

CSE472

Offline-3

Wasif Jalal
1905084

FNN-1 Architecture

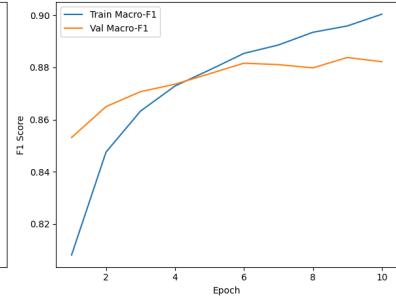
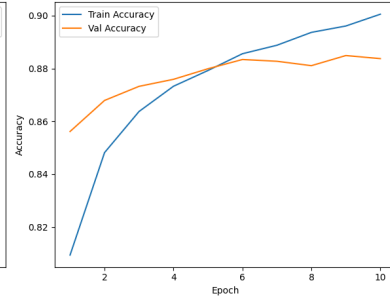
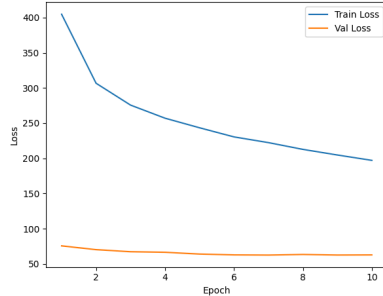
DenseLayer(784, 128),
BatchNorm(128),
ReLU(),
Dropout(0.1),
DenseLayer(128, 64),
BatchNorm(64),
ReLU(),
Dropout(0.1),
DenseLayer(64, 10),
Softmax()

Learning Rate 0.005

Confusion Matrix - Validation Data (Model: <_main_.FNN1 object at 0x7f71f81b2b90>, LR: 0.005)

| | | | | | | | | | | |
|---|-----|------|-----|------|------|------|-----|------|------|------|
| 0 | 932 | 2 | 11 | 24 | 5 | 1 | 208 | 0 | 6 | 0 |
| 1 | 2 | 1130 | 2 | 12 | 2 | 0 | 1 | 0 | 0 | 0 |
| 2 | 12 | 4 | 910 | 8 | 163 | 0 | 137 | 0 | 1 | 0 |
| 3 | 43 | 10 | 6 | 1096 | 65 | 0 | 40 | 0 | 1 | 0 |
| 4 | 2 | 1 | 64 | 18 | 1030 | 0 | 63 | 1 | 4 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1180 | 0 | 26 | 3 | 11 |
| 6 | 98 | 6 | 48 | 23 | 106 | 0 | 919 | 0 | 9 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 34 | 0 | 1149 | 0 | 30 |
| 8 | 1 | 0 | 8 | 0 | 11 | 1 | 19 | 5 | 1081 | 2 |
| 9 | 0 | 0 | 0 | 0 | 0 | 16 | 0 | 51 | 1 | 1145 |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Loss vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.005) Accuracy vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.005) F1 Score vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.005)



Training Loss: 0.2628, Training Accuracy: 0.9006, Training Macro-F1 Score: 0.9004

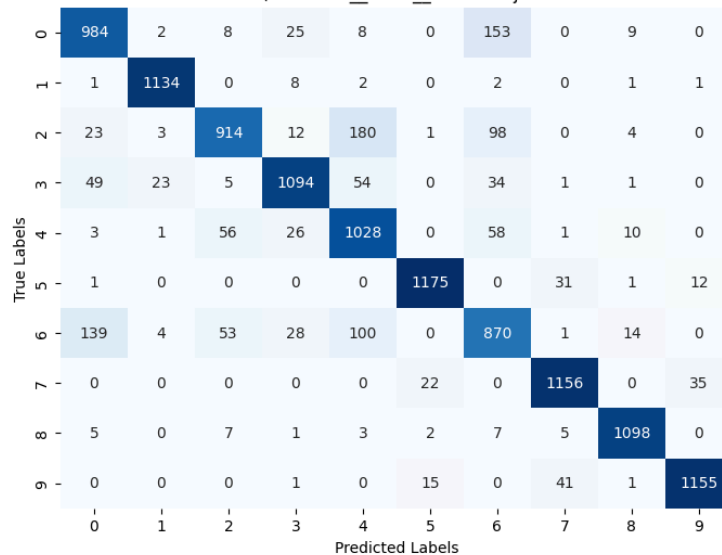
Validation Loss: 0.3344, Validation Accuracy: 0.8838, Validation Macro-F1 Score: 0.8822

FNN-1 Architecture

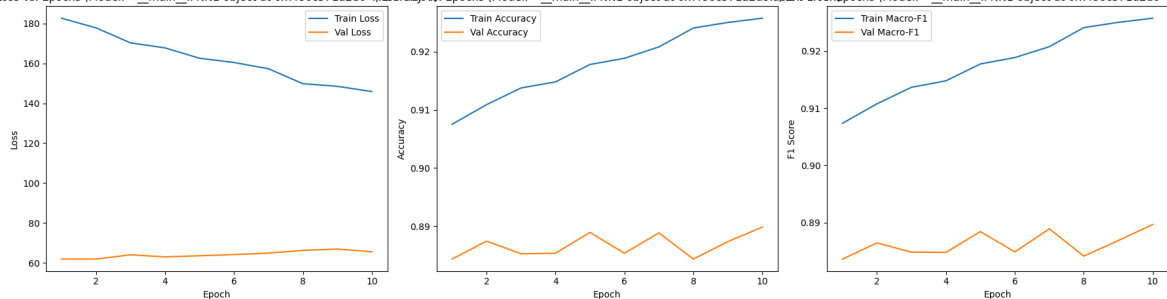
DenseLayer(784, 128),
BatchNorm(128),
ReLU(),
Dropout(0.1),
DenseLayer(128, 64),
BatchNorm(64),
ReLU(),
Dropout(0.1),
DenseLayer(64, 10),
Softmax()

Learning Rate 0.004

Confusion Matrix - Validation Data (Model: <_main_.FNN1 object at 0x7f71f81b2b90>, LR: 0.004)



Loss vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.004), Accuracy vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.004), Macro-F1 vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.004)



Training Loss: 0.1945, Training Accuracy: 0.9258, Training Macro-F1 Score: 0.9258

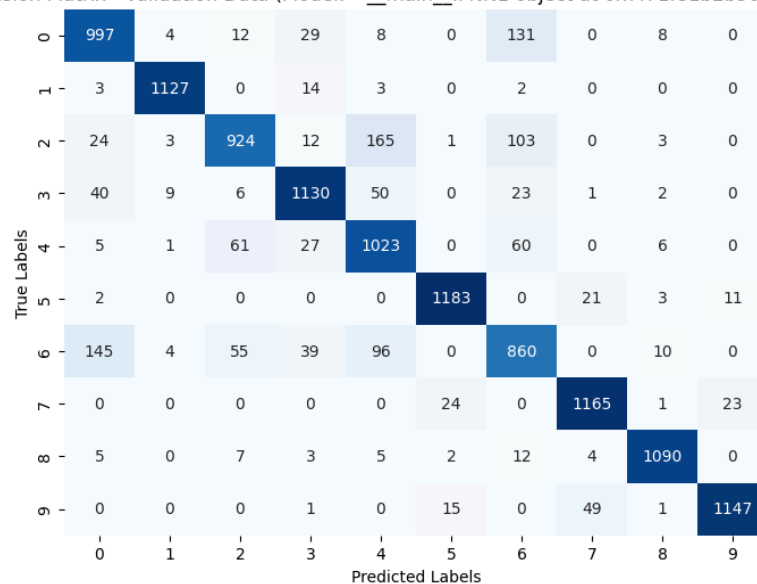
Validation Loss: 0.3492, Validation Accuracy: 0.8898, Validation Macro-F1 Score: 0.8896

FNN-1 Architecture

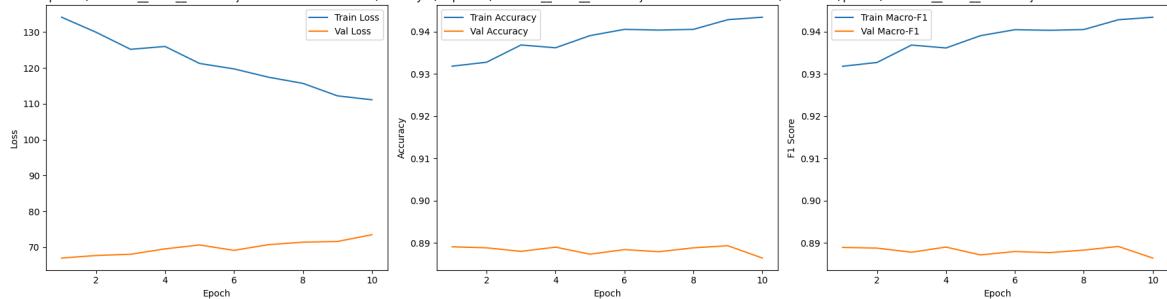
DenseLayer(784, 128),
BatchNorm(128),
ReLU(),
Dropout(0.1),
DenseLayer(128, 64),
BatchNorm(64),
ReLU(),
Dropout(0.1),
DenseLayer(64, 10),
Softmax()

Learning Rate 0.003

Confusion Matrix - Validation Data (Model: <_main_.FNN1 object at 0x7f71f81b2b90>, LR: 0.003)



Loss vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.003) Accuracy vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.003) F1 Score vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.003)



Training Loss: 0.1482, Training Accuracy: 0.9434, Training Macro-F1 Score: 0.9434

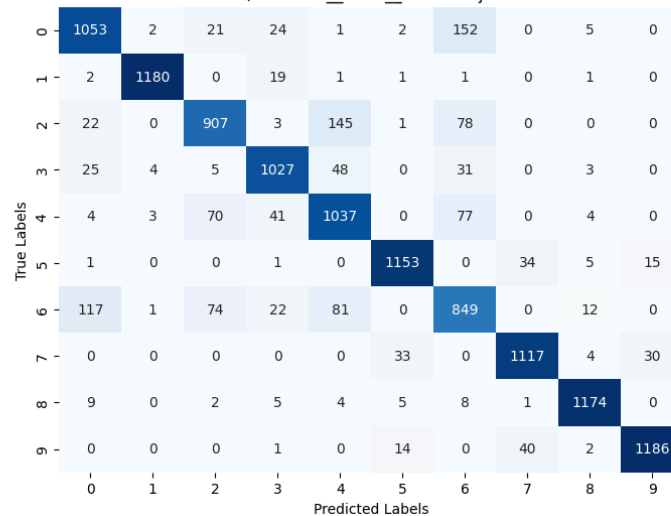
Validation Loss: 0.3910, Validation Accuracy: 0.8864, Validation Macro-F1 Score: 0.8865

FNN-1 Architecture

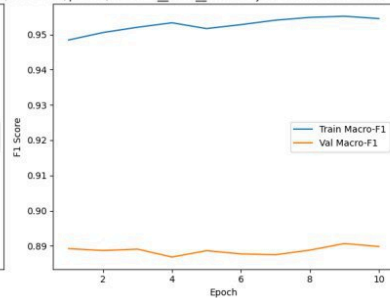
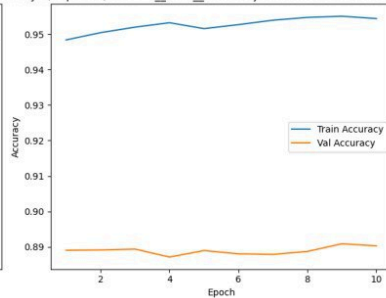
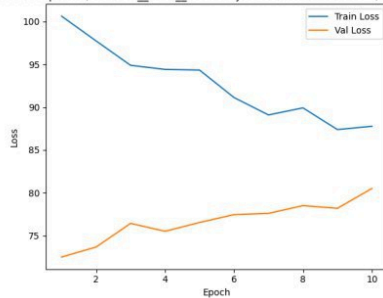
DenseLayer(784, 128),
BatchNorm(128),
ReLU(),
Dropout(0.1),
DenseLayer(128, 64),
BatchNorm(64),
ReLU(),
Dropout(0.1),
DenseLayer(64, 10),
Softmax()

Learning Rate 0.002

Confusion Matrix - Validation Data (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.002)



Loss vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.002) Accuracy vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.002) F1 Score vs. Epochs (Model: <_main_.FNN1 object at 0x7f86c571a2d0>, LR: 0.002)



Training Loss: 0.1170, Training Accuracy: 0.9545, Training Macro-F1 Score: 0.9545

Validation Loss: 0.4282, Validation Accuracy: 0.8902, Validation Macro-F1 Score: 0.8898

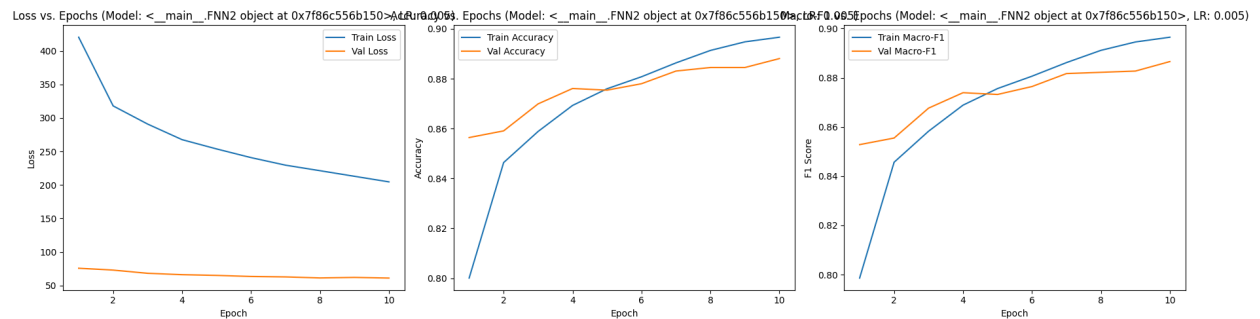
FNN-2 Architecture

```
DenseLayer(784, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.005

Confusion Matrix - Validation Data (Model: <_main__FNN2 object at 0x7f71f8913990>, LR: 0.005)

| | | | | | | | | | | |
|---|-----|------|-----|------|------|------|-----|------|------|------|
| 0 | 981 | 5 | 16 | 25 | 11 | 4 | 136 | 0 | 11 | 0 |
| 1 | 2 | 1130 | 1 | 13 | 1 | 0 | 1 | 0 | 1 | 0 |
| 2 | 14 | 2 | 895 | 11 | 200 | 0 | 113 | 0 | 0 | 0 |
| 3 | 34 | 9 | 8 | 1114 | 72 | 1 | 21 | 0 | 2 | 0 |
| 4 | 5 | 0 | 52 | 16 | 1048 | 1 | 55 | 0 | 6 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1186 | 0 | 22 | 3 | 9 |
| 6 | 150 | 4 | 63 | 28 | 124 | 1 | 822 | 0 | 17 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 55 | 0 | 1121 | 2 | 35 |
| 8 | 2 | 1 | 8 | 0 | 6 | 4 | 6 | 3 | 1098 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 46 | 0 | 1143 |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |



Training Loss: 0.2728, Training Accuracy: 0.8966, Training Macro-F1 Score: 0.8965

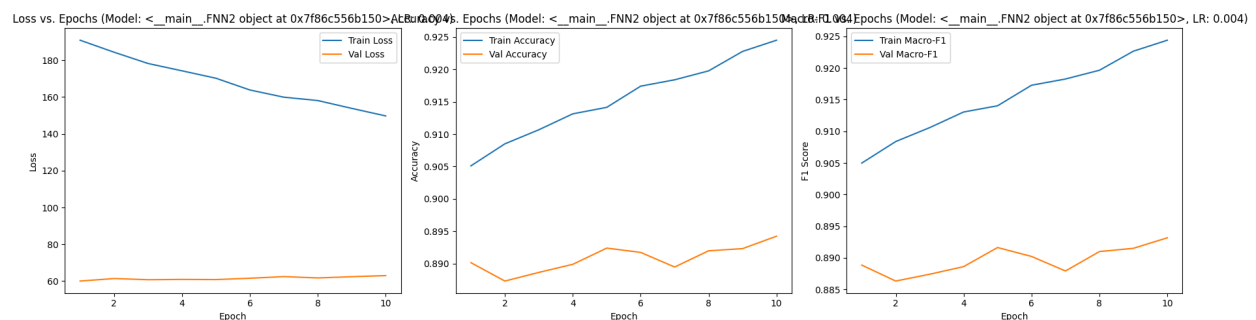
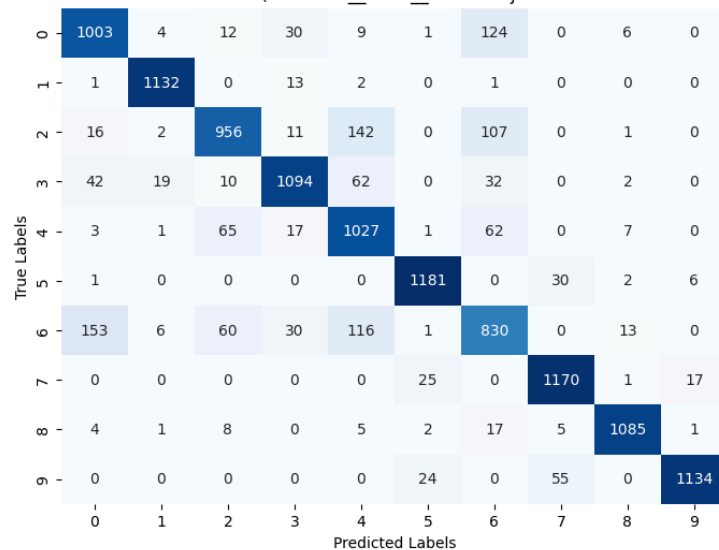
Validation Loss: 0.3255, Validation Accuracy: 0.8881, Validation Macro-F1 Score: 0.8866

FNN-2 Architecture

```
DenseLayer(784, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.004

Confusion Matrix - Validation Data (Model: <_main_.FNN2 object at 0x7f71f8913990>, LR: 0.004)



Training Loss: 0.1997, Training Accuracy: 0.9245, Training Macro-F1 Score: 0.9244

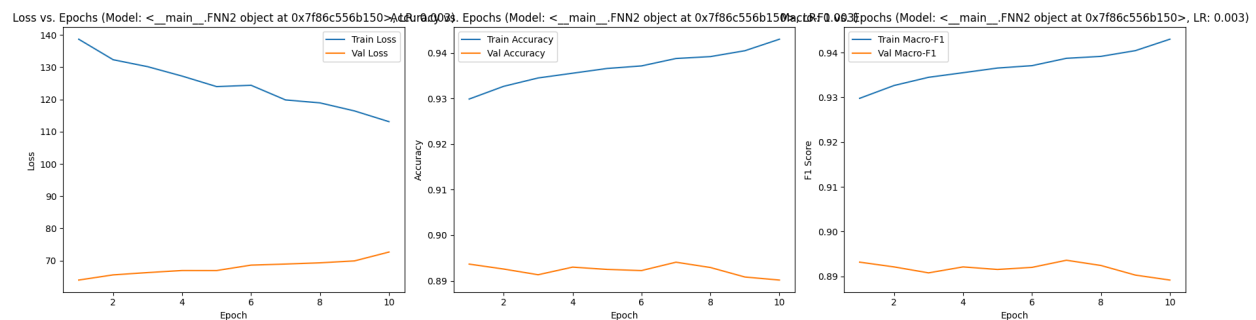
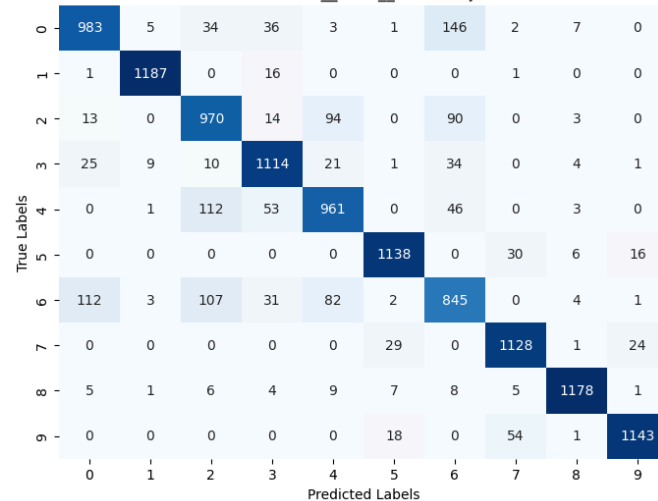
Validation Loss: 0.3352, Validation Accuracy: 0.8942, Validation Macro-F1 Score: 0.8932

FNN-2 Architecture

```
DenseLayer(784, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.003

Confusion Matrix - Validation Data (Model: <_main_.FNN2 object at 0x7f71fd7499d0>, LR: 0.003)



Training Loss: 0.1508, Training Accuracy: 0.9430, Training Macro-F1 Score: 0.9430

Validation Loss: 0.3864, Validation Accuracy: 0.8902, Validation Macro-F1 Score: 0.8892

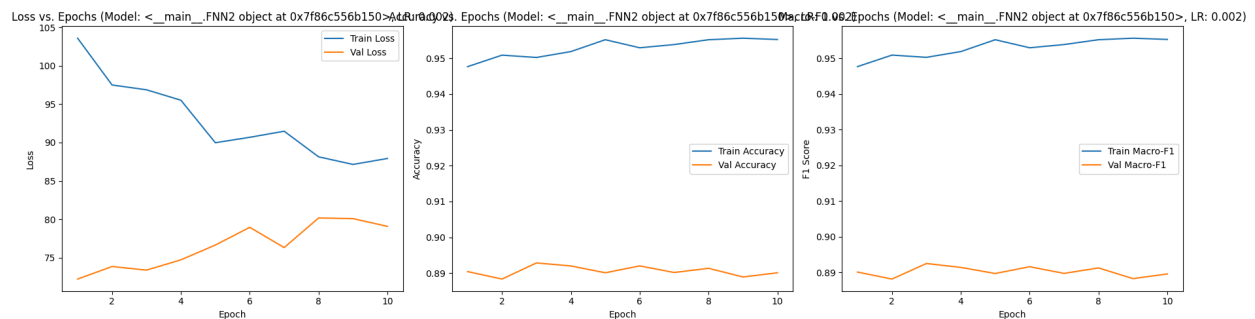
FNN-2 Architecture

```
DenseLayer(784, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.002

Confusion Matrix - Validation Data (Model: <__main__.FNN2 object at 0x7f86c556b150>, LR: 0.002)

| | | | | | | | | | | |
|---|------|------|-----|------|------|------|-----|------|------|------|
| 0 | 1031 | 2 | 19 | 34 | 1 | 3 | 162 | 0 | 8 | 0 |
| 1 | 3 | 1181 | 0 | 14 | 3 | 1 | 2 | 0 | 1 | 0 |
| 2 | 12 | 2 | 973 | 9 | 91 | 0 | 67 | 0 | 2 | 0 |
| 3 | 25 | 7 | 8 | 1027 | 43 | 0 | 31 | 0 | 2 | 0 |
| 4 | 0 | 2 | 93 | 47 | 1009 | 0 | 80 | 0 | 5 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1156 | 0 | 30 | 7 | 16 |
| 6 | 113 | 1 | 101 | 26 | 65 | 1 | 840 | 0 | 9 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 38 | 0 | 1095 | 2 | 49 |
| 8 | 8 | 0 | 6 | 3 | 5 | 1 | 9 | 2 | 1173 | 1 |
| 9 | 0 | 0 | 1 | 0 | 0 | 14 | 0 | 31 | 1 | 1196 |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |



Training Loss: 0.1172, Training Accuracy: 0.9552, Training Macro-F1 Score: 0.9552

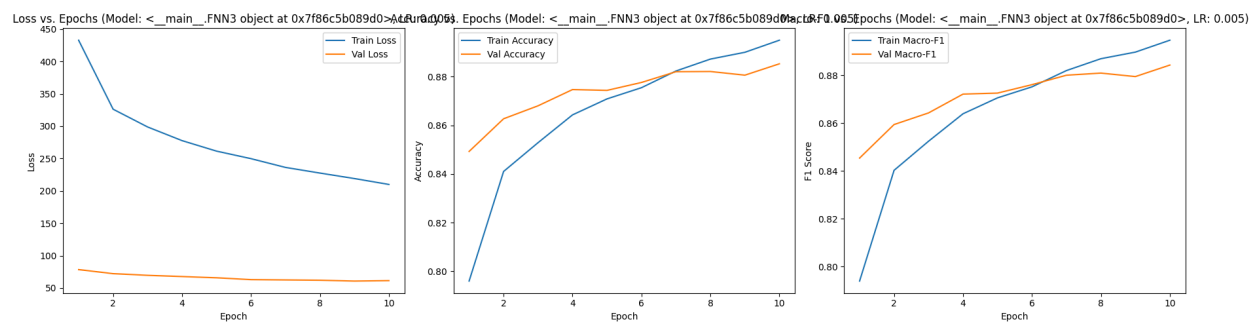
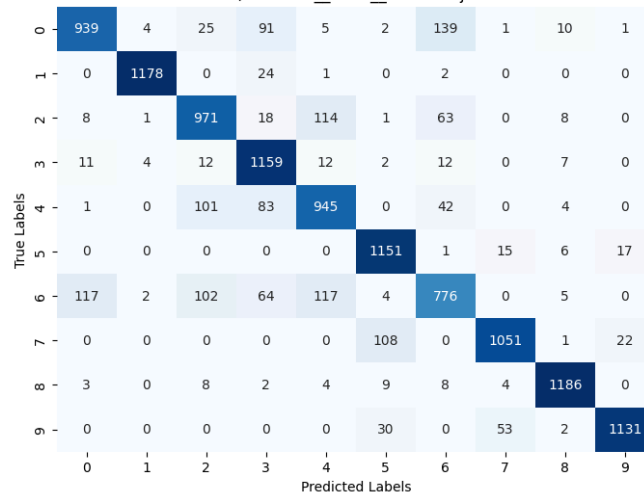
Validation Loss: 0.4207, Validation Accuracy: 0.8901, Validation Macro-F1 Score: 0.8895

FNN-3 Architecture

```
DenseLayer(784, 512),  
BatchNorm(512),  
ReLU(),  
Dropout(0.1),  
DenseLayer(512, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.005

Confusion Matrix - Validation Data (Model: <_main_.FNN3 object at 0x7f71f82e4790>, LR: 0.005)



Training Loss: 0.2800, Training Accuracy: 0.8949, Training Macro-F1 Score: 0.8947

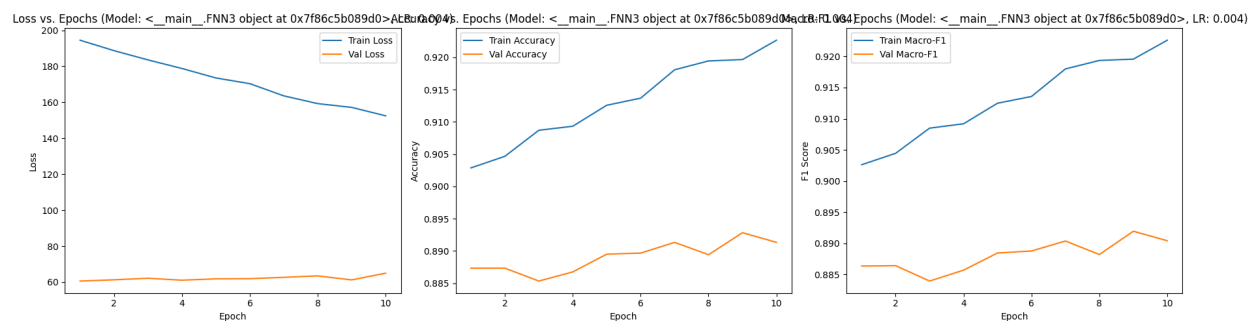
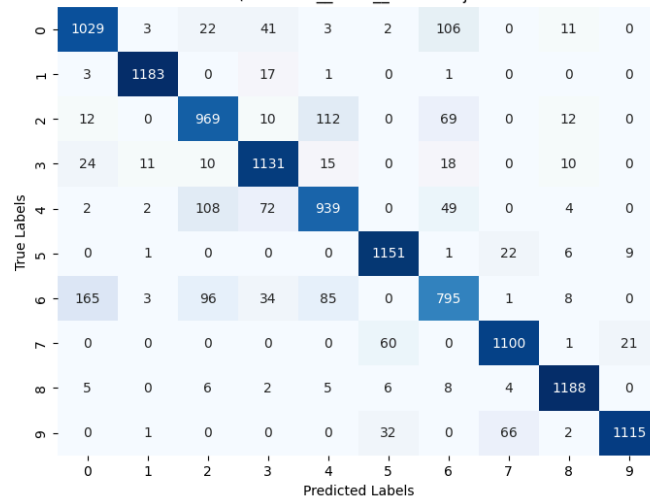
Validation Loss: 0.3267, Validation Accuracy: 0.8852, Validation Macro-F1 Score: 0.8843

FNN-3 Architecture

```
DenseLayer(784, 512),  
BatchNorm(512),  
ReLU(),  
Dropout(0.1),  
DenseLayer(512, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.004

Confusion Matrix - Validation Data (Model: <_main_.FNN3 object at 0x7f71f82e4790>, LR: 0.004)



Training Loss: 0.2033, Training Accuracy: 0.9227, Training Macro-F1 Score: 0.9226

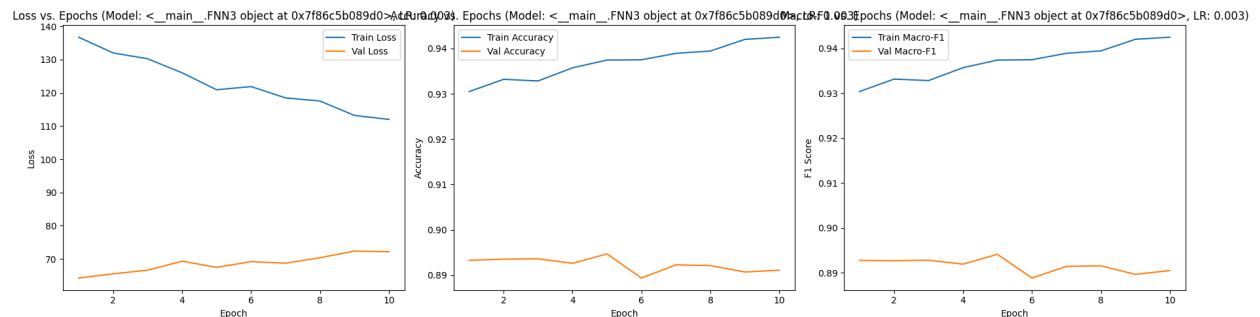
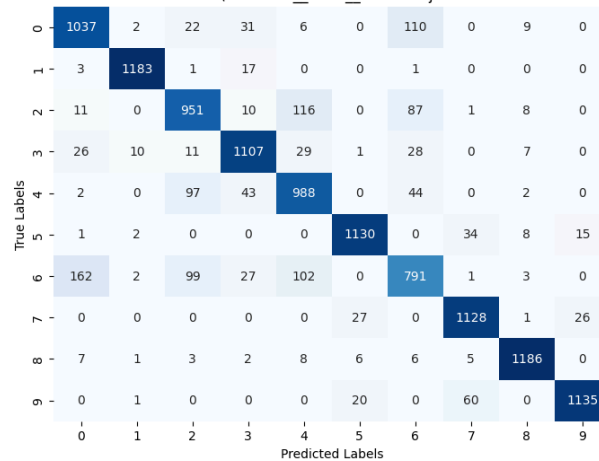
Validation Loss: 0.3457, Validation Accuracy: 0.8913, Validation Macro-F1 Score: 0.8904

FNN-3 Architecture

```
DenseLayer(784, 512),  
BatchNorm(512),  
ReLU(),  
Dropout(0.1),  
DenseLayer(512, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.003

Confusion Matrix - Validation Data (Model: <_main_.FNN3 object at 0x7f71f82e4790>, LR: 0.003)



Training Loss: 0.1494, Training Accuracy: 0.9425, Training Macro-F1 Score: 0.9425

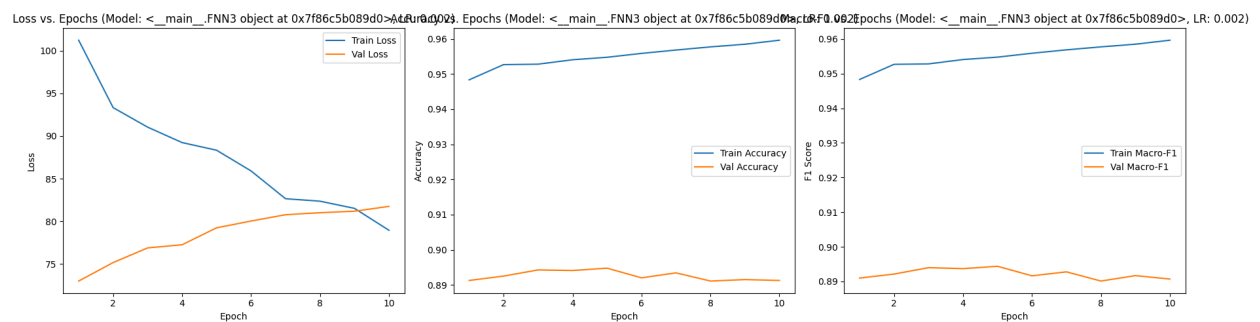
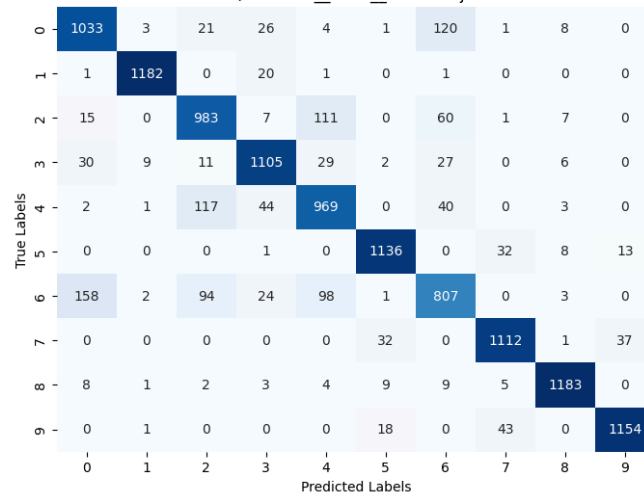
Validation Loss: 0.3840, Validation Accuracy: 0.8911, Validation Macro-F1 Score: 0.8905

FNN-3 Architecture

```
DenseLayer(784, 512),  
BatchNorm(512),  
ReLU(),  
Dropout(0.1),  
DenseLayer(512, 256),  
BatchNorm(256),  
ReLU(),  
Dropout(0.1),  
DenseLayer(256, 128),  
BatchNorm(128),  
ReLU(),  
Dropout(0.1),  
DenseLayer(128, 64),  
BatchNorm(64),  
ReLU(),  
Dropout(0.1),  
DenseLayer(64, 10),  
Softmax()
```

Learning Rate 0.002

Confusion Matrix - Validation Data (Model: <__main__.FNN3 object at 0x7f71f82e4790>, LR: 0.002)



Training Loss: 0.1053, Training Accuracy: 0.9596, Training Macro-F1 Score: 0.9597

Validation Loss: 0.4349, Validation Accuracy: 0.8912, Validation Macro-F1 Score: 0.8907

Best Model

From the above analysis it is observed that FNN-2 with learning rate 0.004 performs best on the validation test, so we use it to predict on the independent test set and get the following results:

Training Loss: 0.2360, Training Accuracy: 0.9116, Training Macro-F1 Score: 0.9114

Test Loss: 0.3031, Test Accuracy: 0.9068, Test Macro-F1 Score: 0.9063