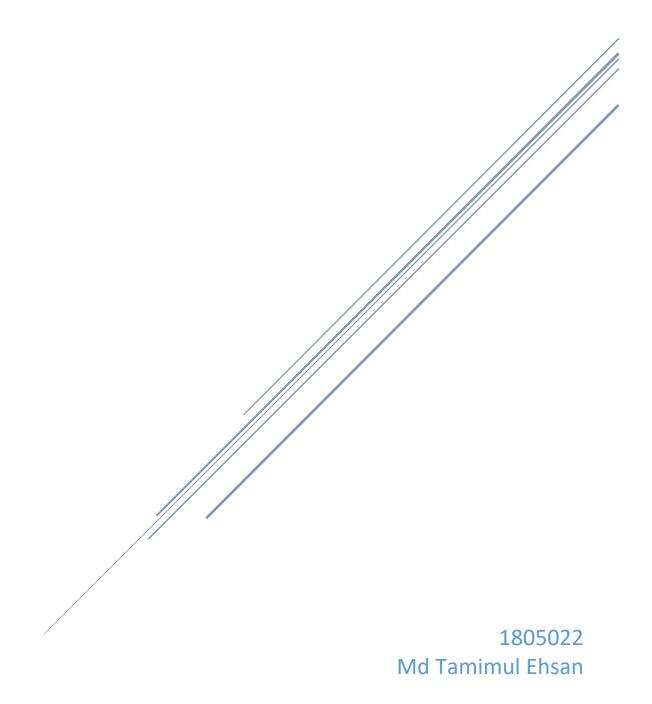
FNN FROM SCRATCH

CSE 472 Offline 3



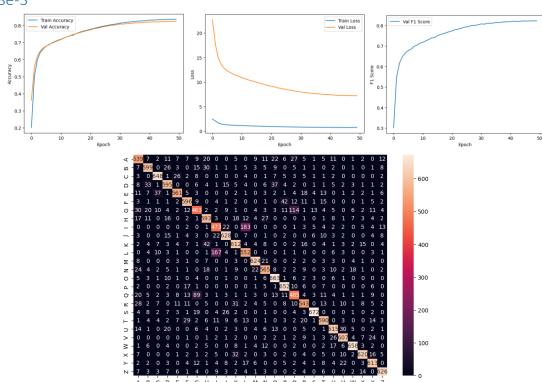
Hyperparameters:

- 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. Linear LR Scheduler
- 7. Default Adam Parameters
- 8. He Initialization

Model 1:

```
Dense(28 * 28, 256),
ReLU(),
Dense(256, 26),
Softmax()
],
```

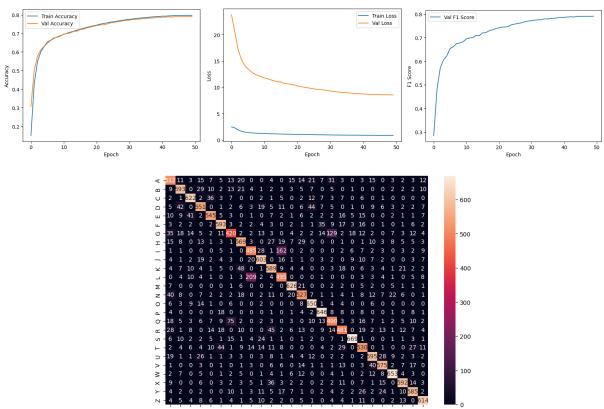
LR = 8e-3



Test Accuracy: 82.69711538461539 %

Test f1-macro score: 82.65141266486083 %

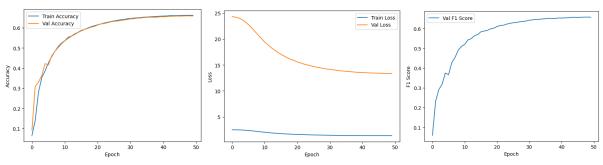


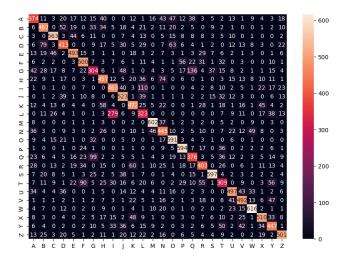


Test accuracy: 79.16826923076923 %

Test f1-macro score: 79.10643150377055 %

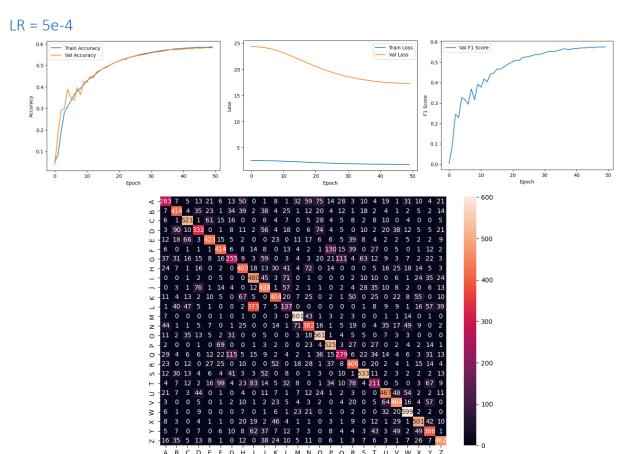
LR = 1e-3





Test accuracy: 66.15865384615385 %

Test f1-macro score: 65.94972808300558 %



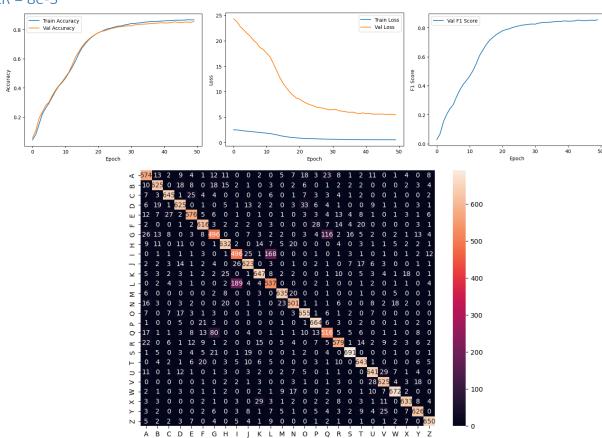
Test accuracy: 58.7548076923077 %

Test f1-macro score: 58.01596721494972 %

Model 2(Best Model):

```
[
    Dense(28*28, 512),
    ReLU(),
    Dropout(0.8),
    Dense(512, 256),
    ReLU(),
    Dropout(0.8),
    Dense(256, 26),
    Softmax()
],
```

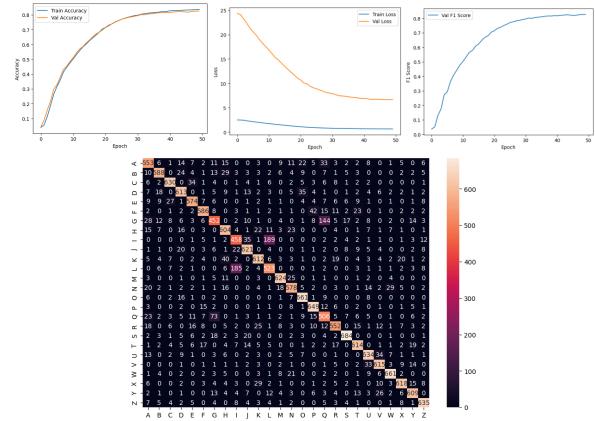
LR = 8e-3



Test accuracy: 85.1201923076923 %

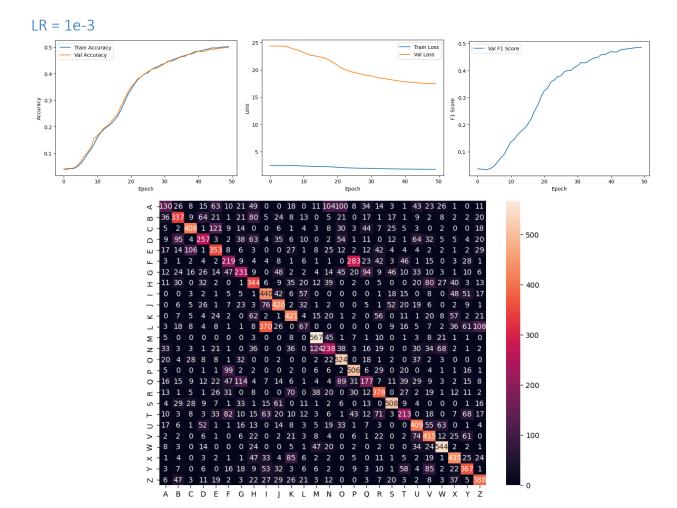
Test f1-macro score: 85.09334133852762 %





Test accuracy: 82.54326923076924 %

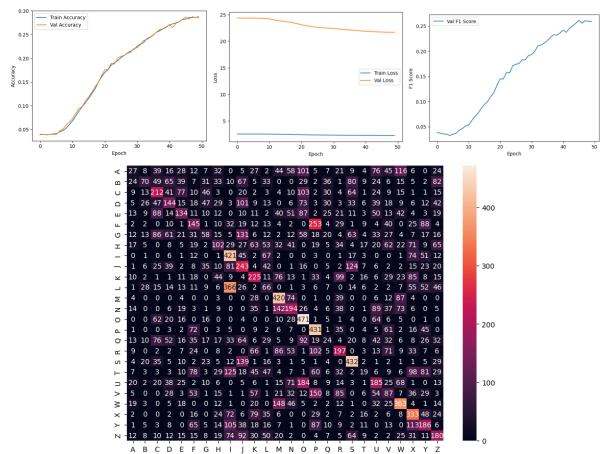
Test f1-macro score: 82.50098869480294 %



Test accuracy: 50.03846153846154 %

Test f1-macro score: 48.659485235266295 %

LR = 5e-4

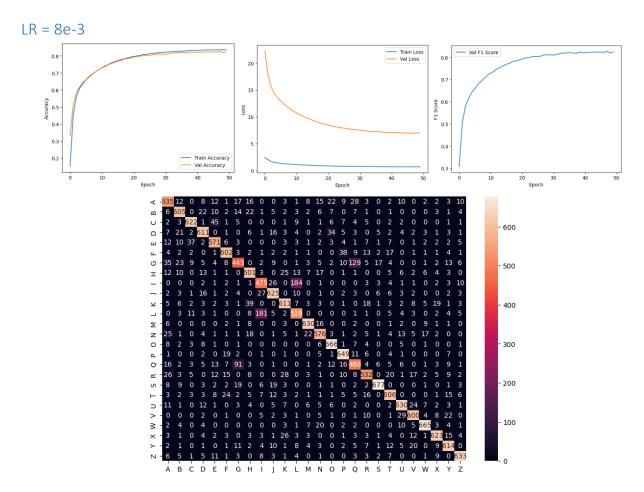


Test accuracy: 28.60576923076923 %

Test f1-macro score: 25.72176785953546 %

Model 3:

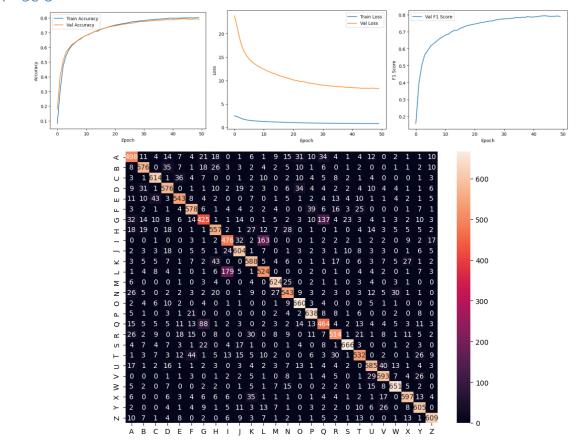
```
Dense(28*28, 256),
ReLU(),
Dropout(0.8),
Dense(256, 26),
Softmax()
```



Test accuracy: 82.84615384615385 %

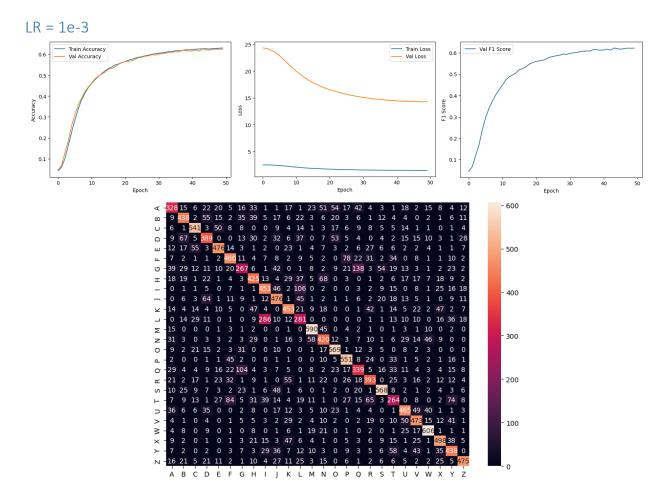
Test f1-macro score: 82.79709319179057 %

LR = 5e-3



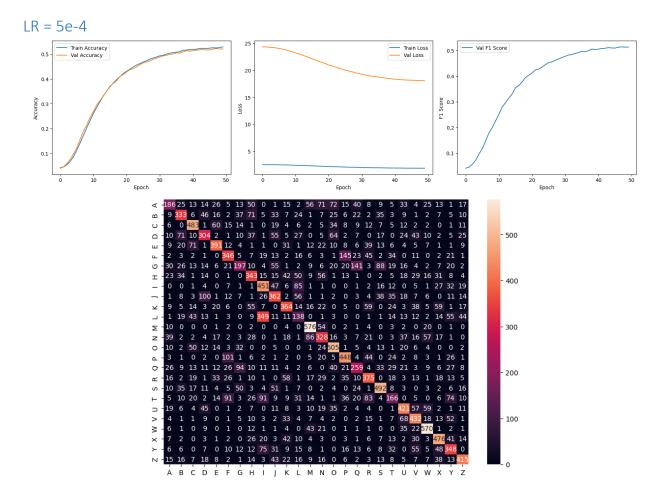
Test accuracy: 79.54326923076923 %

Test f1-macro score: 79.4672237202777 %



Test accuracy: 62.73076923076923 %

Test f1-macro score: 62.39666940664379 %



Test accuracy: 52.67307692307692 %

Test f1-macro score: 51.84128860828774 %