Feed Forward Neural Network from Scratch

Report on Assignment 3

Md. Asif Haider
1805112

Hyperparameters for Experiment:

1. **Learning Rate:** [5e-3, 1e-3, 5e-4]
    a. Starting from these values with a gradually decaying linear LR scheduler
2. **Dense Layer Units:** [256, 512, 1024]
3. **Dropout Rate:** [No, 0.1, 0.2, 0.3]
4. **Mini Batch Size:** [128, 512, 1024]

Other Hyperparameters:

5. **Epoch:** 100
6. **Adam Optimizer:** Default Parameters (beta1 = 0.9, beta2 = 0.999)
7. **He Initialization**
8. **Overflow and Error Handling:** 1e-8
9. **Random Seed:** 112

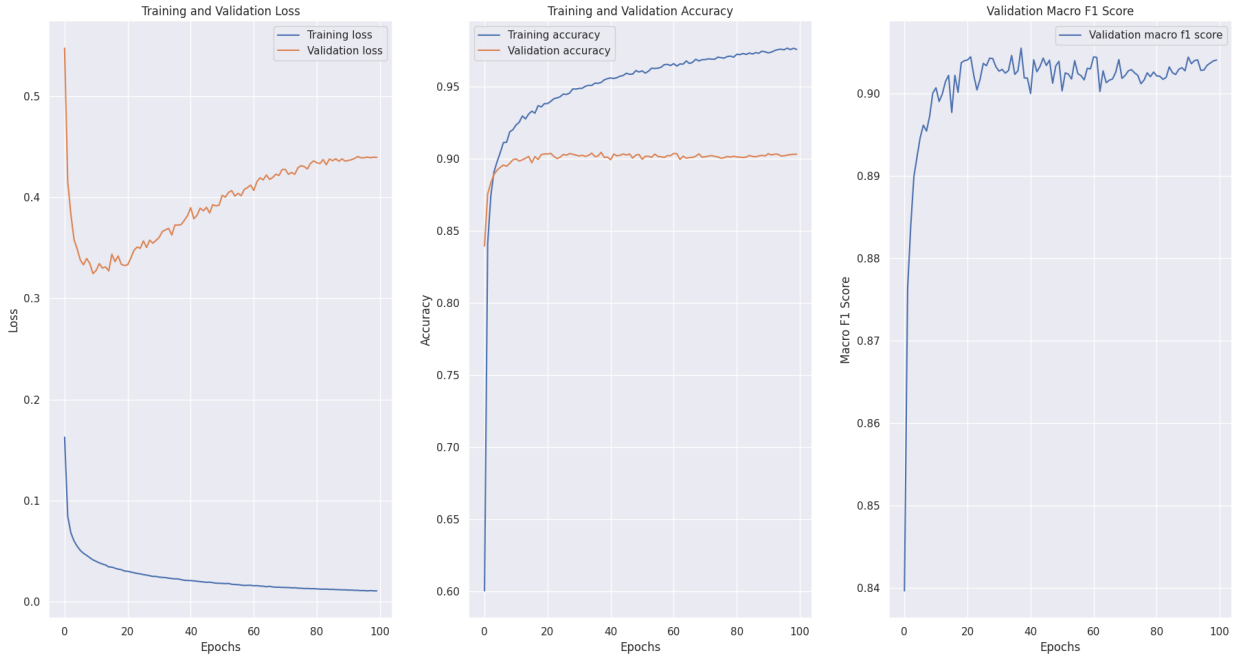
Model 1:

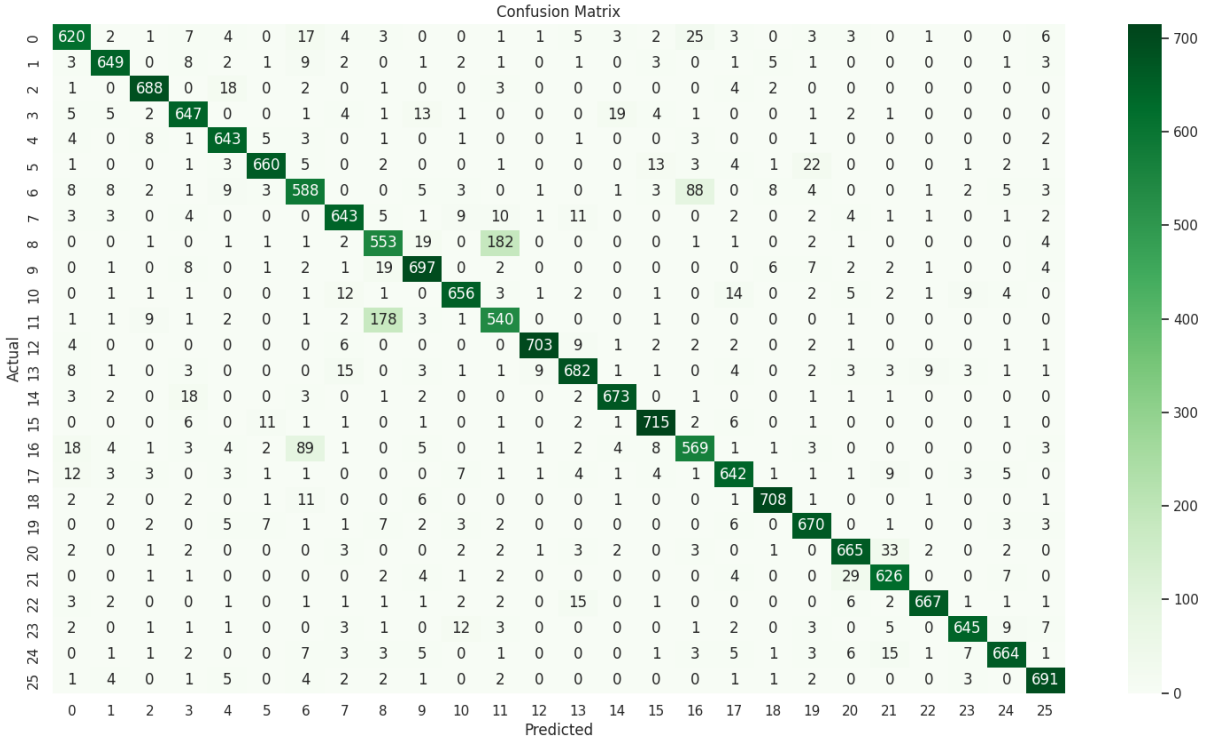
```
# define the network
dense1 = Dense(784, 256)
dense2 = Dense(256, 26)
relu1 = ReLU()
softmax = Softmax()
dropout = Dropout(0.1)

network = [
    dense1,
    relu1,
    dropout,
    dense2,
    softmax
]
```

LR: 5e-3

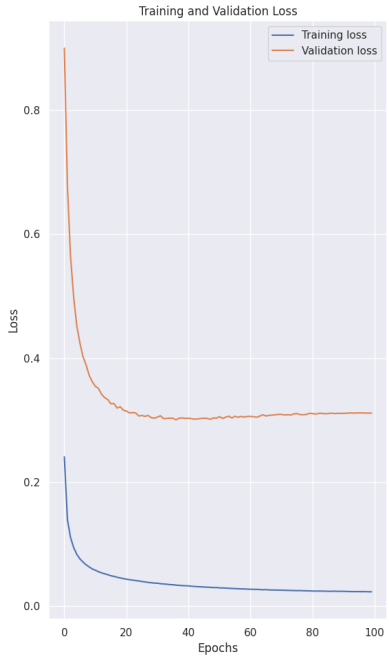
Training Loss: 0.01086
Training Accuracy: 97.6%
Validation Loss: 0.439
Validation Accuracy: 90.3%
Validation Macro F1 Score: 0.904

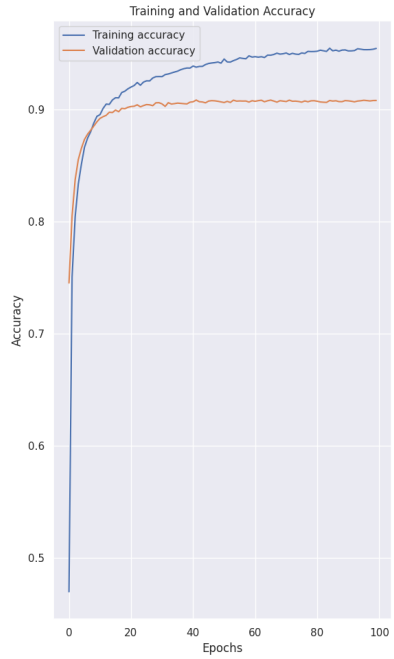

Training and Validation Loss
Training loss
Validation loss
Loss
Epochs
Training and Validation Accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs
Validation Macro F1 Score
Validation macro f1 score
Macro F1 Score
Epochs

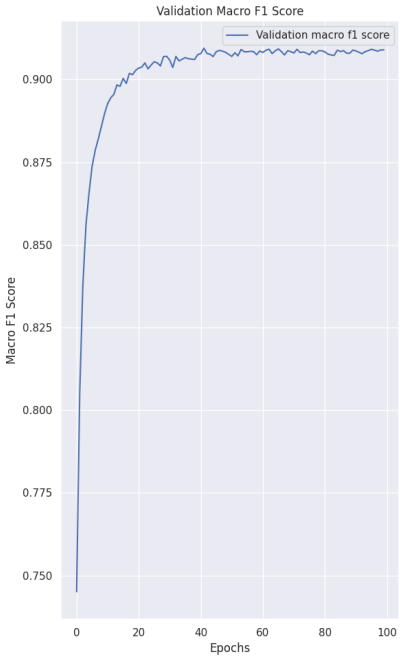

Confusion Matrix
Actual
Predicted

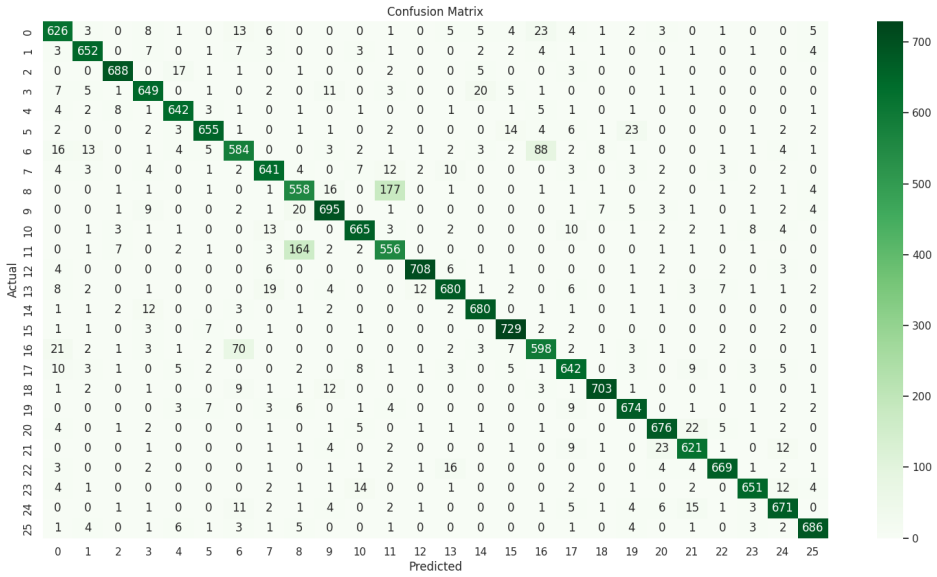
LR: 1e-3

Training Loss: 0.02292
Training Accuracy: 95.4%
Validation Loss: 0.311
Validation Accuracy: 90.8%
Validation Macro F1 Score: **0.909**



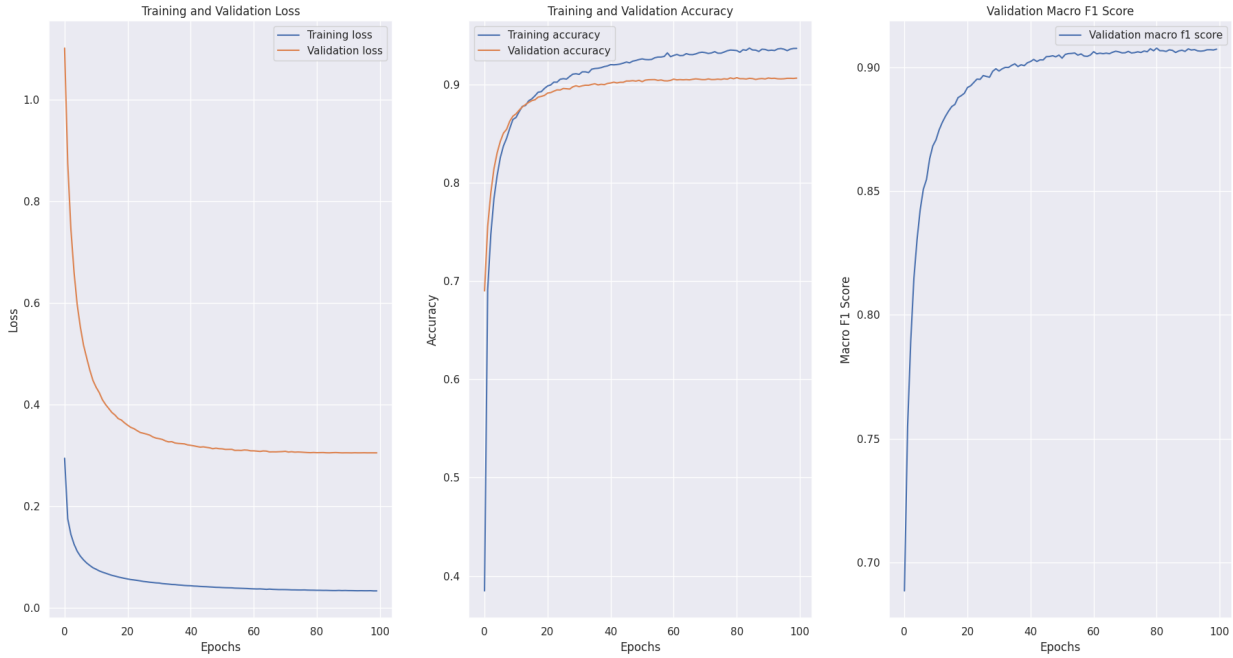


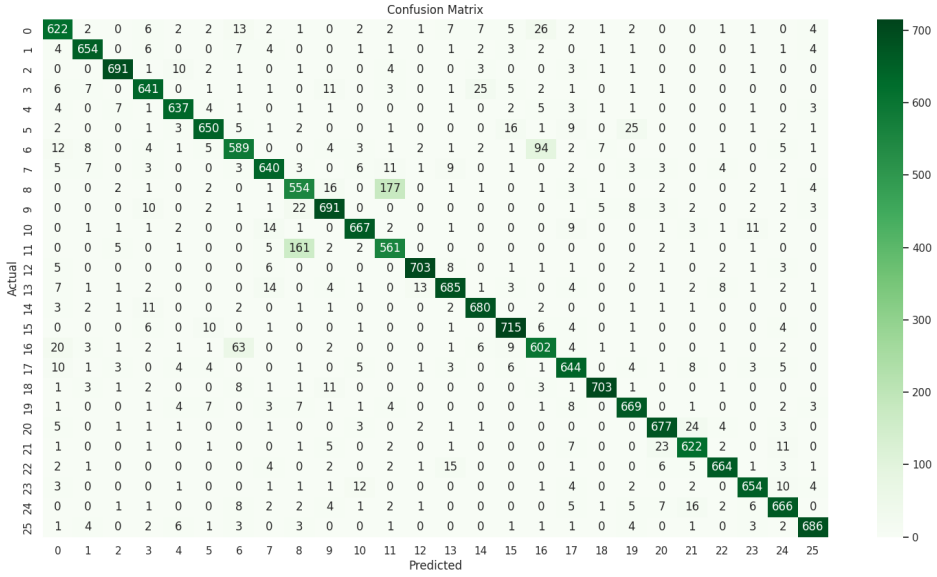




LR: 5e-4

Training Loss: 0.0334
Training Accuracy: 93.7%
Validation Loss: 0.305
Validation Accuracy: 90.6%
Validation Macro F1 Score: 0.907


Training and Validation Loss
Training loss
Validation loss
Loss
Epochs
Training and Validation Accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs
Validation Macro F1 Score
Validation macro f1 score
Macro F1 Score
Epochs


Confusion Matrix
Actual
Predicted

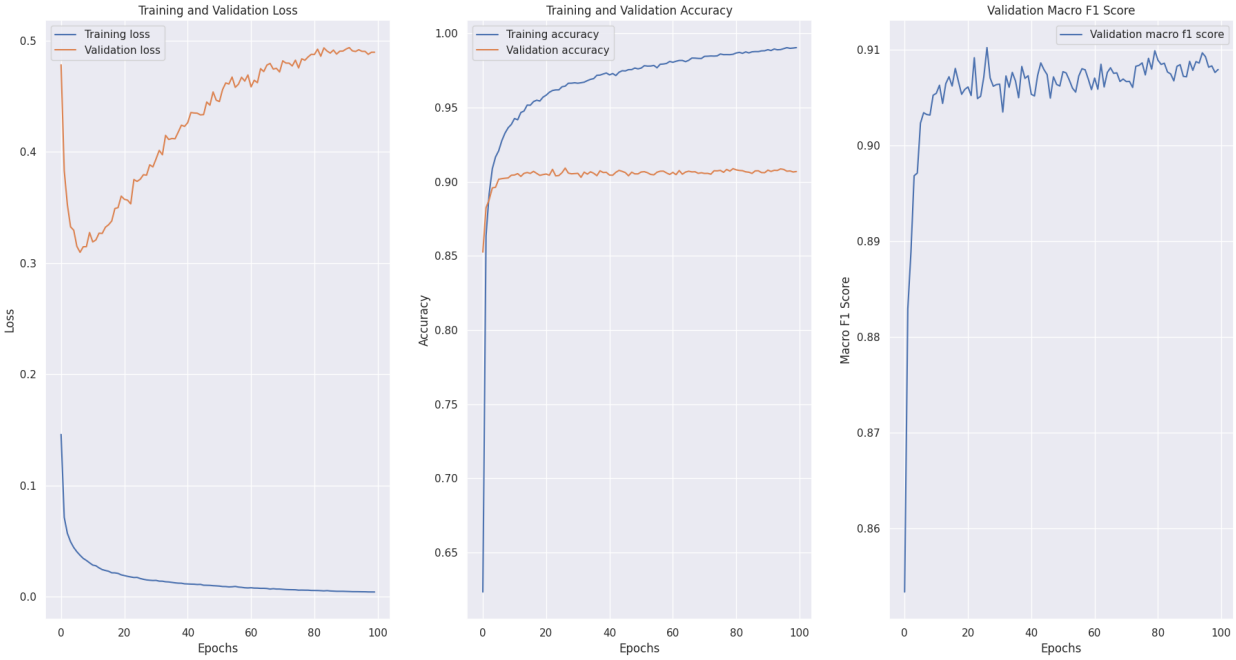
Model 2:

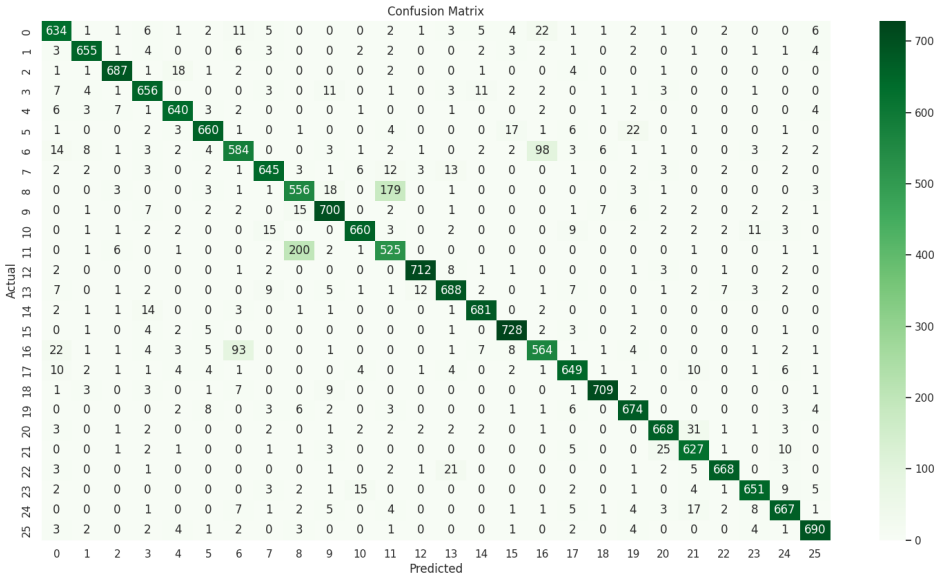
```
# define the network
dense1 = Dense(784, 512)
dense2 = Dense(512, 26)
relu1 = ReLU()
softmax = Softmax()
dropout = Dropout(0.1)

network = [
    dense1,
    relu1,
    dropout,
    dense2,
    softmax
]
```

LR: 5e-3

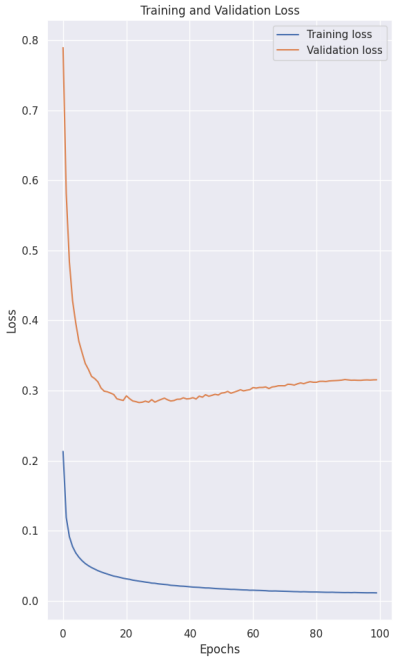
Training Loss: 0.00439
Training Accuracy: 99.0%
Validation Loss: 0.490
Validation Accuracy: 90.7%
Validation Macro F1 Score: 0.908

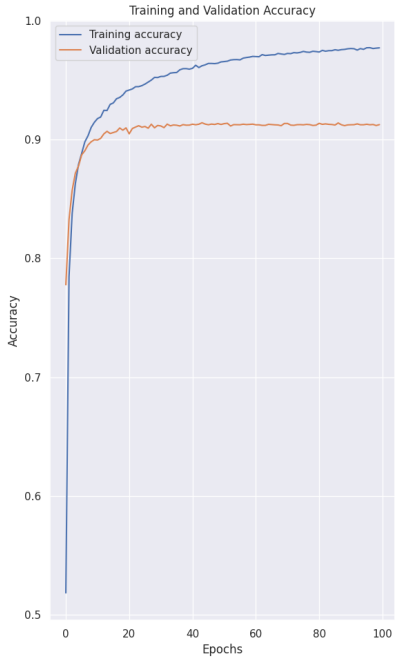

Training and Validation Loss
Training loss
Validation loss
Loss
Epochs
Training and Validation Accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs
Validation Macro F1 Score
Validation macro f1 score
Macro F1 Score
Epochs

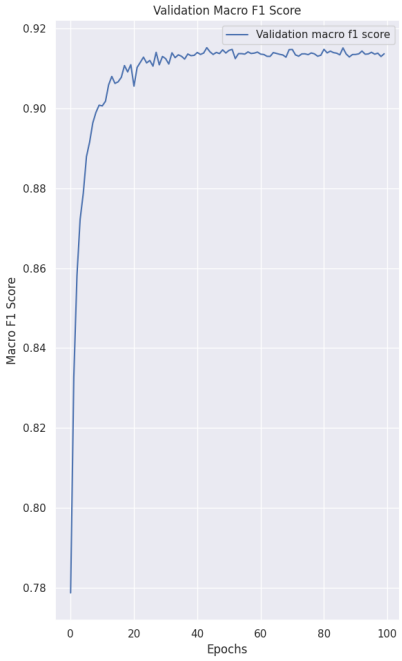

Confusion Matrix
Actual
Predicted

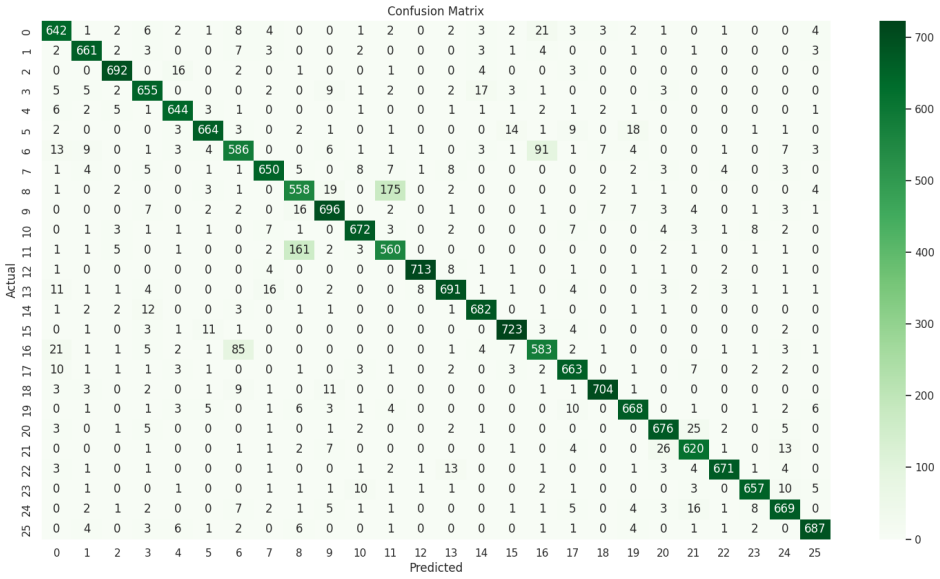
LR: 1e-3

Training Loss: 0.01127
Training Accuracy: 97.7%
Validation Loss: 0.315
Validation Accuracy: 91.3%
Validation Macro F1 Score: **0.914**



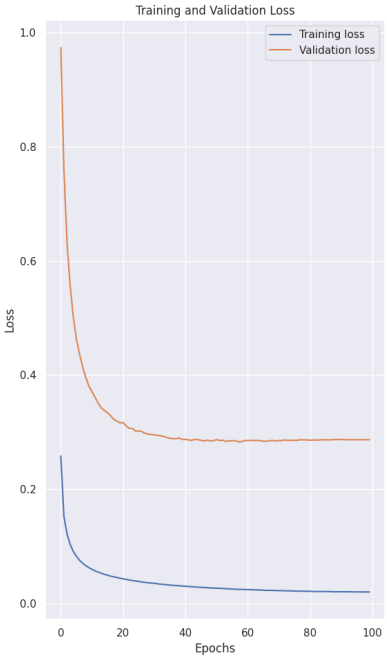


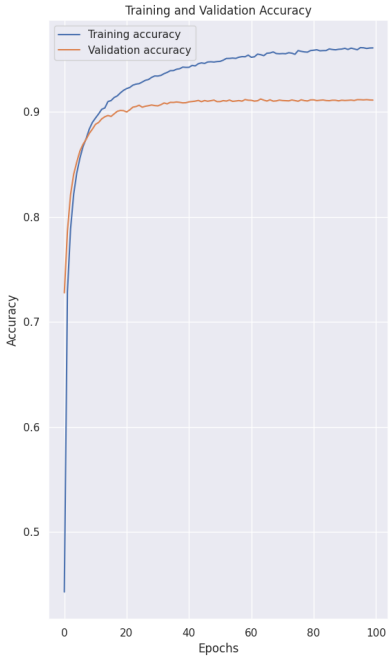


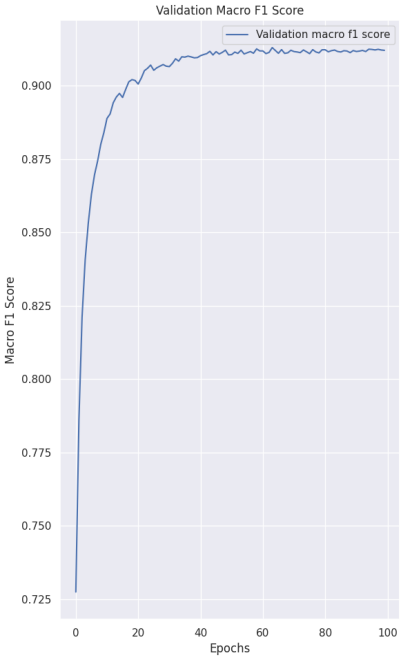


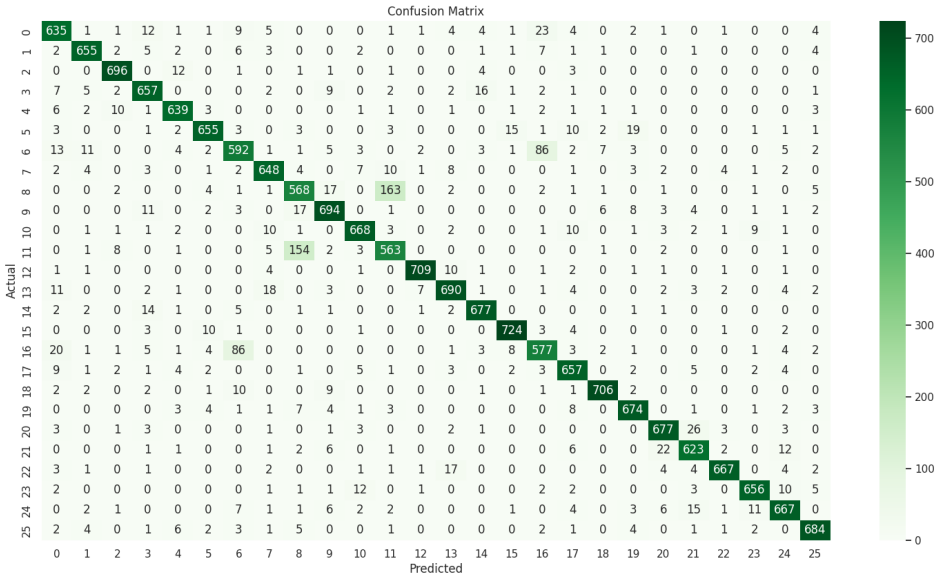
LR: 5e-4

Training Loss: 0.0205
Training Accuracy: 96.1%
Validation Loss: 0.287
Validation Accuracy: 91.1%
Validation Macro F1 Score: 0.912


Training and Validation Loss
Training loss
Validation loss
Loss
Epochs


Training and Validation Accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs


Validation Macro F1 Score
Validation macro f1 score
Macro F1 Score
Epochs


Confusion Matrix
Actual
Predicted

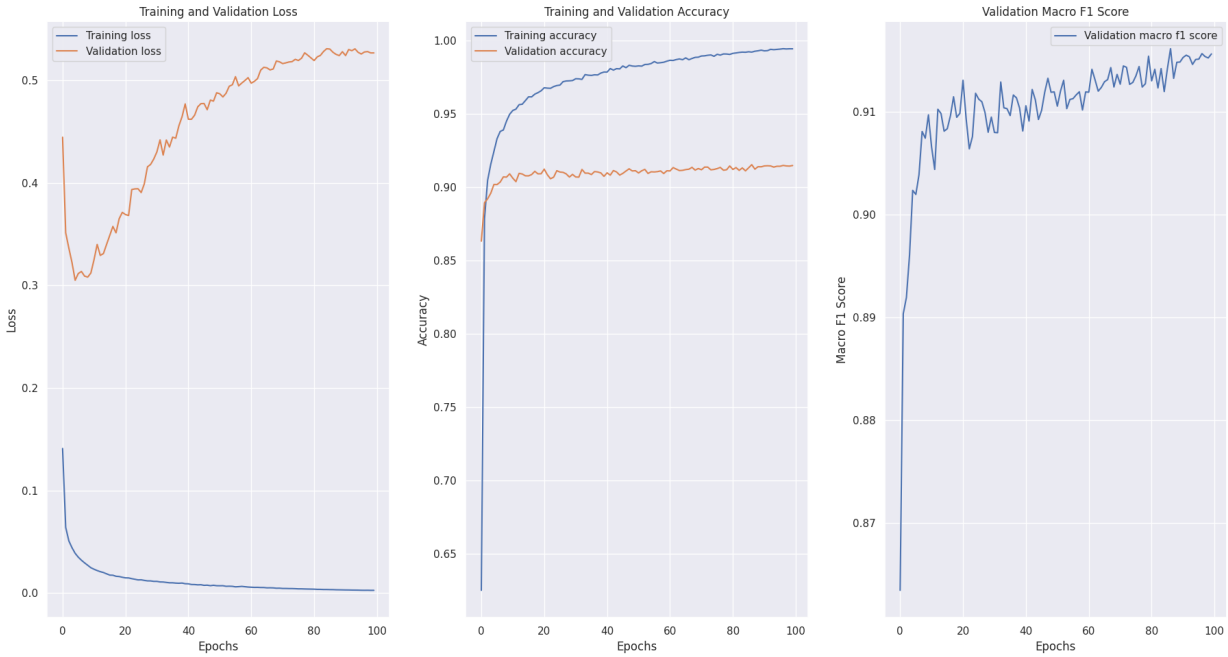
Model 3:

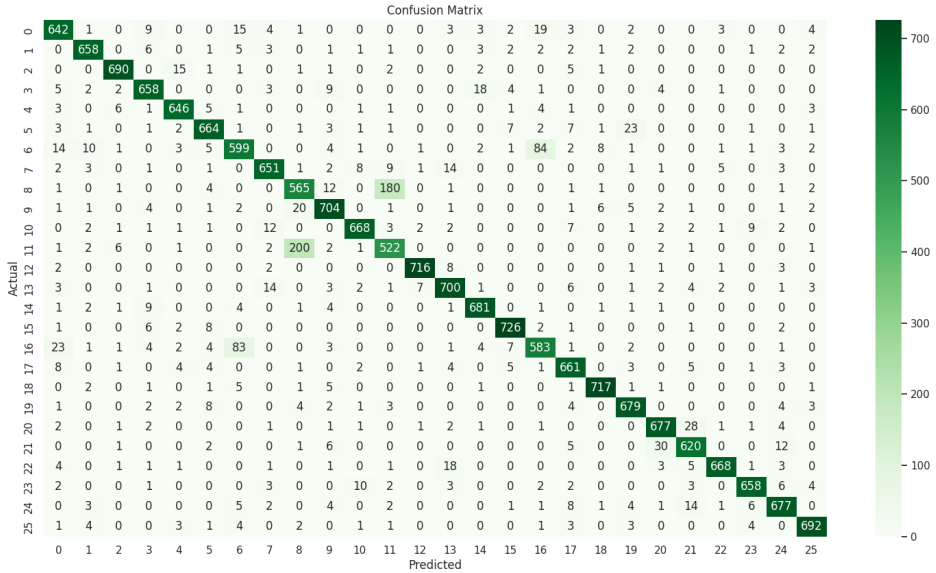
```
# define the network
dense1 = Dense(784, 1024)
dense2 = Dense(1024, 26)
relu1 = ReLU()
softmax = Softmax()
dropout = Dropout(0.1)

network = [
    dense1,
    relu1,
    dropout,
    dense2,
    softmax
]
```

LR: 5e-3 (Best Model Version)

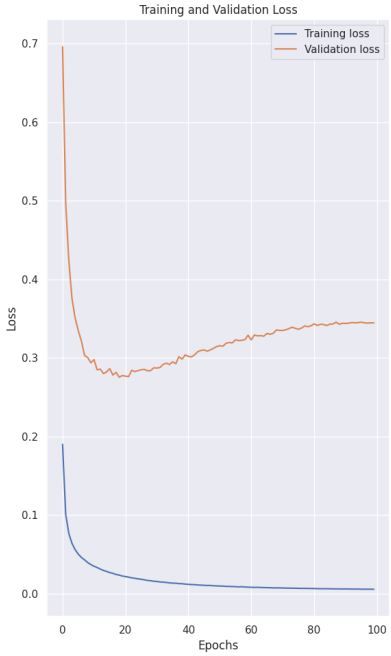
Training Loss: 0.00262
Training Accuracy: **99.4%**
Validation Loss: 0.527
Validation Accuracy: **91.5%**
Validation Macro F1 Score: **0.916**
Test Accuracy: **91.7%**
Test Macro F1 Score: **0.917**

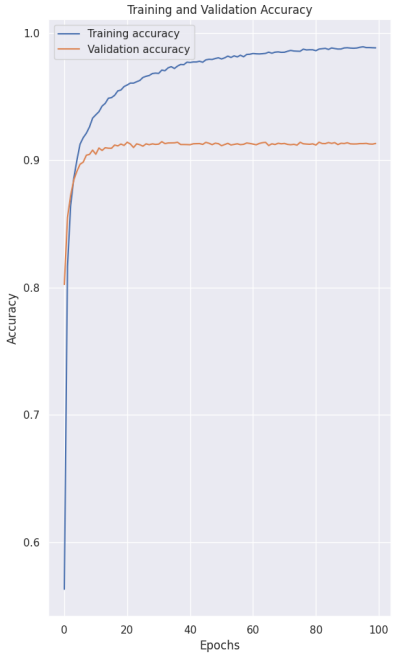

Training and Validation Loss
Training loss
Validation loss
Loss
Epochs
Training and Validation Accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs
Validation Macro F1 Score
Validation macro f1 score
Macro F1 Score
Epochs

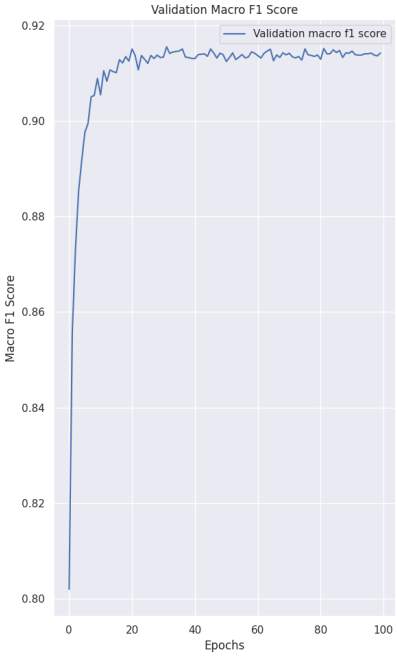

Confusion Matrix
Actual
Predicted

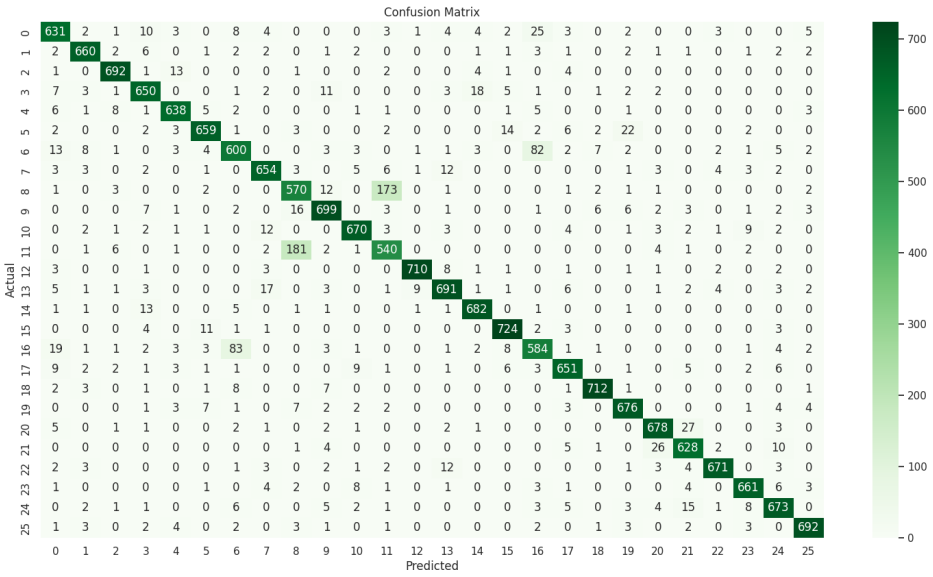
LR: 1e-3

Training Loss: 0.00552
Training Accuracy: 98.8%
Validation Loss: 0.344
Validation Accuracy: 91.3%
Validation Macro F1 Score: 0.914



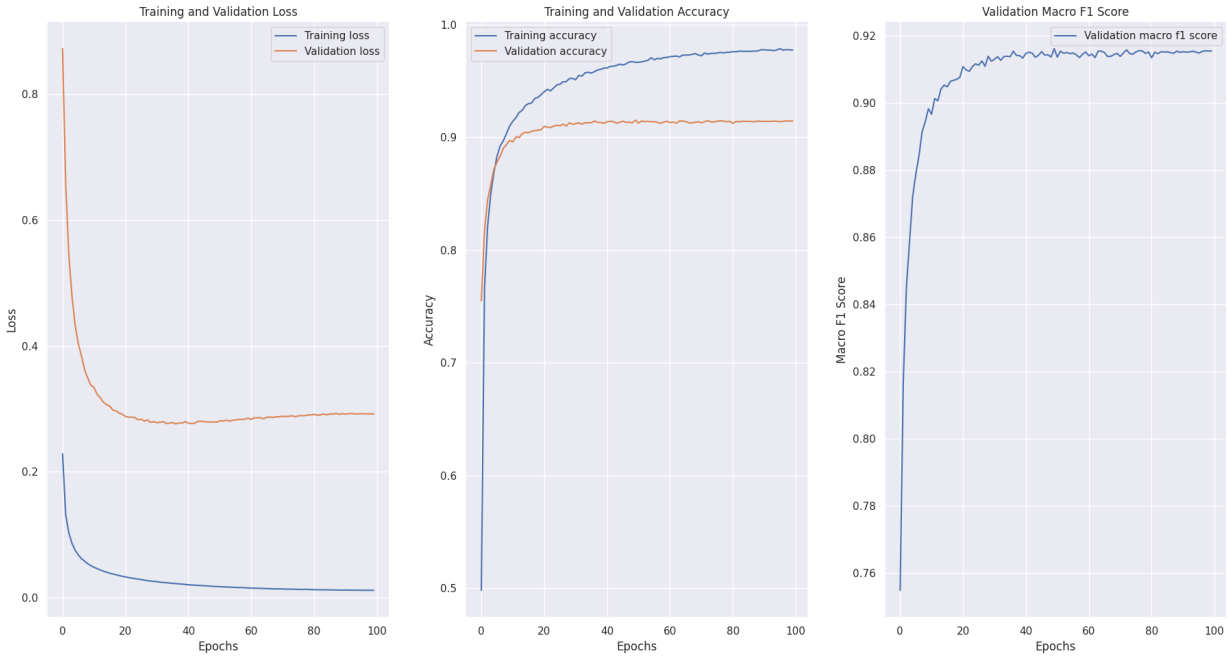


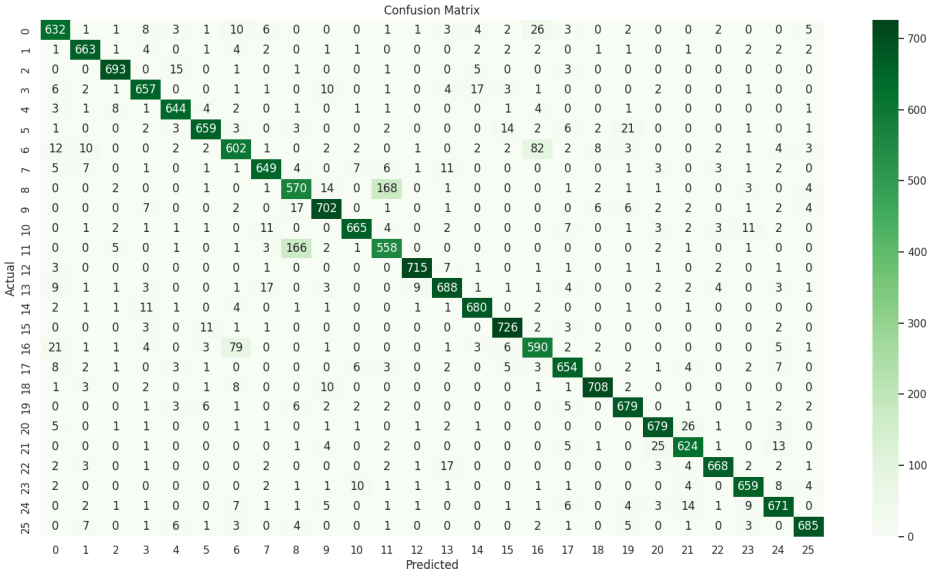




LR: 5e-4

Training Loss: 0.01152
Training Accuracy: 97.8%
Validation Loss: 0.292
Validation Accuracy: 91.5%
Validation Macro F1 Score: 0.915


Training and Validation Loss
Training loss
Validation loss
Loss
Epochs
Training and Validation Accuracy
Training accuracy
Validation accuracy
Accuracy
Epochs
Validation Macro F1 Score
Validation macro f1 score
Macro F1 Score
Epochs


Confusion Matrix
Actual
Predicted