

CSE 472

FNN from Scratch

Offline 3

Mashiyat Mahjabin Prapty
1805117

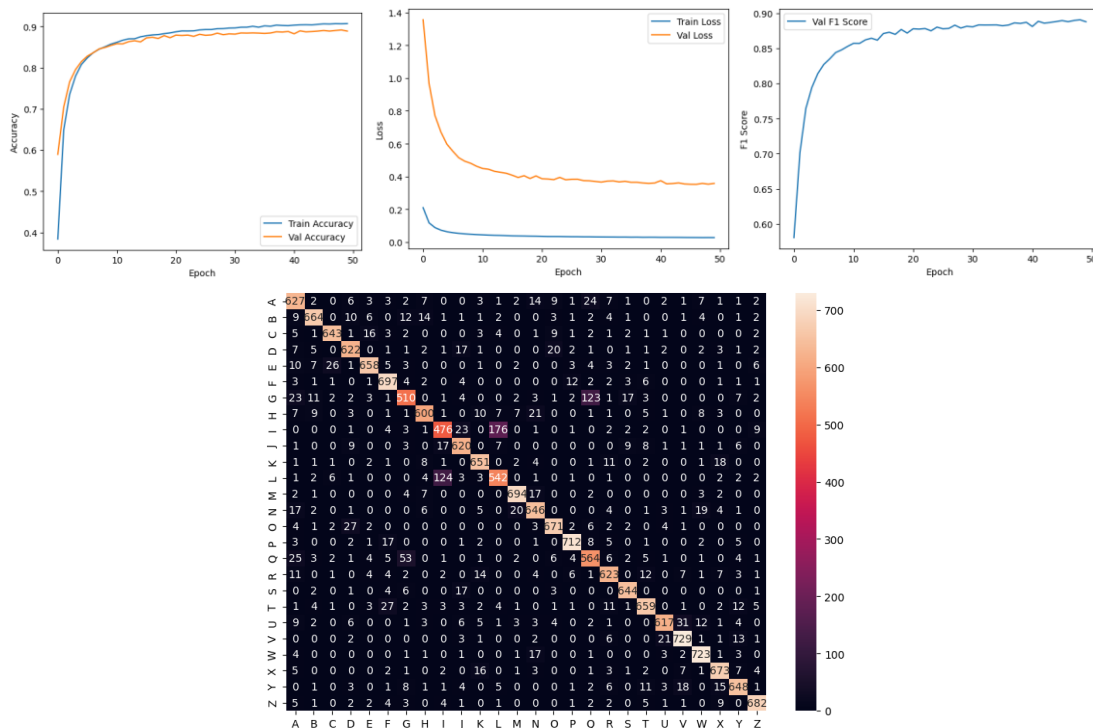
Hyperparameters

1. Starting Learning Rate 8e-3
2. Ending Learning Rate 5e-4
3. Batch size 1024
4. Epochs 50
5. Weight saved on Best Validation Accuracy
6. Hybrid LR Scheduler
7. Default Adam Parameters
8. He Initialization

Model 1(Best Model):

```
[  
    Dense('dense1', 784, 256),  
    ReLU('relu1'),  
    Dropout('dropout1', 0.8),  
    Dense('dense2', 256, 256),  
    ReLU('relu2'),  
    Dropout('dropout2', 0.8),  
    Dense('dense3', 256, 26),  
    Softmax('softmax')  
],
```

LR = 8e-3

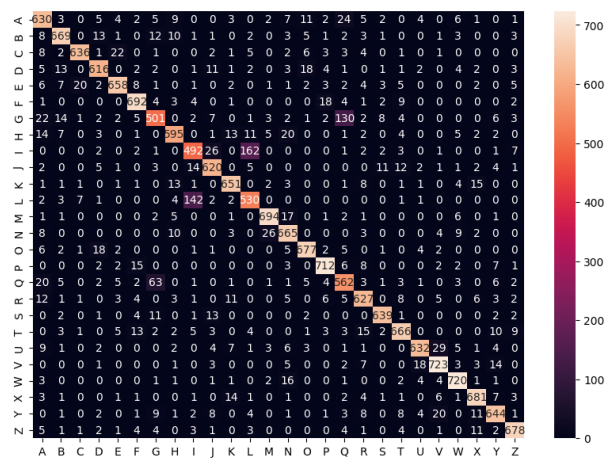
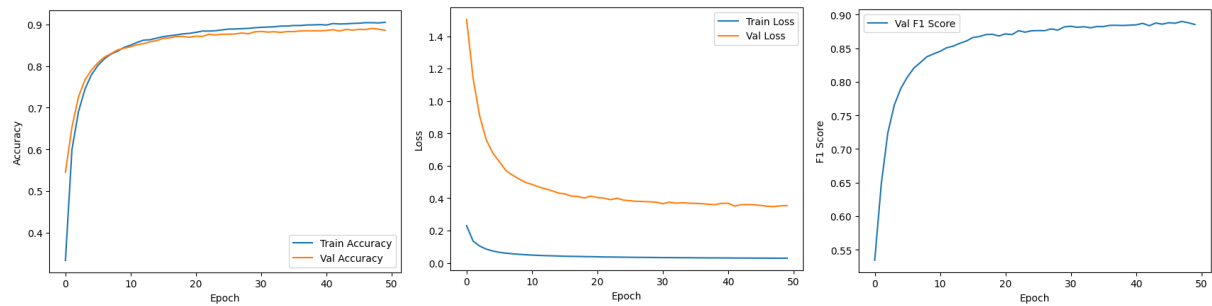


Best Validation Accuracy: 89.14529914529915 %

Test Accuracy: 88.51442307692308 %

Test F1 Score: 88.49860220062828 %

LR = 5e-3

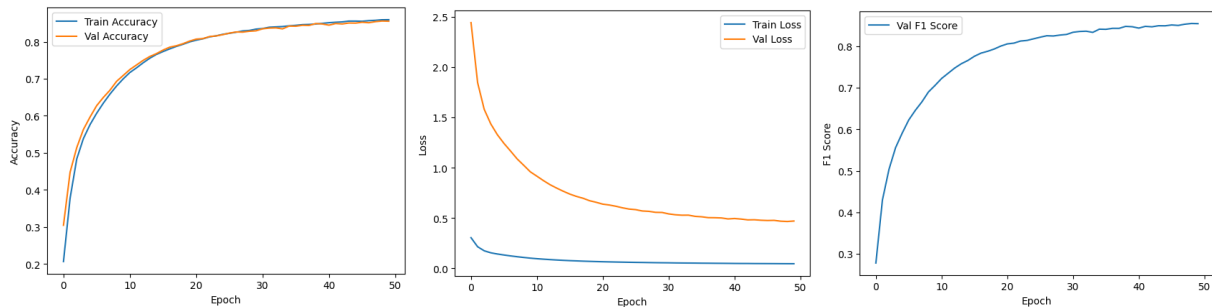


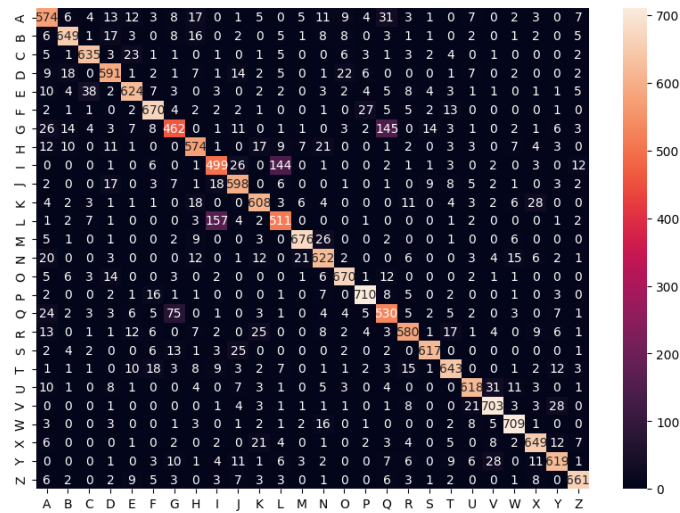
Best Validation Accuracy: 89.07051282051283 %

Test Accuracy: 88.36057692307693 %

Test F1 Score: 88.35513019904049 %

LR = 1e-3



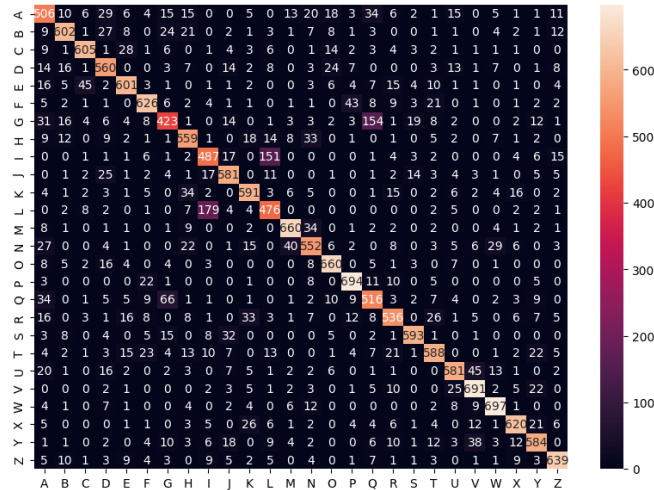
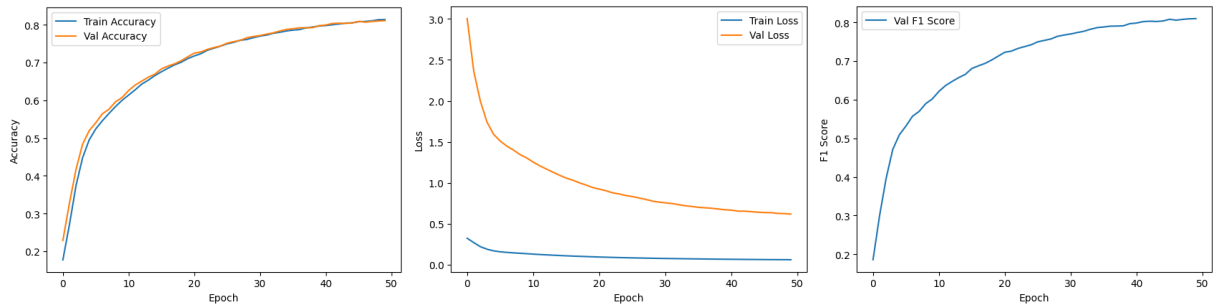


Best Validation Accuracy: 85.58760683760683 %

Test Accuracy: 85.03846153846155 %

Test F1 Score: 85.00598945885577 %

LR = 5e-4



Best Validation Accuracy: 81.08974358974359 %

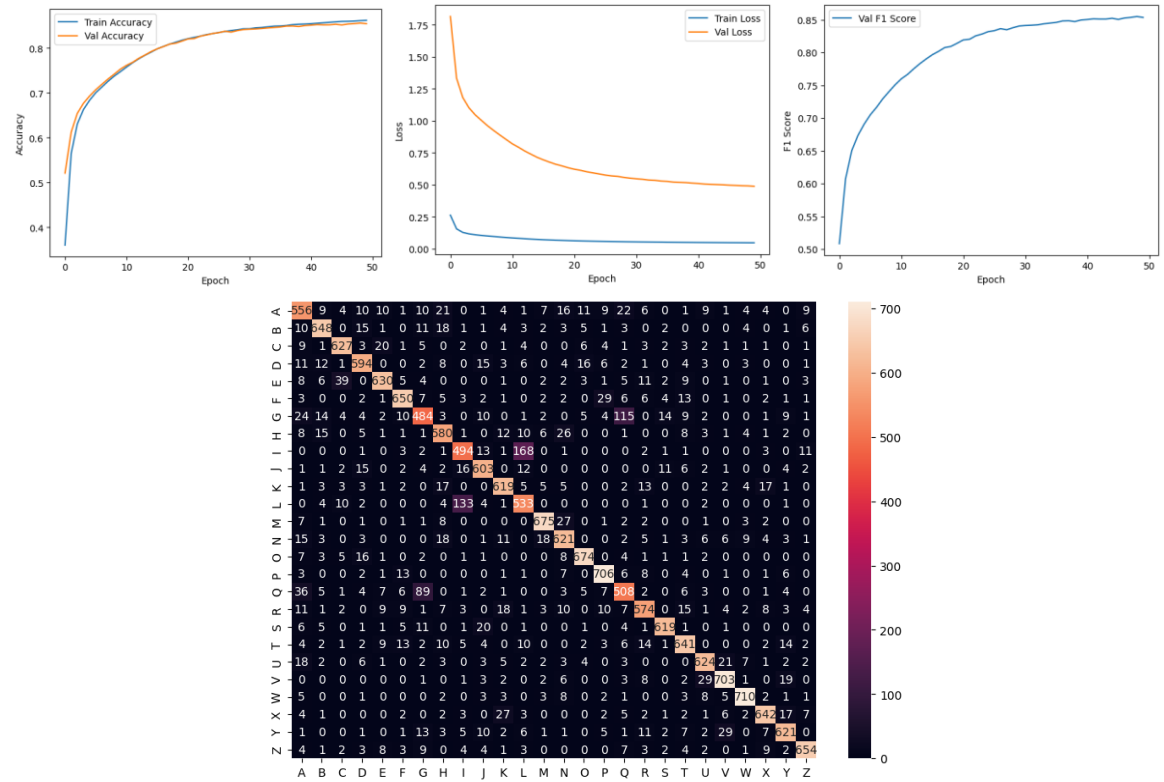
Test Accuracy: 80.96153846153847 %

Test F1 Score: 80.89687099086281 %

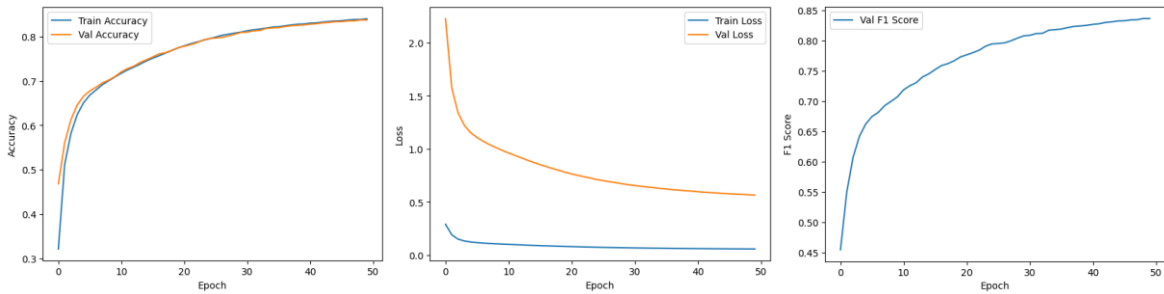
Model 2:

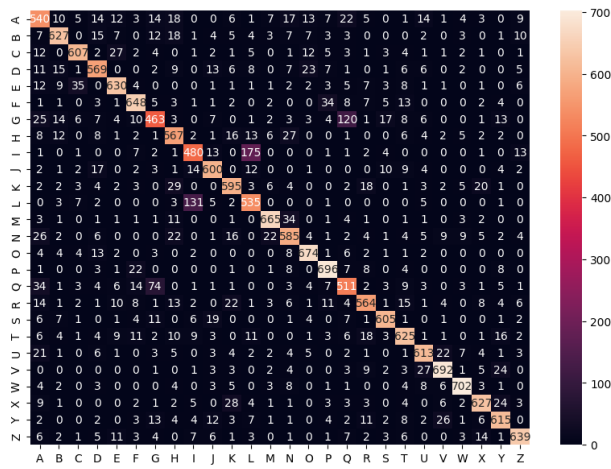
```
[
    Dense('dense1', 784, 128),
    ReLU('relu1'),
    Dense('dense3', 128, 26),
    Softmax('softmax')
],
```

LR = 8e-3

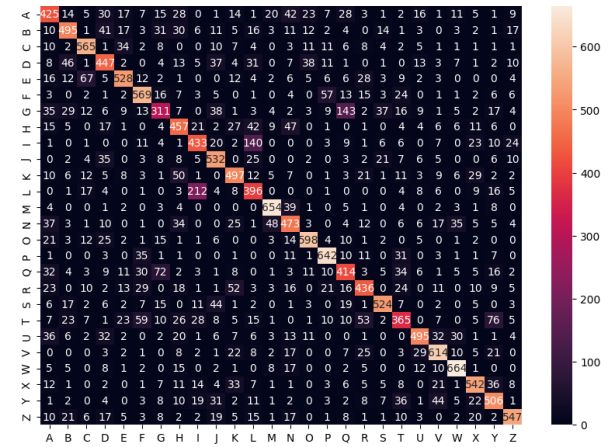
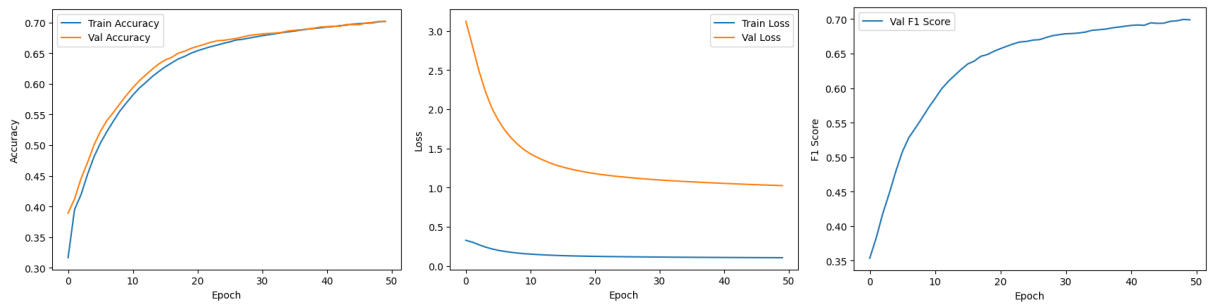


LR = 5e-3

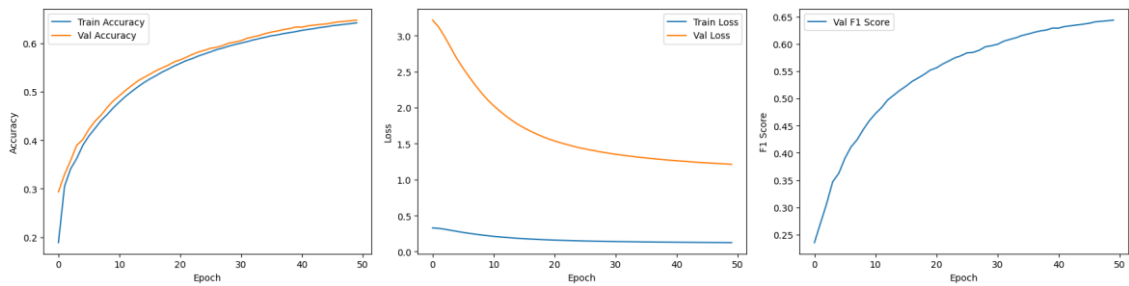


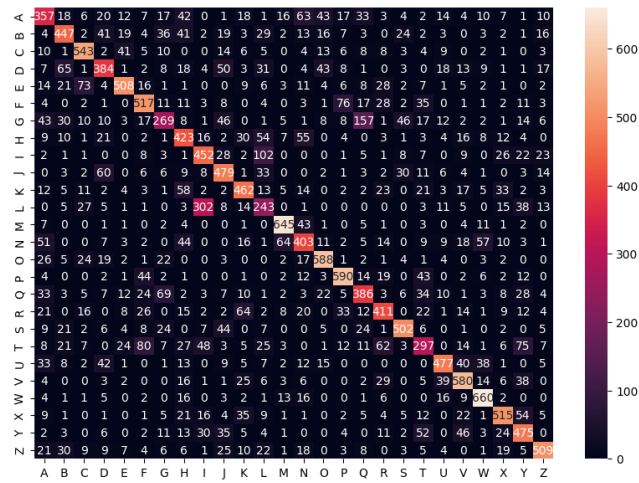


LR = 1e-3



LR = 5e-4

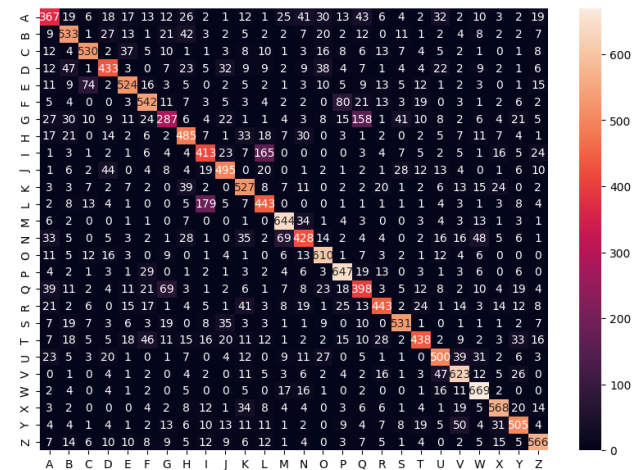
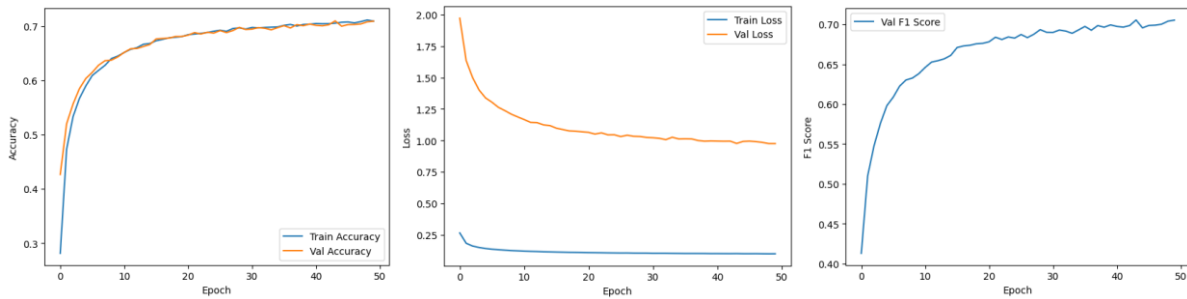




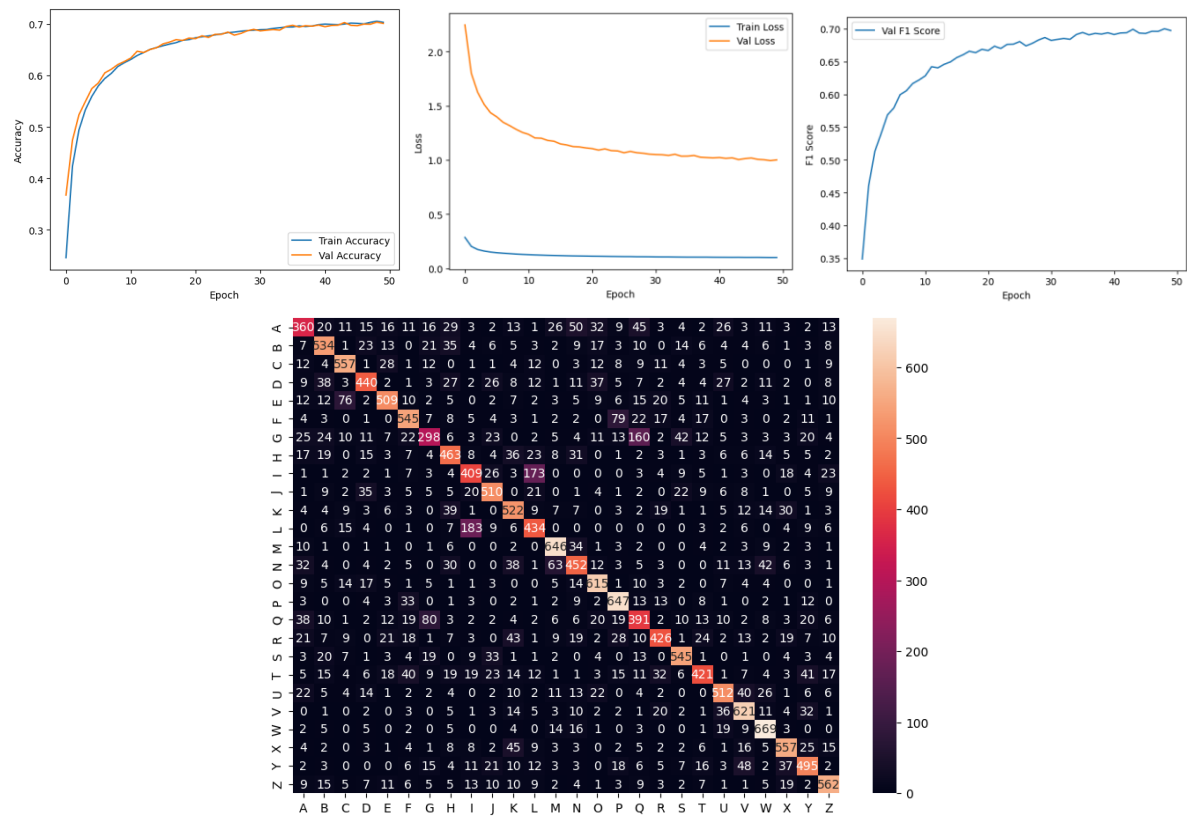
Model 3:

```
[
    Dense('dense1', 784, 256),
    ReLU('relu1'),
    Dropout('dropout1', 0.3),
    Dense('dense3', 256, 26),
    Softmax('softmax')
]
```

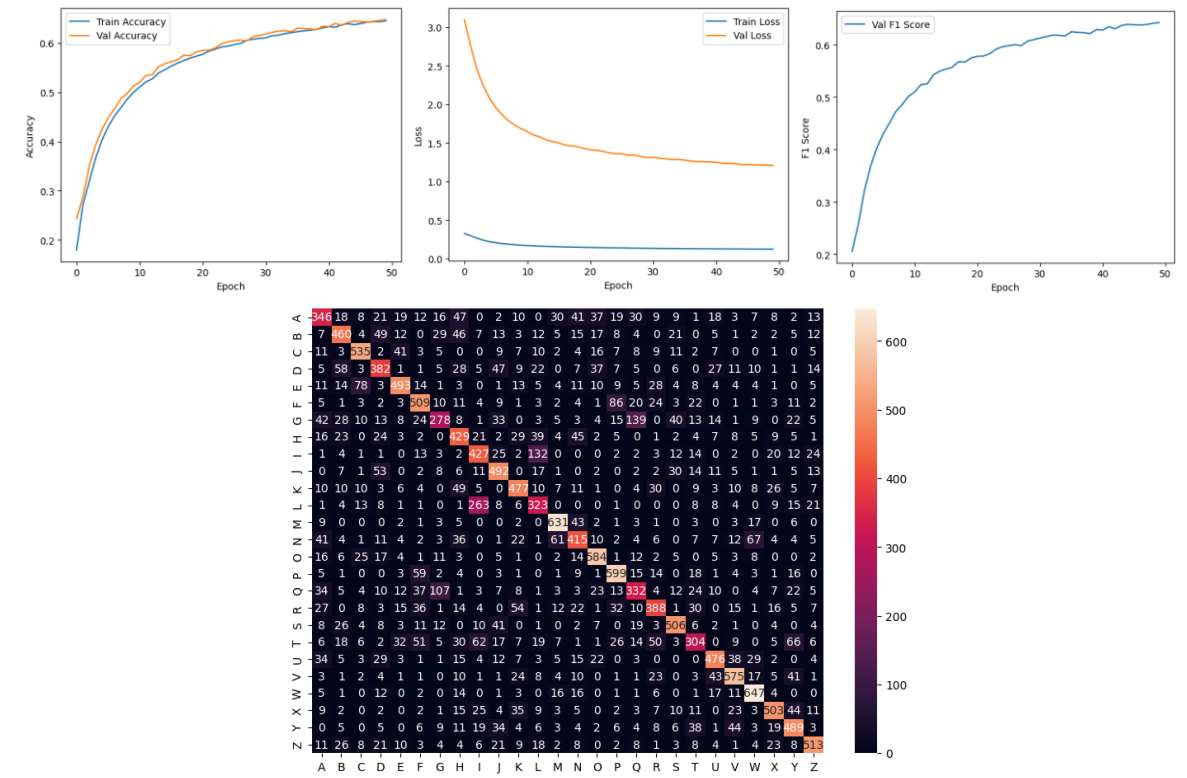
LR = 8e-3



LR = 5e-3



LR = 1e-3



LR = 5e-4

