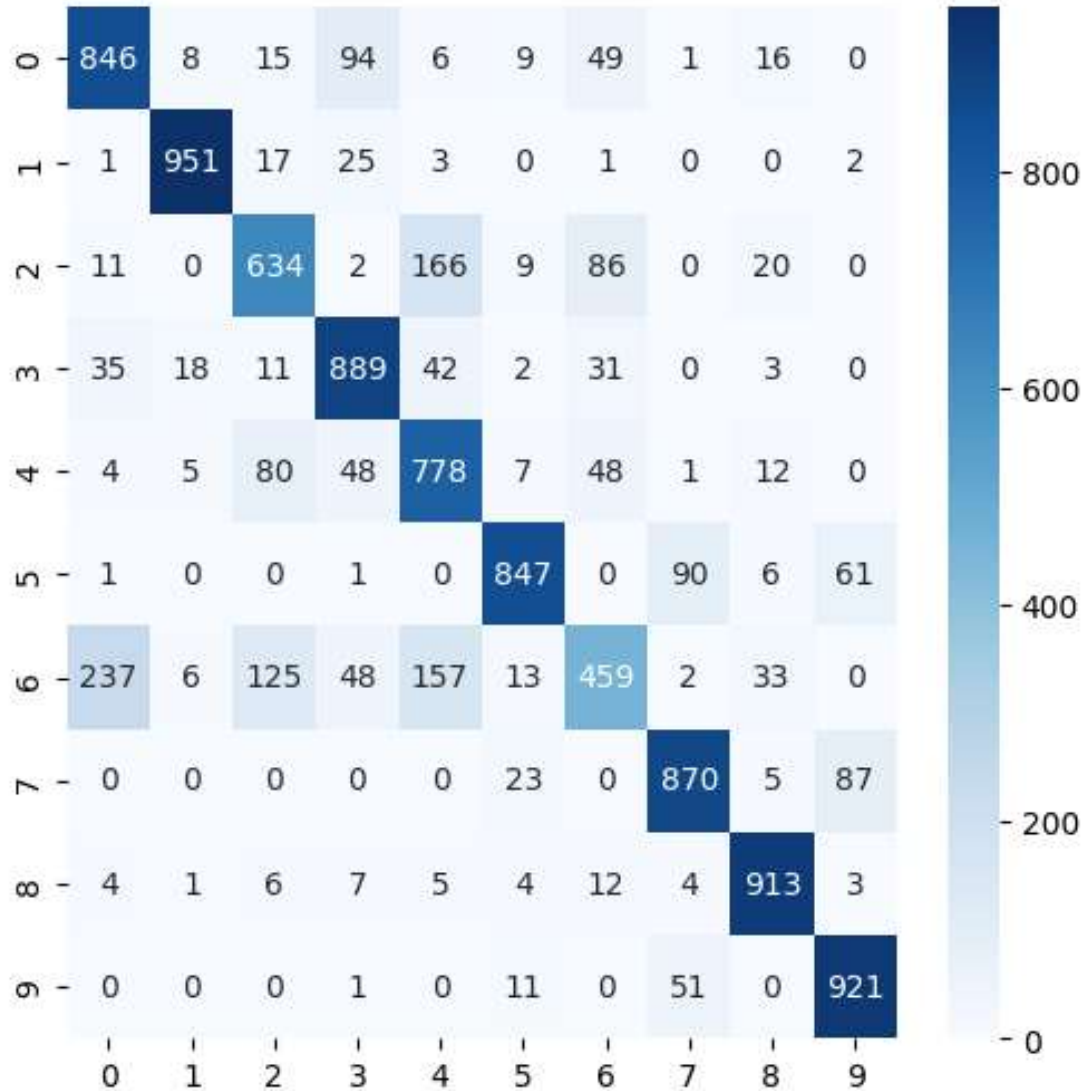
```
Epoch 1/10 - Train Loss: 191.0827, Acc: 0.6861 | Val Loss: 131.8190, Acc: 0.8108, F1: 0.8074
Early stopping triggered.
```

2. Graphs





3. Confusion Matrix

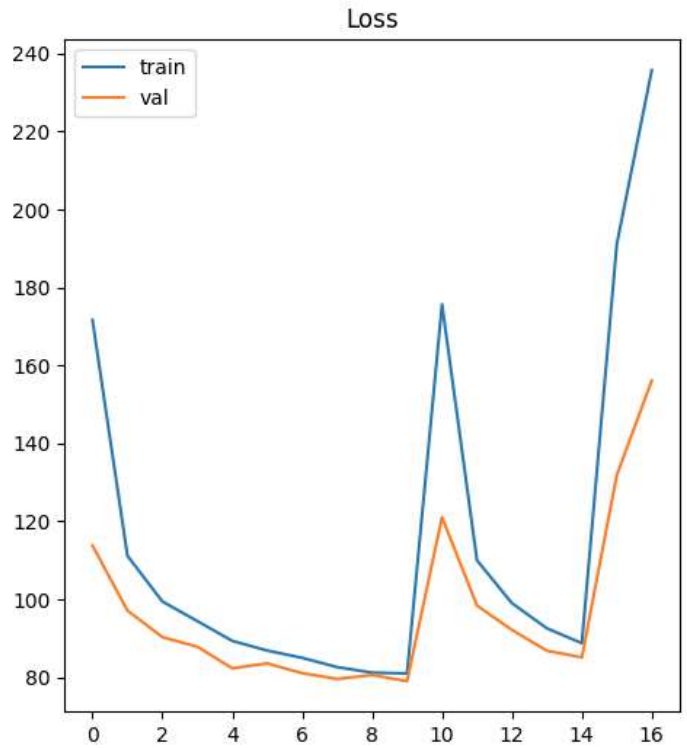


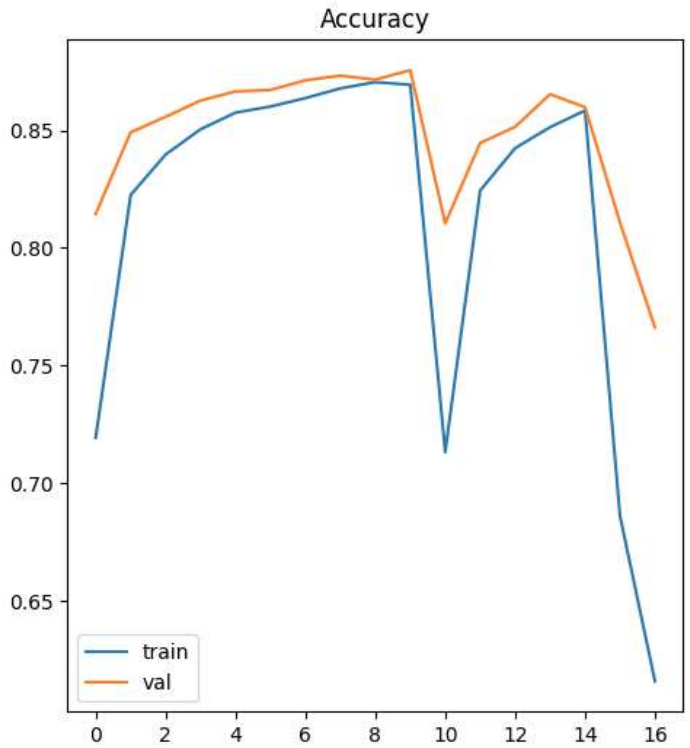
Model 3(LR = 0.000625)

1. Training loss and accuracy, Validation loss and accuracy, F1 scores:

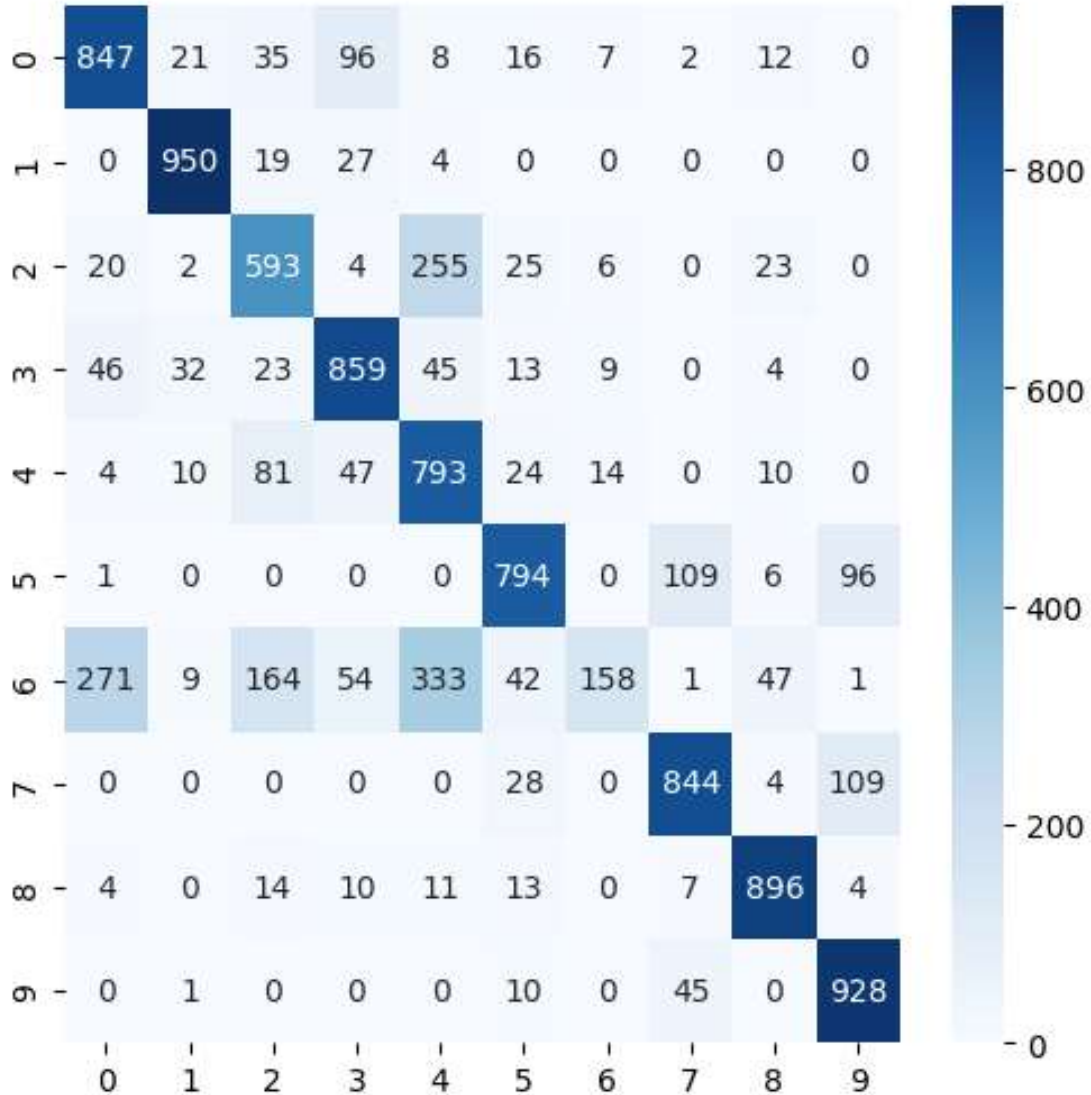
```
Epoch 1/10 - Train Loss: 235.7122, Acc: 0.6159 | Val Loss: 156.1241, Acc: 0.7662, F1: 0.7483
Early stopping triggered.
```

2. Graphs





3. Confusion Matrix



Best (chosen) model:

We have chosen the best model as model because it gave the best F1 score on validation set among all the models.

```
Best Model:  {'model_2'}
Best Model Val F1 Score:  {0.8779592893548 9}
```

Independent test performance of the best model

```
Test Accuracy: 0.8552, Test F1: 0.8555
```