ML Report(Offline 3)
Student ID : 1805027

The train losses,validation losses,train accuracy,validation accuracy and the confusion matrices are reported in the following pages for each of the 12 models that were trained.It is to be noted here that ReLU activation layers were placed after each dense layer except the last one which is followed by softmax layer.

Implemented,

->Dense Layers

->ReLU activation

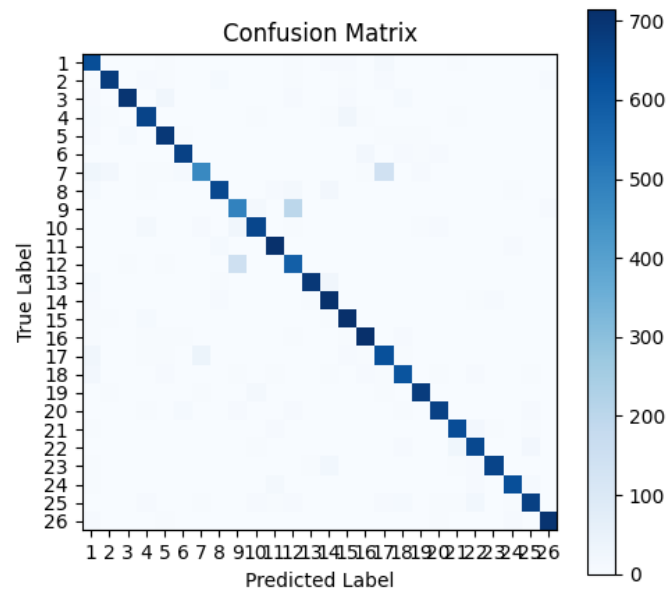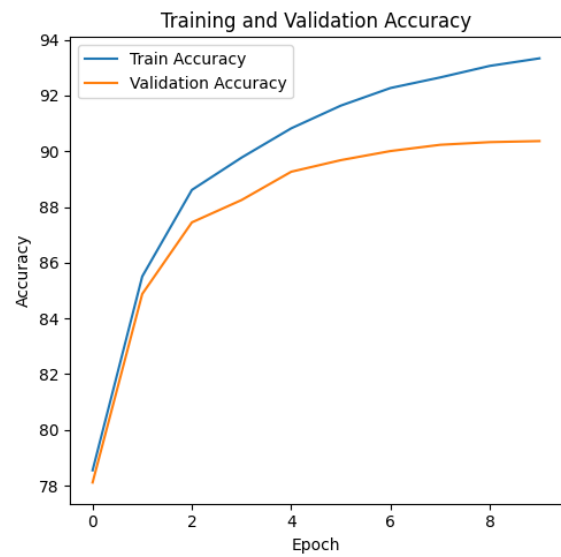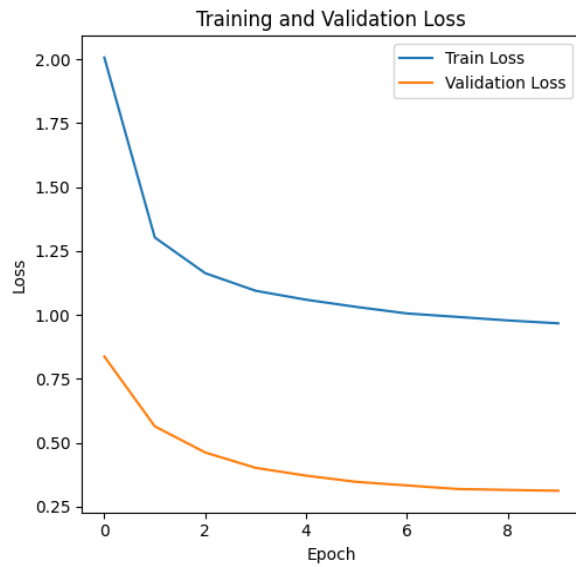->Softmax layer

->Cross entropy loss

->Adam's optimization

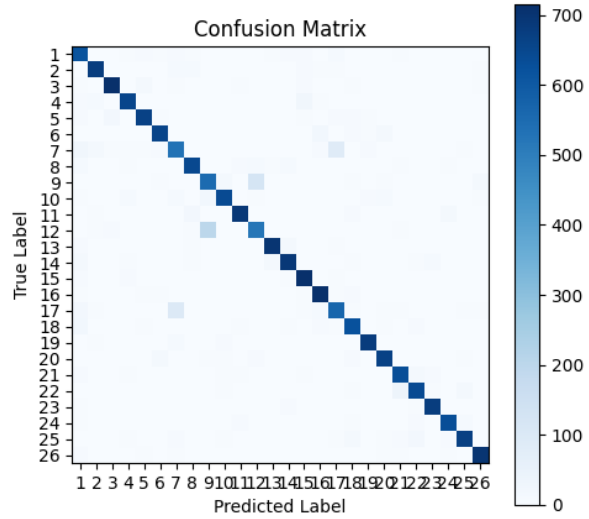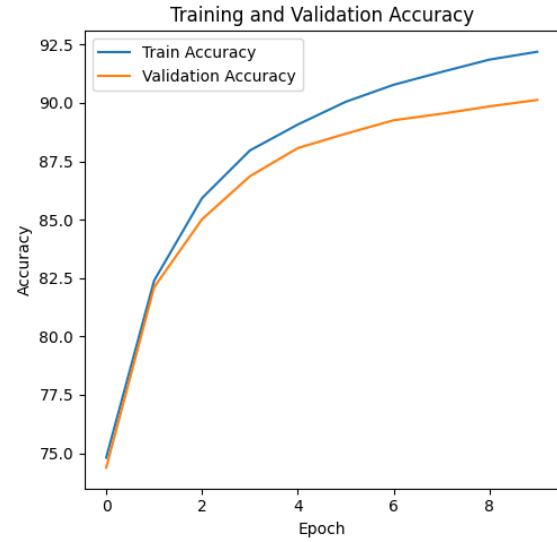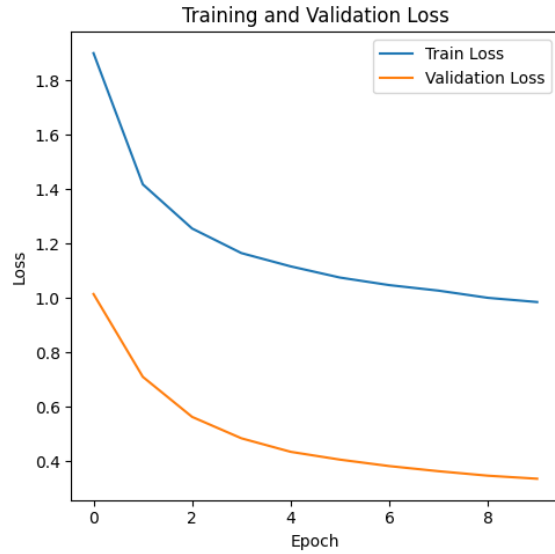->Xavier Initialization

->Mini-batch gradient descent

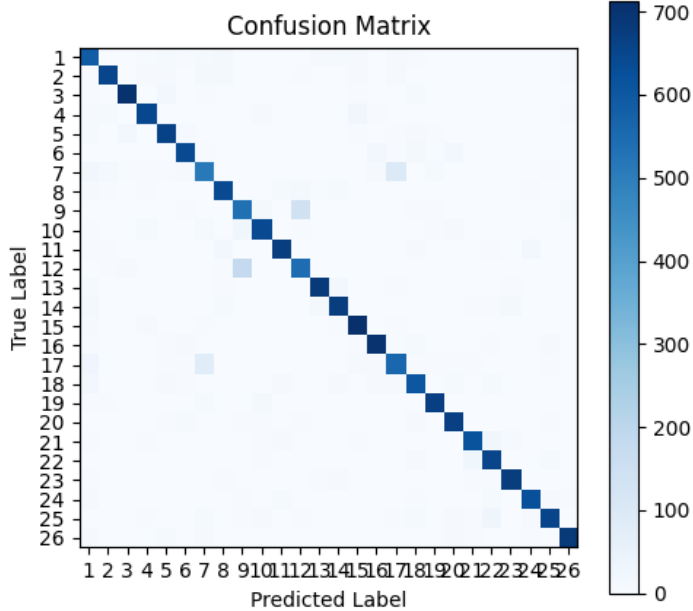Starting learning rate = 0.005
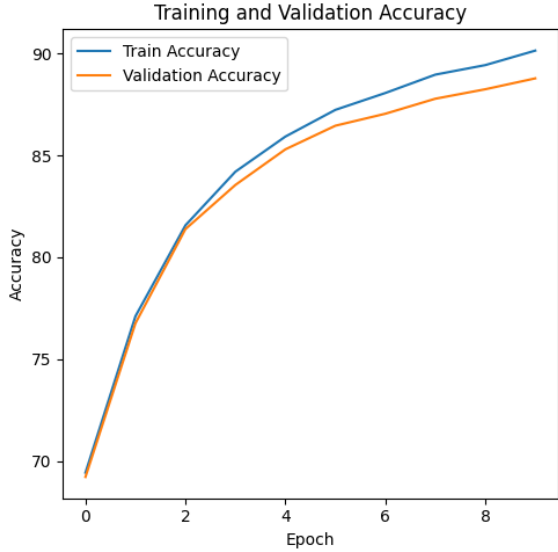
Ending learning rate = 0.0005

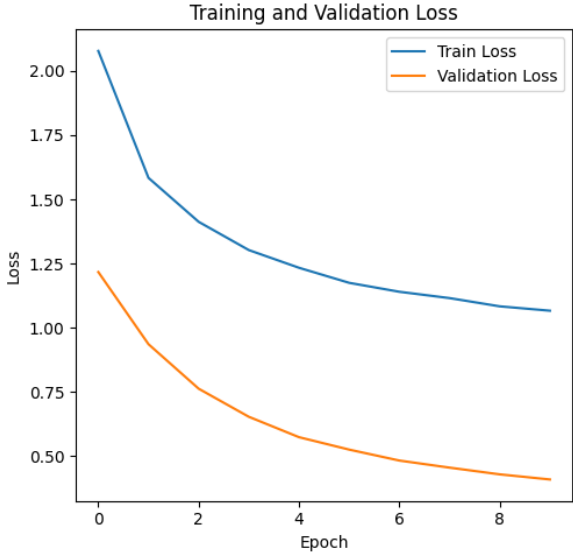## 1) Architecture : (784, 1024, 26) , Learning rate = .005, Epochs = 10

Training and Validation Loss

Training and Validation Accuracy
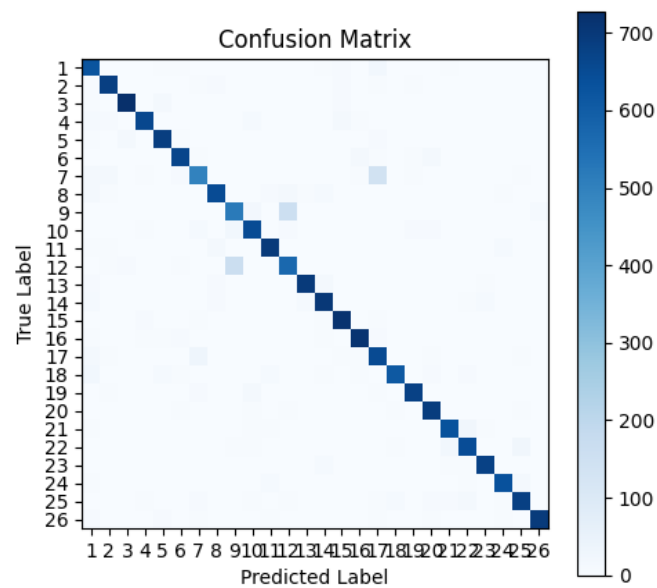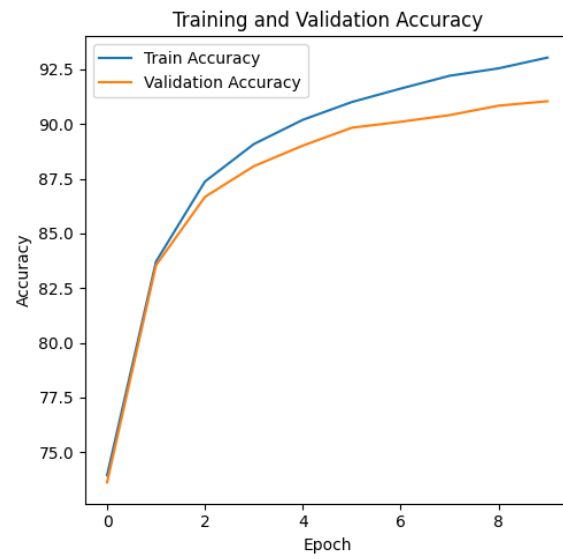

Confusion Matrix

2) Architecture : (784, 1024, 26) , Learning rate = .001, Epochs = 10

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

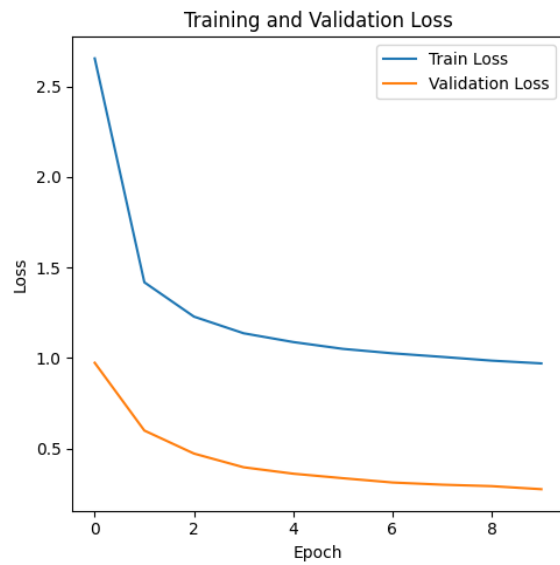### 3) Architecture : (784, 1024, 26) , Learning rate = .0005, Epochs = 10

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

4) Architecture : (784, 1024,2048, 26) , Learning rate = .005, Epochs = 10

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

5) Architecture : (784, 1024,2048, 26) , Learning rate = .001, Epochs = 10

Training and Validation Loss
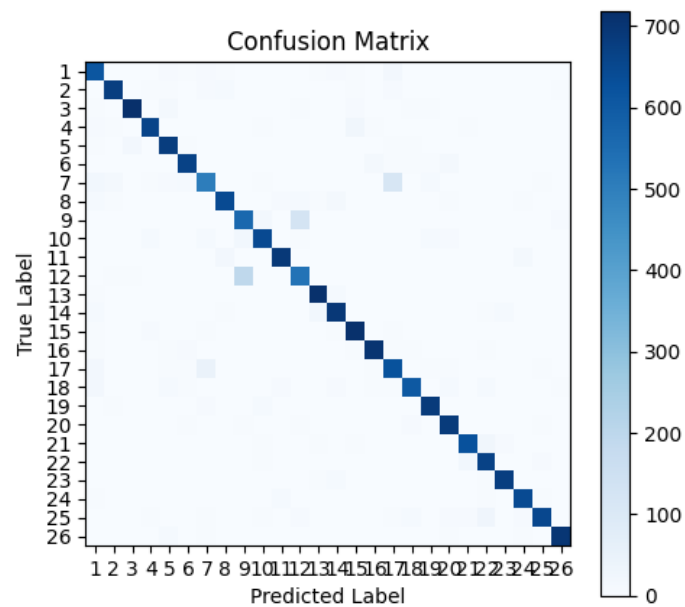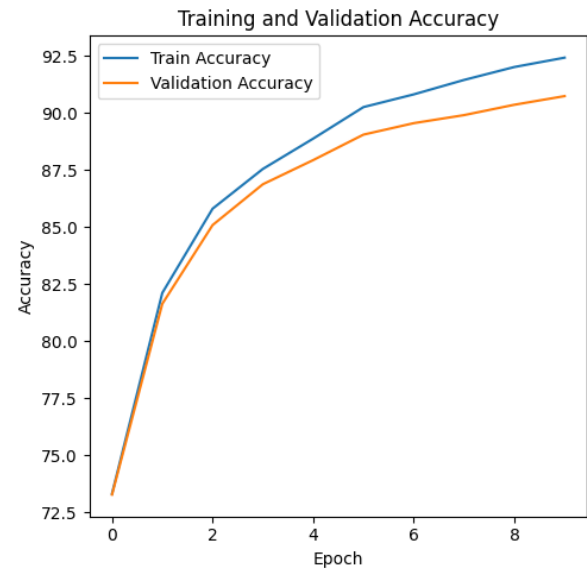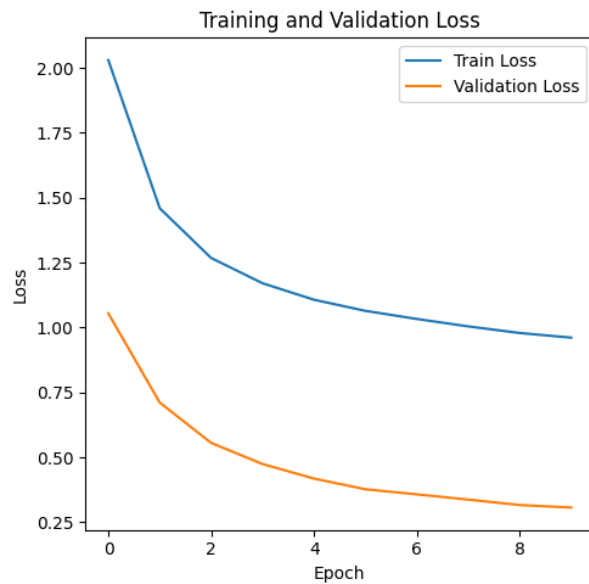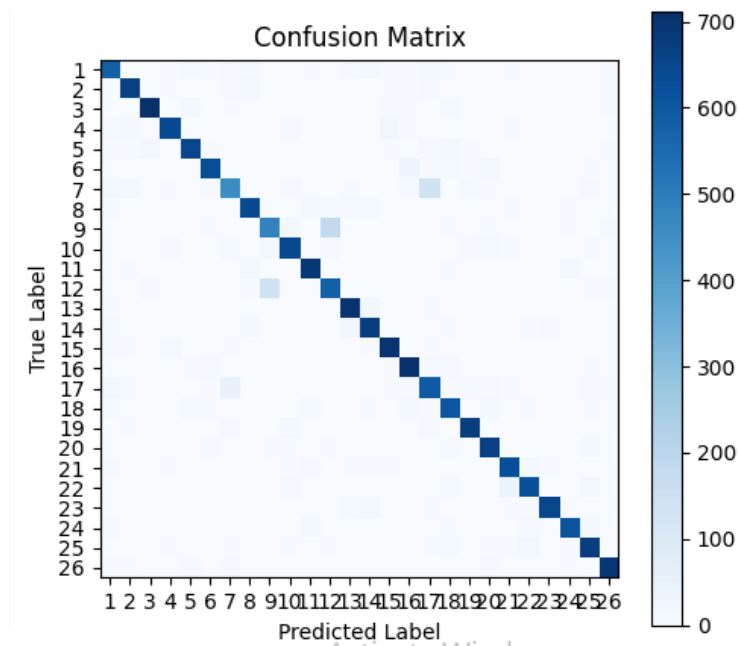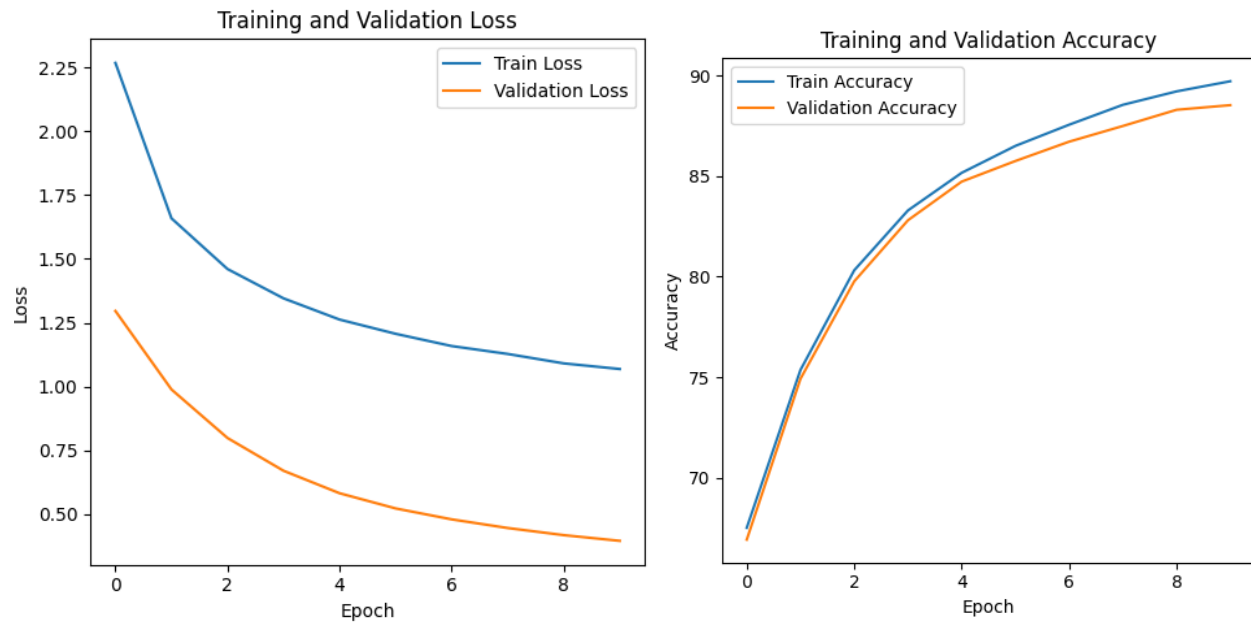
Training and Validation Accuracy

Confusion Matrix

6) Architecture : (784, 1024,2048, 26) , Learning rate = .0005, Epochs = 10

7) Architecture : (784, 1024, 26) , Learning rate = .005, Epochs = 10 [Without Dropout]
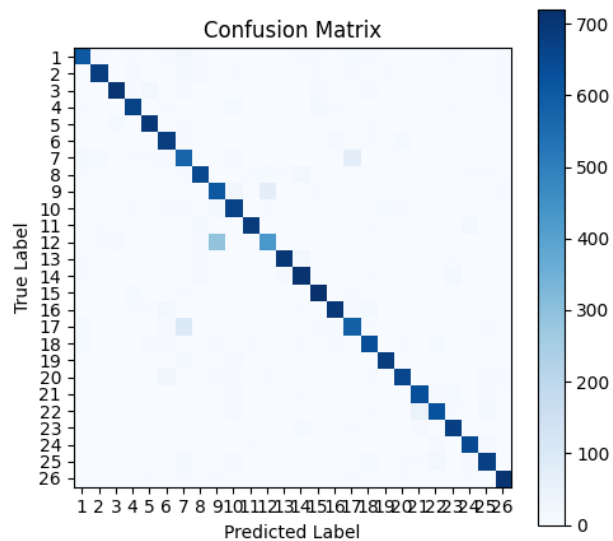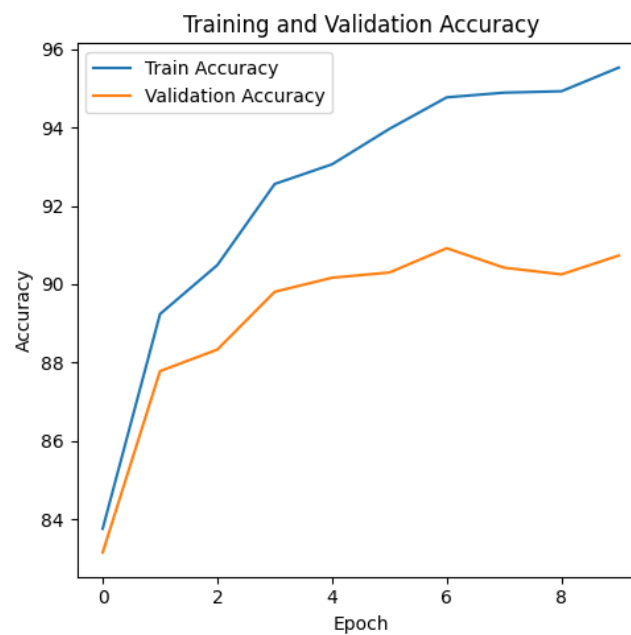

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

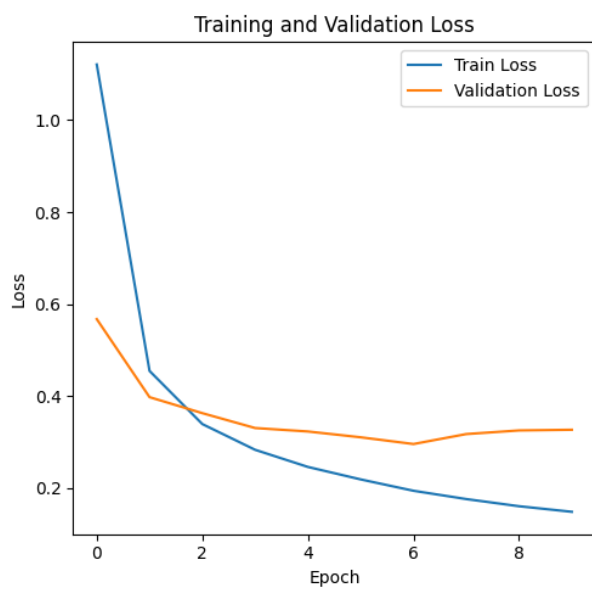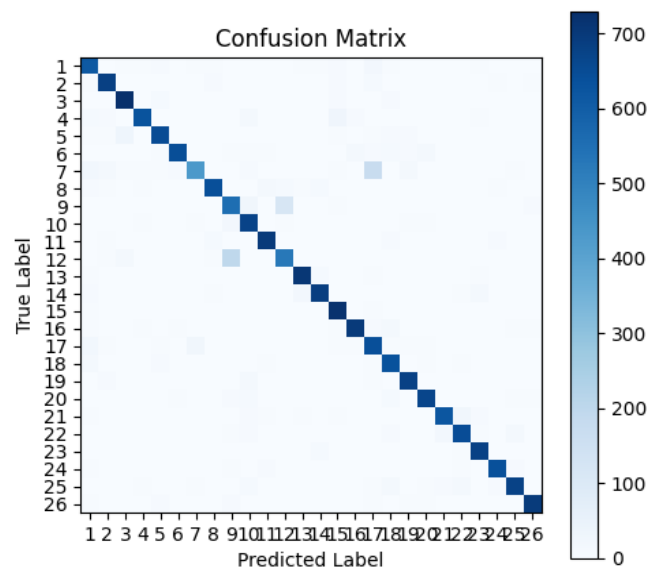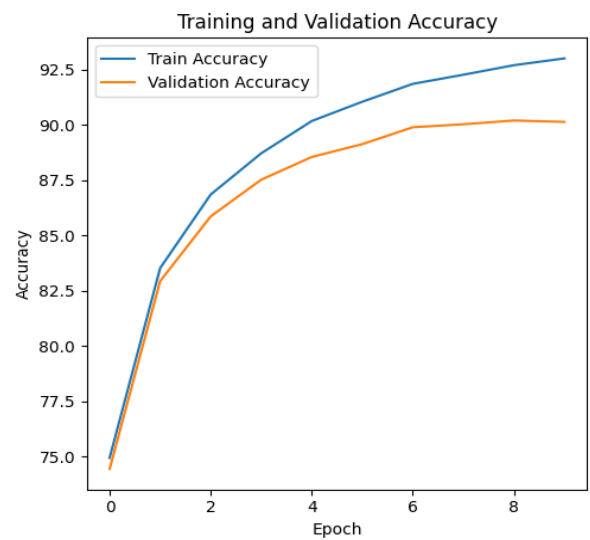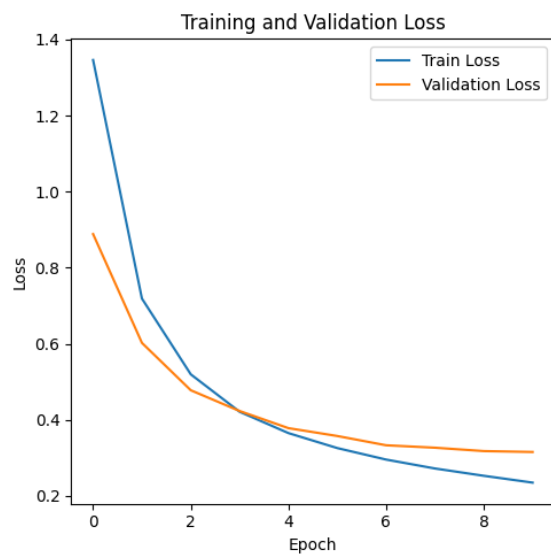**8) Architecture : (784, 1024, 26) , Learning rate = .001, Epochs = 10 [Without Dropout]**

Training and Validation Loss
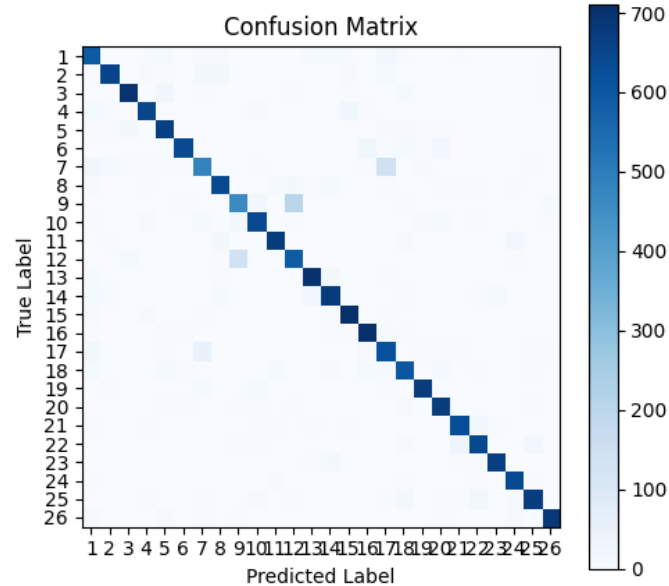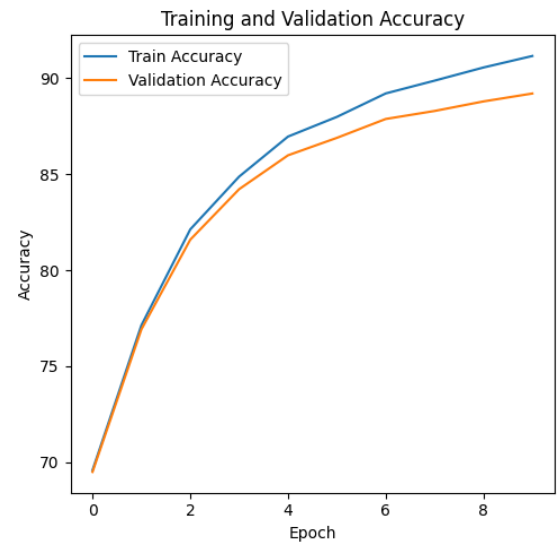

Training and Validation Accuracy

Confusion Matrix

9) Architecture : (784, 1024, 26) , Learning rate = .0005, Epochs = 10 [Without Dropout]

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

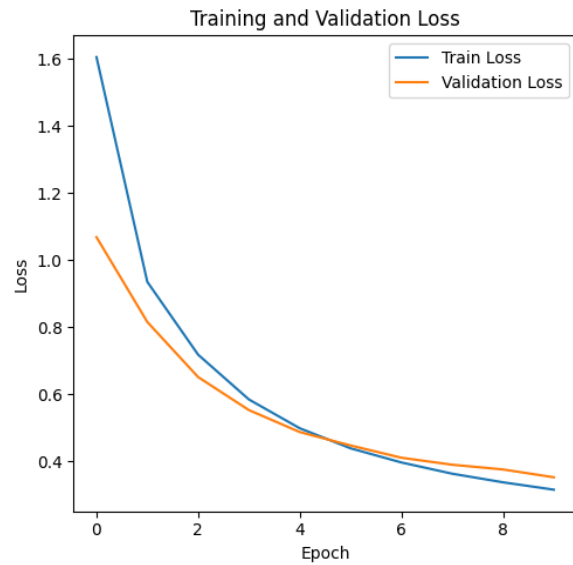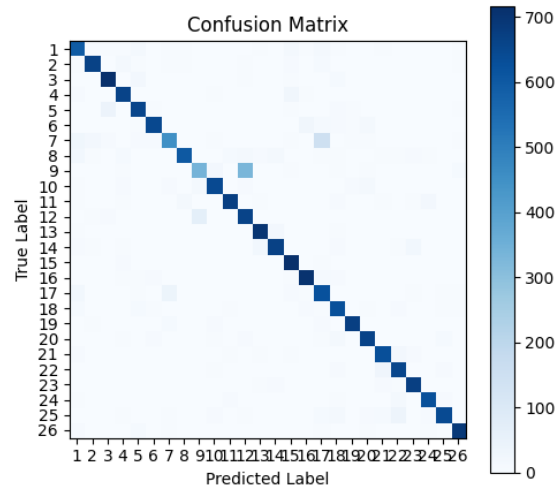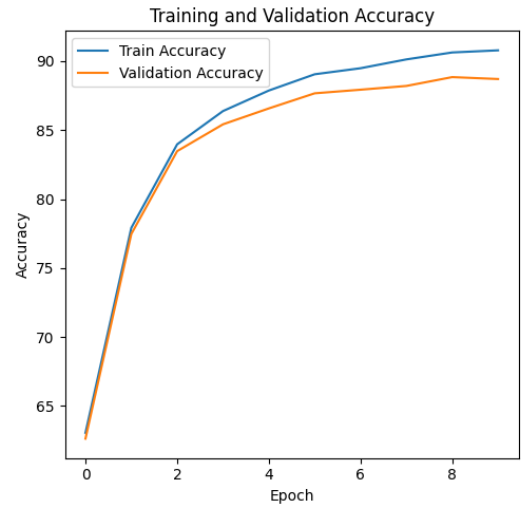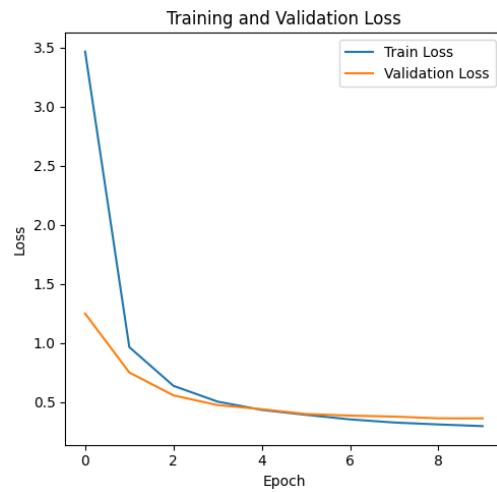**10) Architecture : (784, 1024, 2048, 26) , Learning rate = .005, Epochs = 10 [Without Dropout]**

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

11) Architecture : (784, 1024, 2048, 26) , Learning rate = .001, Epochs = 10 [Without Dropout]

Training and Validation Loss

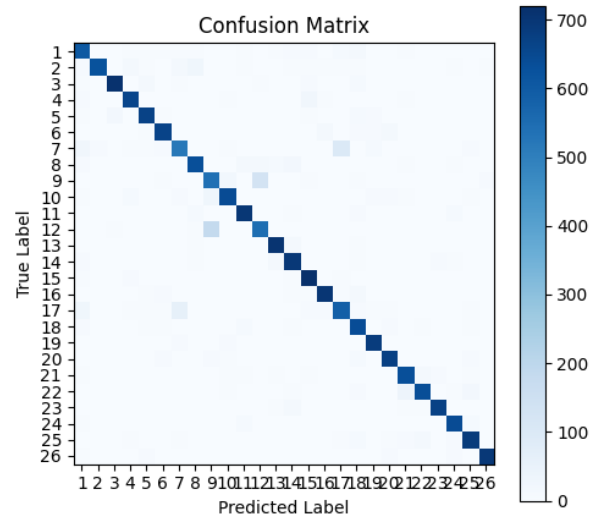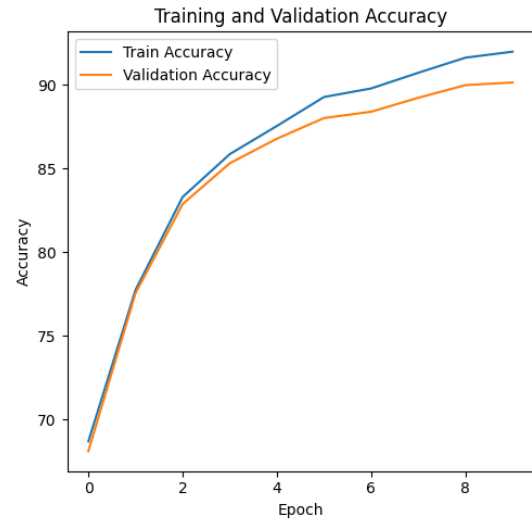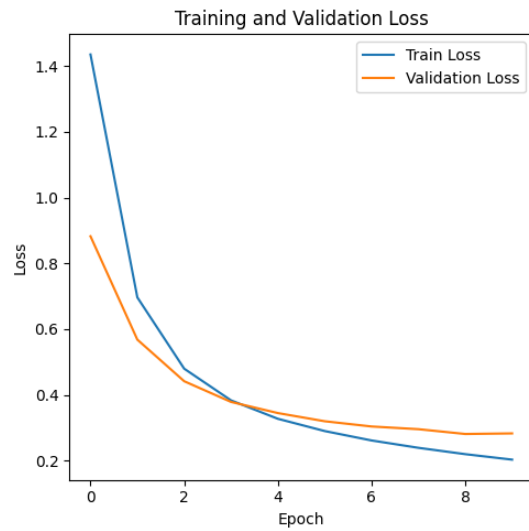Training and Validation Accuracy

Confusion Matrix

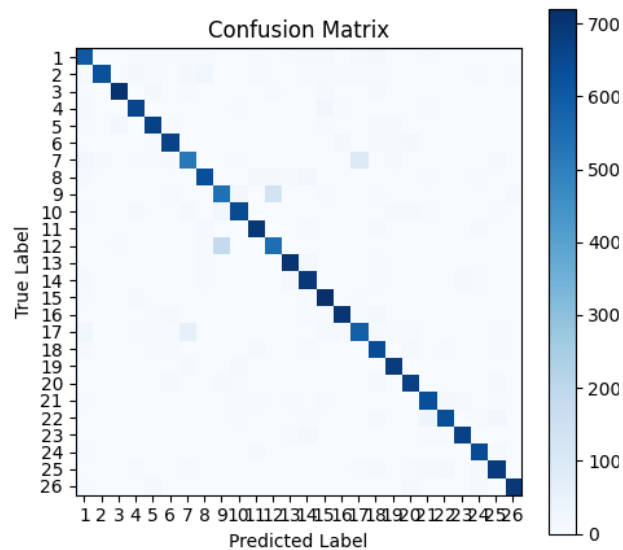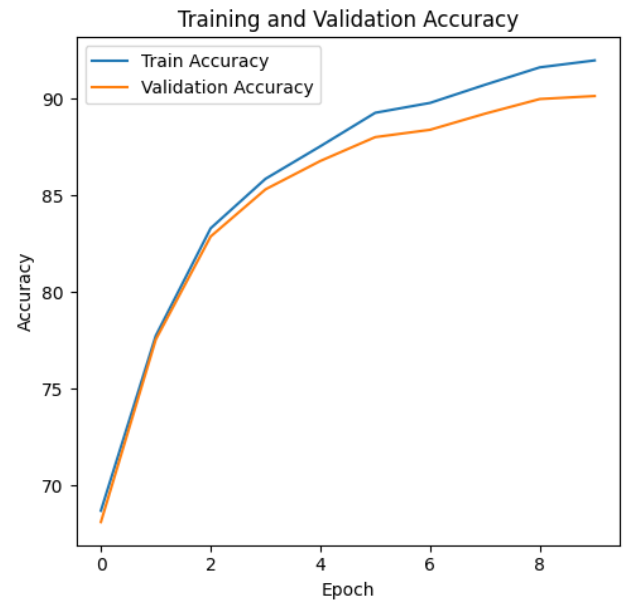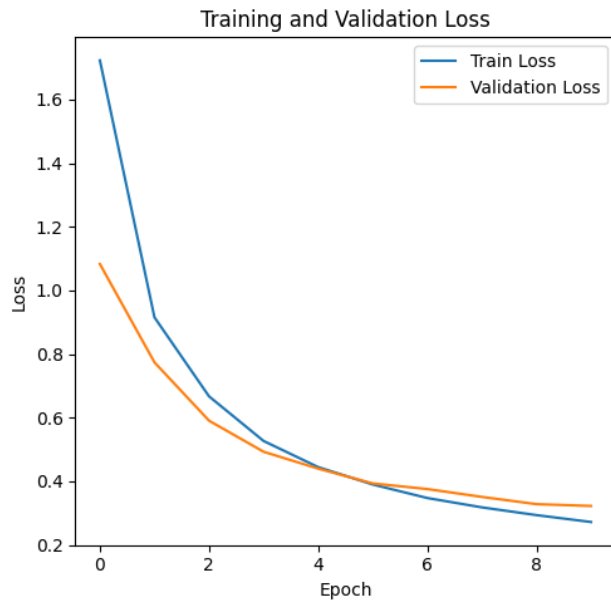**12) Architecture : (784, 1024, 2048, 26) , Learning rate = .0005, Epochs = 10 [Without Dropout]**

Training and Validation Loss

Training and Validation Accuracy

Confusion Matrix

Comments:

The best model turned out to be the **4th one (784,102,2048,26),LR=0.005,**

```
Test accuracy:  91.17788461538461
Test F1:  91.17059700094961
```

It gave a test accuracy of 91.18 and f1 score of 91.17 respectively.