

# CSE472 : Machine Learning

## Offline 3 : Feed-forward Neural Network

Date: 3 January 2024

Submitted by:  
Kazi Ababil Azam  
Student No: 1805077  
CSE, L-4/T-2, B1

## Hyperparameters:

1. Epochs
2. Batch Size
3. Learning Rate
4. Adam Optimizer:
  - a. Beta1: 0.9
  - b. Beta2: 0.999
  - c. Epsilon: 1E-8

## Model 1:

### Architecture:

DenseLayer(784, 512),  
ReLU(),  
DropoutLayer(0.5),  
DenseLayer(512, 128),  
ReLU(),  
DropoutLayer(0.25),  
DenseLayer(128, 64),  
ReLU(),  
DenseLayer(64, 26),  
Softmax()

### Hyperparameters:

Epochs: 25

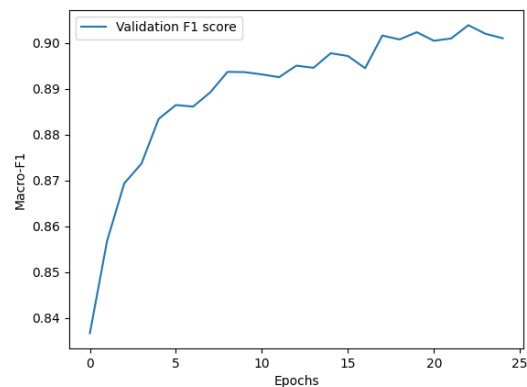
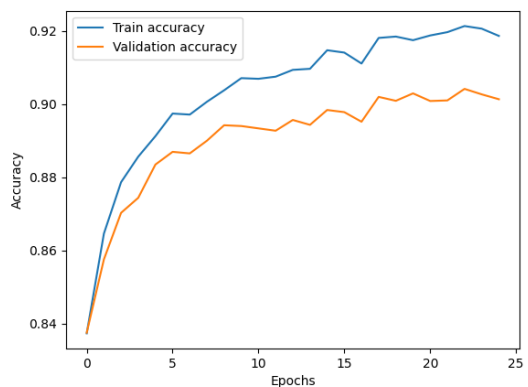
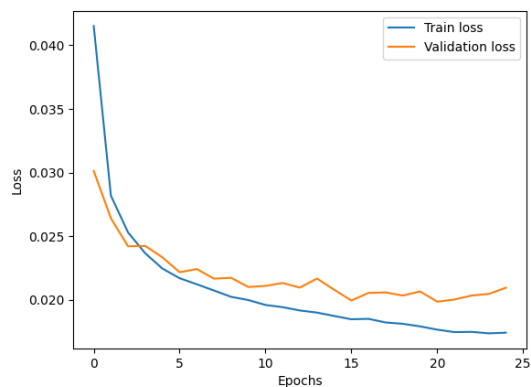
Batch Size: 256

*Learning rate: 0.005*

Training Accuracy: 0.9186368778280543

Validation Accuracy: 0.9013354700854701

**Validation F1 score: 0.900996605092391**

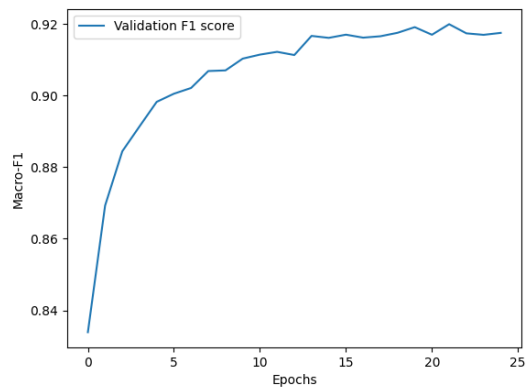
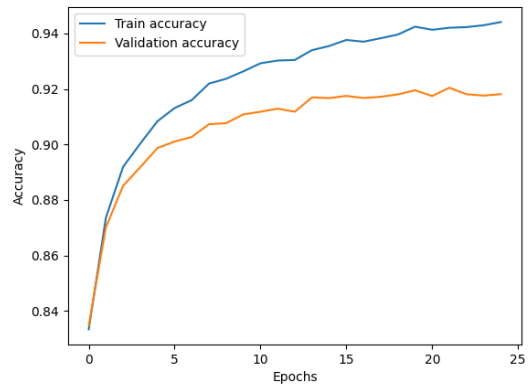
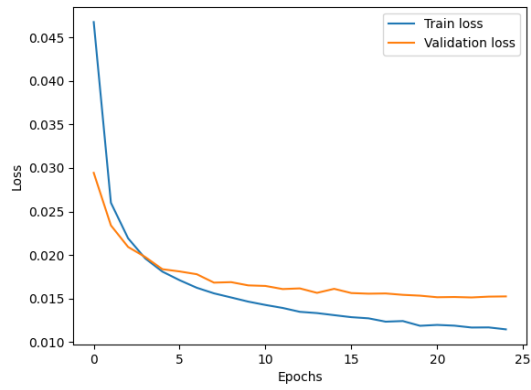


*Learning rate: 0.001*

Training Accuracy: 0.9441365007541478

Validation Accuracy: 0.9181623931623931

**Validation F1 score: 0.917484849884212**

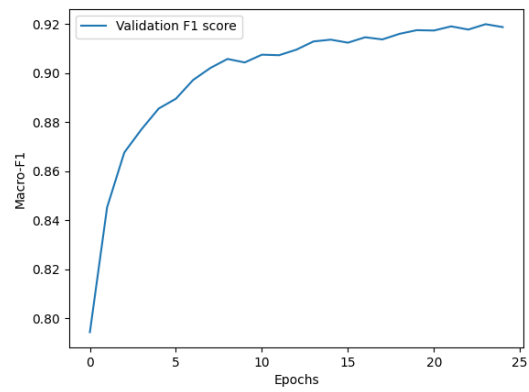
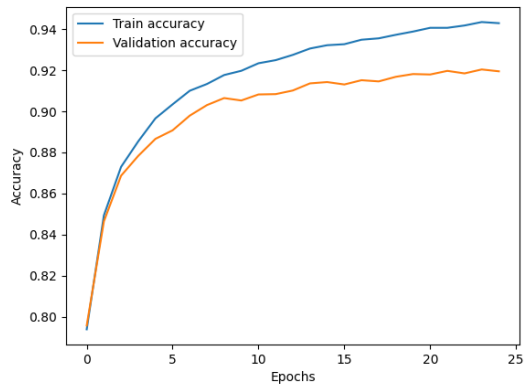
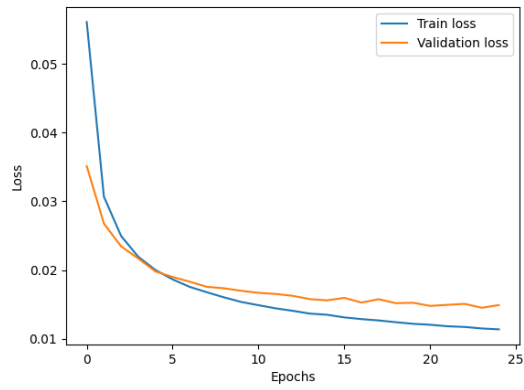


*Learning rate: 0.0005*

Training Accuracy: 0.9429487179487179

Validation Accuracy: 0.9194978632478632

**Validation F1 score: 0.9187430319379778 (Best F1 Score)**

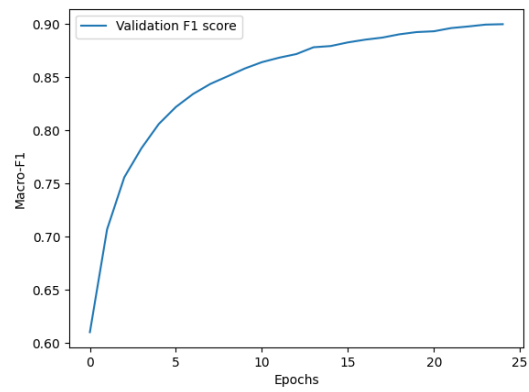
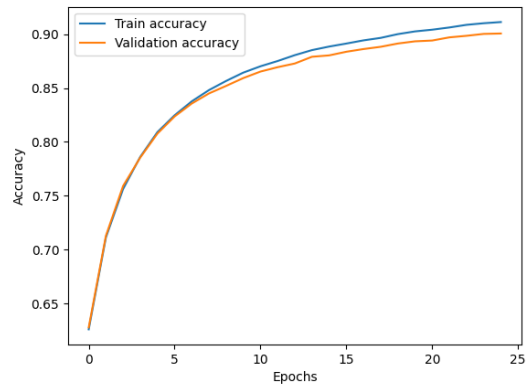
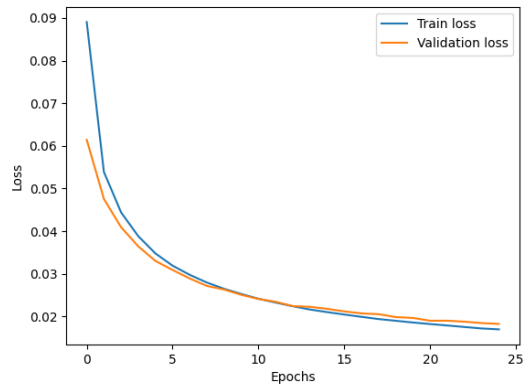


*Learning rate: 0.0001*

Training Accuracy: 0.9112085218702866

Validation Accuracy: 0.9005876068376069

**Validation F1 score: 0.9000456710896209**



## Model 2:

### Architecture:

DenseLayer(784, 128),  
ReLU(),  
DropoutLayer(0.5),  
DenseLayer(128, 32),  
ReLU(),  
DenseLayer(32, 26),  
Softmax()

### Hyperparameters:

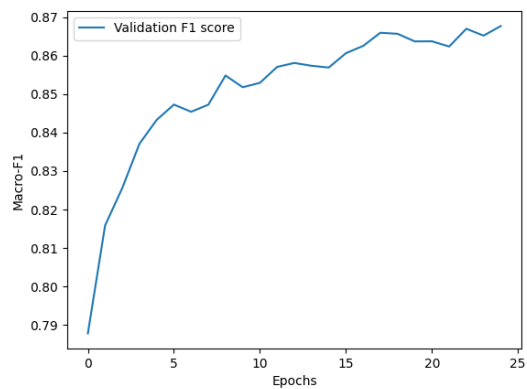
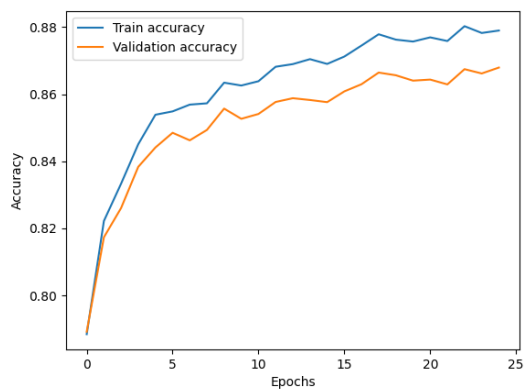
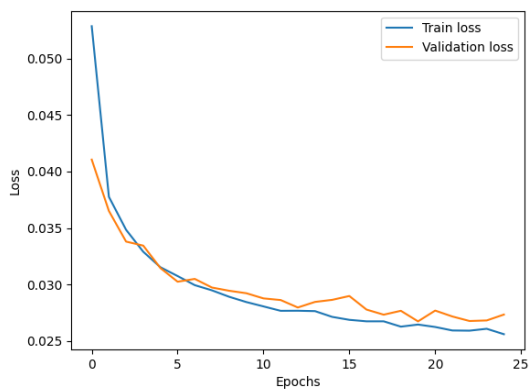
Epochs: 25  
Batch Size: 256

*Learning rate: 0.005*

Training Accuracy: 0.878987556561086

Validation Accuracy: 0.867948717948718

**Validation F1 score: 0.8676702518628798**

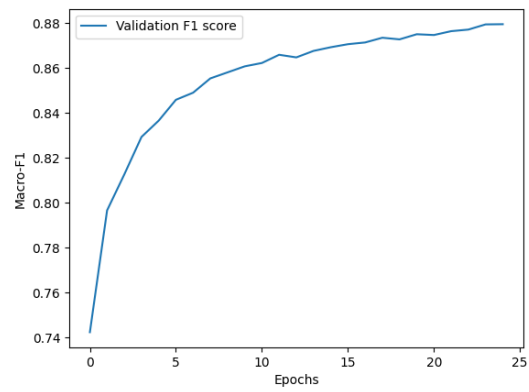
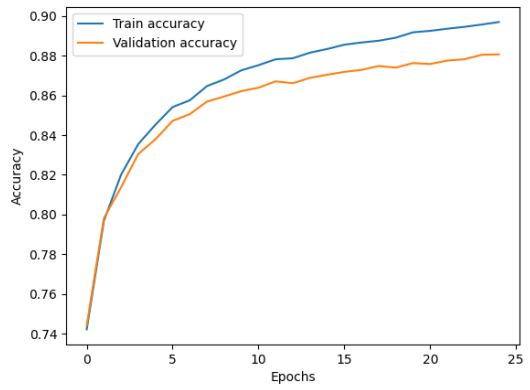
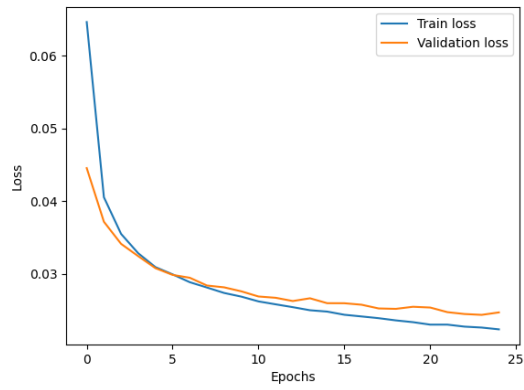


*Learning rate: 0.001*

Training Accuracy: 0.8968608597285068

Validation Accuracy: 0.8806089743589743

**Validation F1 score: 0.8795661373558642**



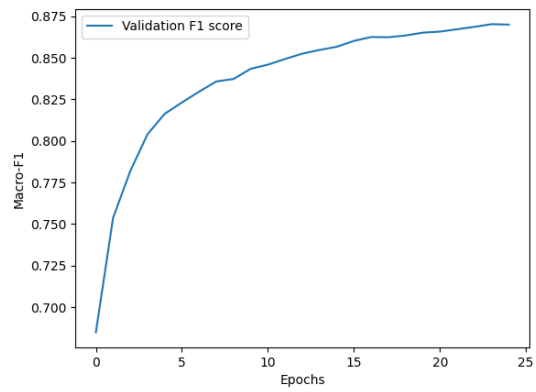
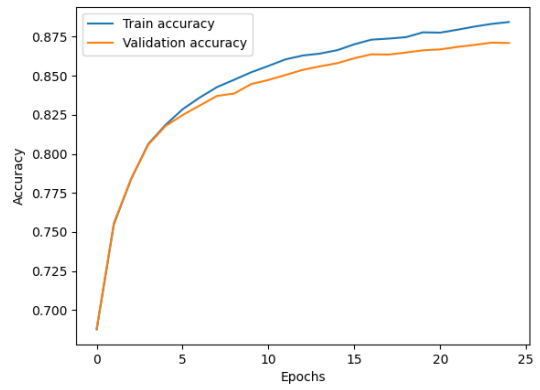
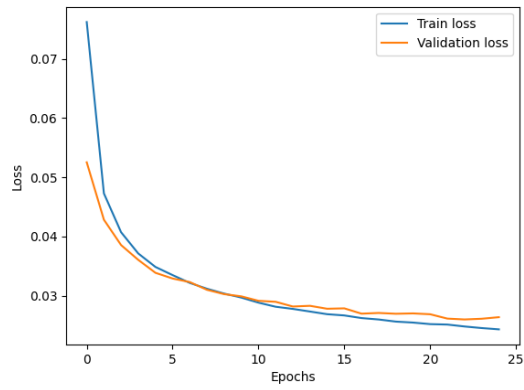


*Learning rate: 0.0005*

Training Accuracy: 0.8843891402714932

Validation Accuracy: 0.8709935897435898

**Validation F1 score: 0.8700275575123337**

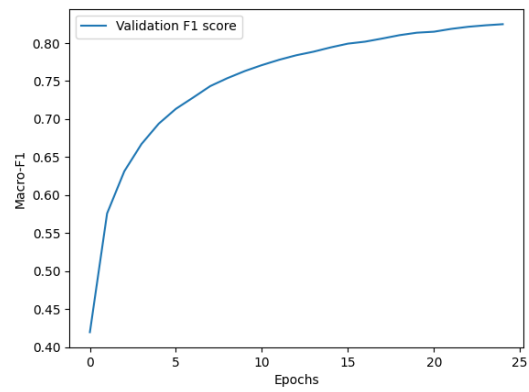
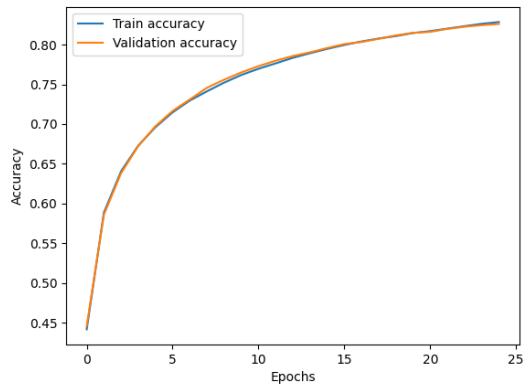
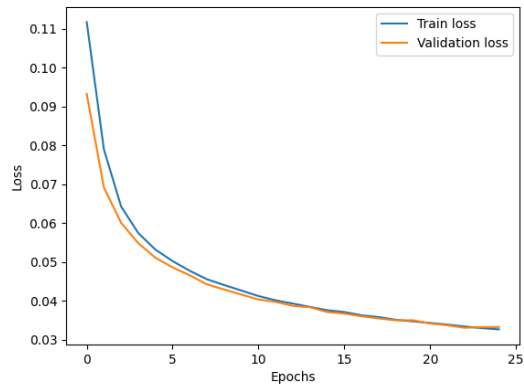


*Learning rate: 0.0001*

Training Accuracy: 0.8284879336349925

Validation Accuracy: 0.826068376068376

**Validation F1 score: 0.8249093780693262**



## Model 3:

### Architecture:

DenseLayer(784, 128),  
ReLU(),  
DropoutLayer(0.5),  
DenseLayer(128, 128),  
ReLU(),  
DropoutLayer(0.5),  
DenseLayer(128, 64),  
ReLU(),  
DropoutLayer(0.5),  
DenseLayer(64, 26),  
Softmax()

### Hyperparameters:

Epochs: 25

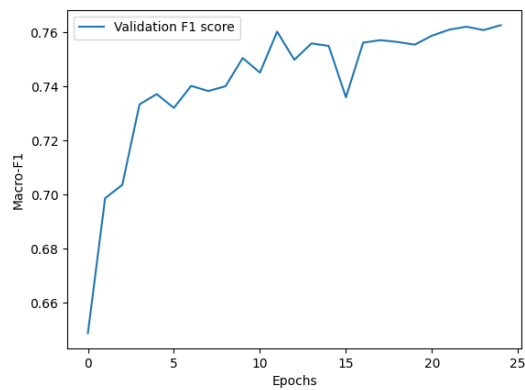
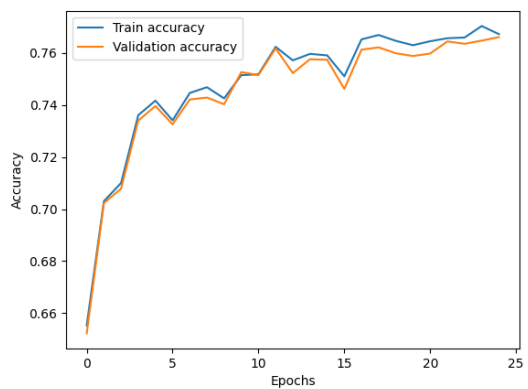
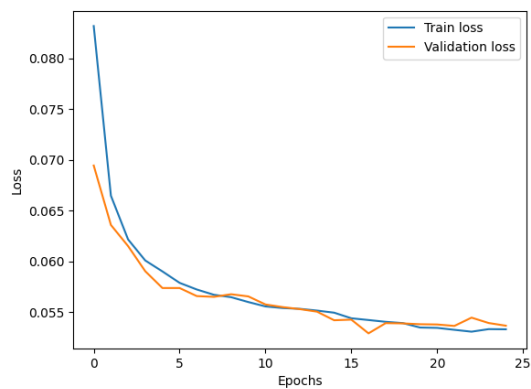
Batch Size: 256

*Learning rate: 0.005*

Training Accuracy: 0.7672888386123681

Validation Accuracy: 0.7661324786324787

**Validation F1 score: 0.7625766354029706**

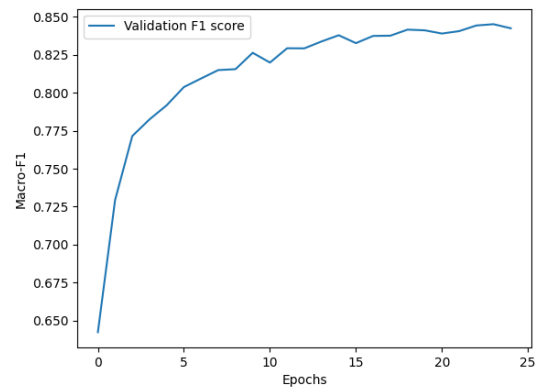
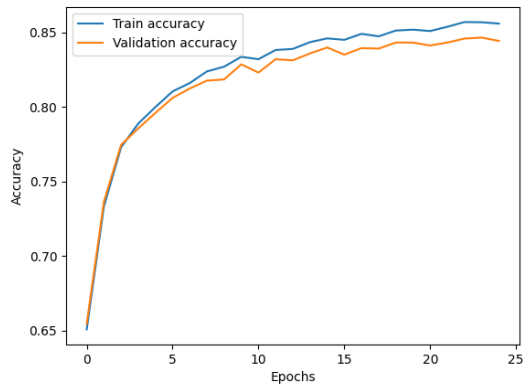
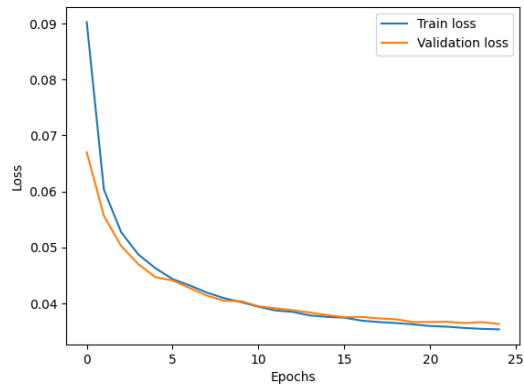


*Learning rate: 0.001*

Training Accuracy: 0.8558257918552036

Validation Accuracy: 0.8441773504273504

**Validation F1 score: 0.842437551432689**

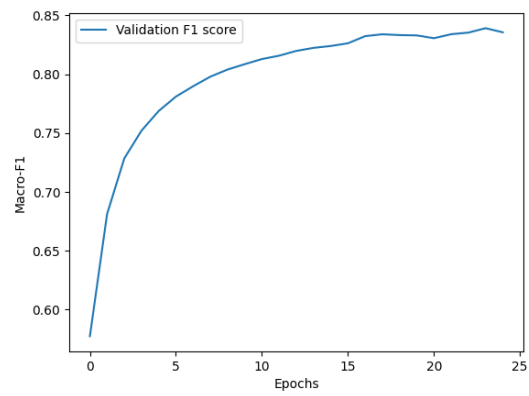
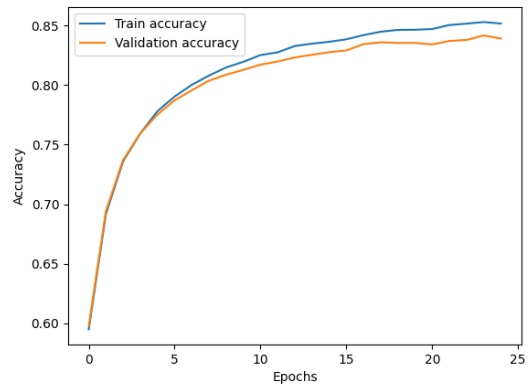
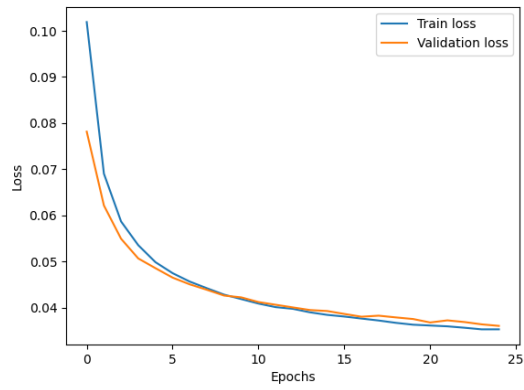


*Learning rate: 0.0005*

Training Accuracy: 0.8517816742081448

Validation Accuracy: 0.8391559829059829

**Validation F1 score: 0.8356123920580967**

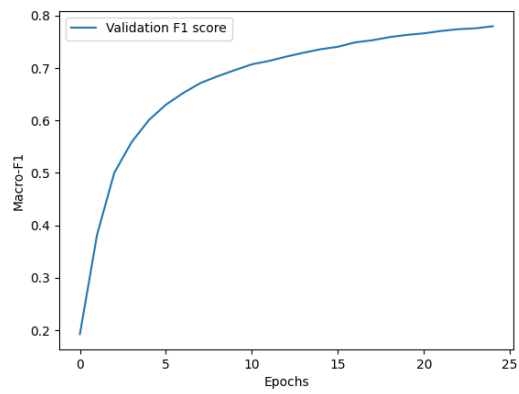
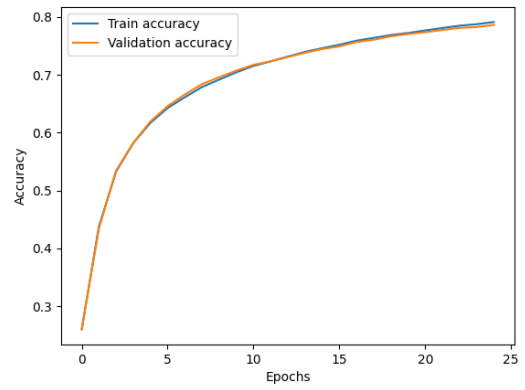
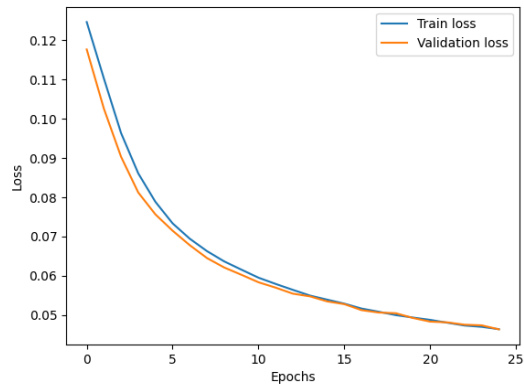


*Learning rate: 0.0001*

Training Accuracy: 0.7911953242835595

Validation Accuracy: 0.7862179487179487

**Validation F1 score: 0.7798927919642318**



Best F1 score:

Model 1 with learning rate 0.0005

Validation F1 score: 0.9187430319379778

Independent Test Dataset Results:

Accuracy: 0.9152403846153846

F1 score: 0.9148286294239791

Confusion Matrix:

