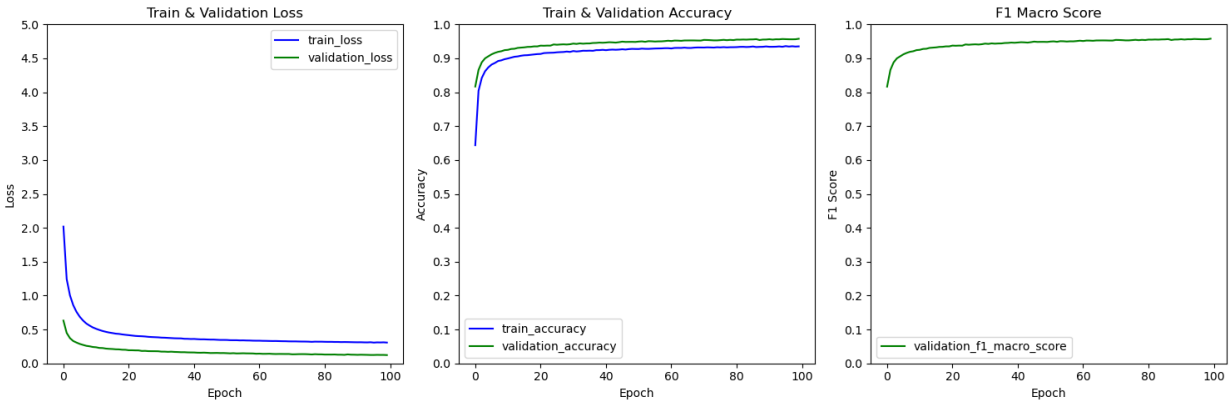
CSE 472: Offline 3 Report

Student ID: 1805115

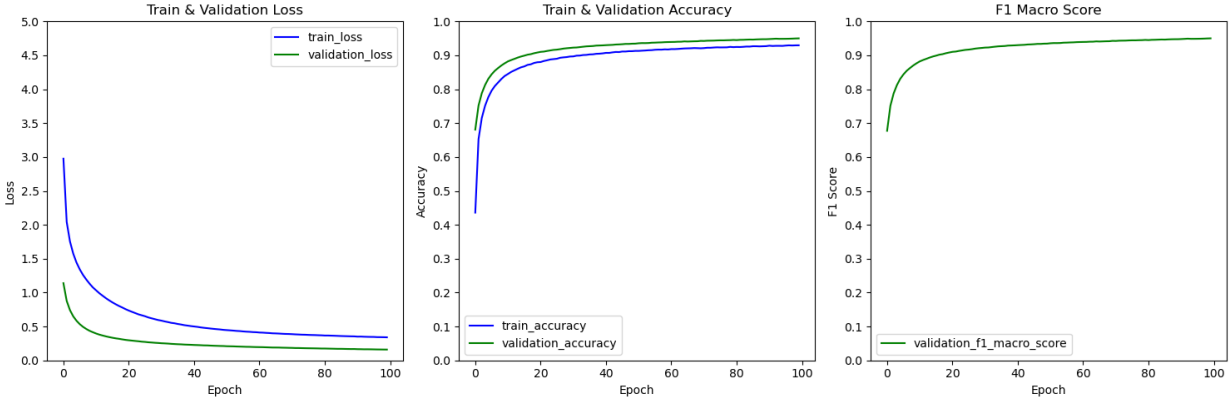
Model 1: Dense1 (784, 784) -> ReLU -> Dropout (0.3) -> Dense2(784, 256) -> ReLU -> Dropout(0.2) -> Dense(256, 26) -> Softmax

Minibatch size: 624

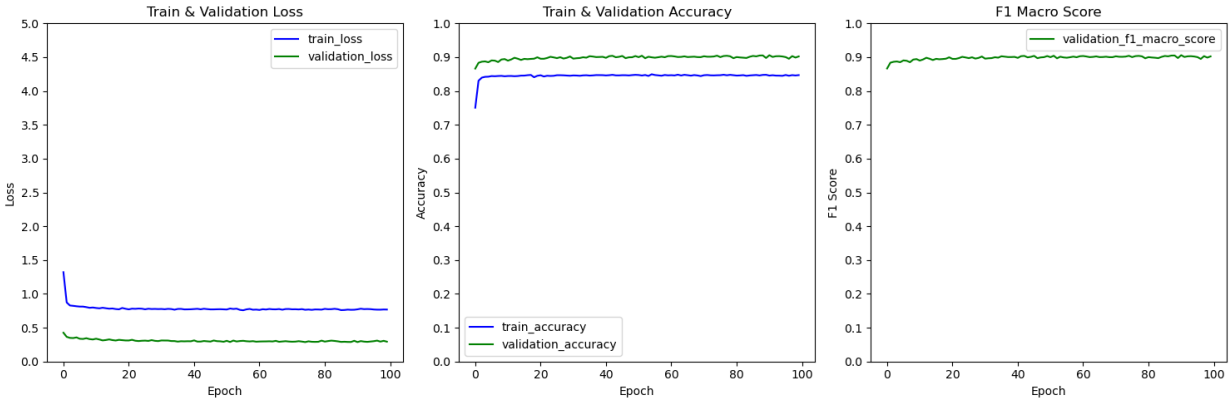
Optimizer: Adam, learning rate decay: 5e-7



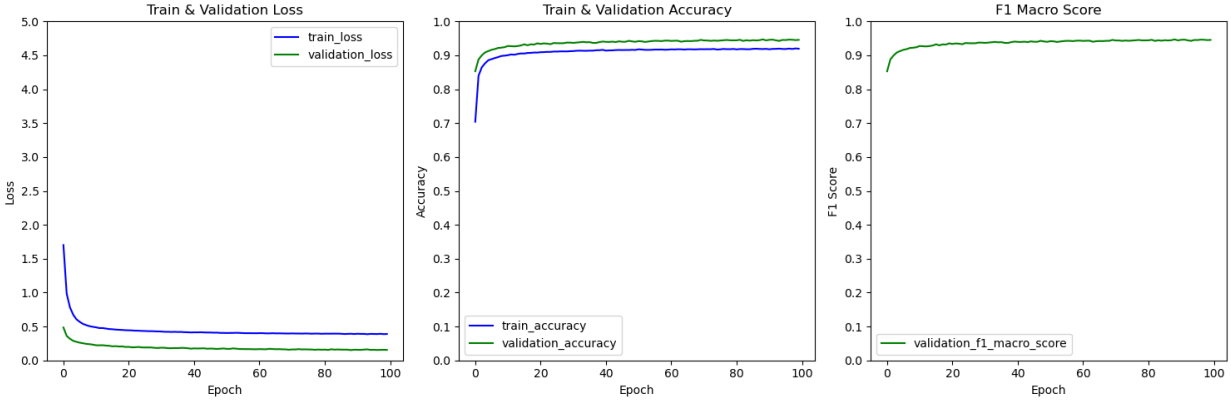
Learning rate = 0.0005



Learning rate = 0.0001



Learning rate = 0.005

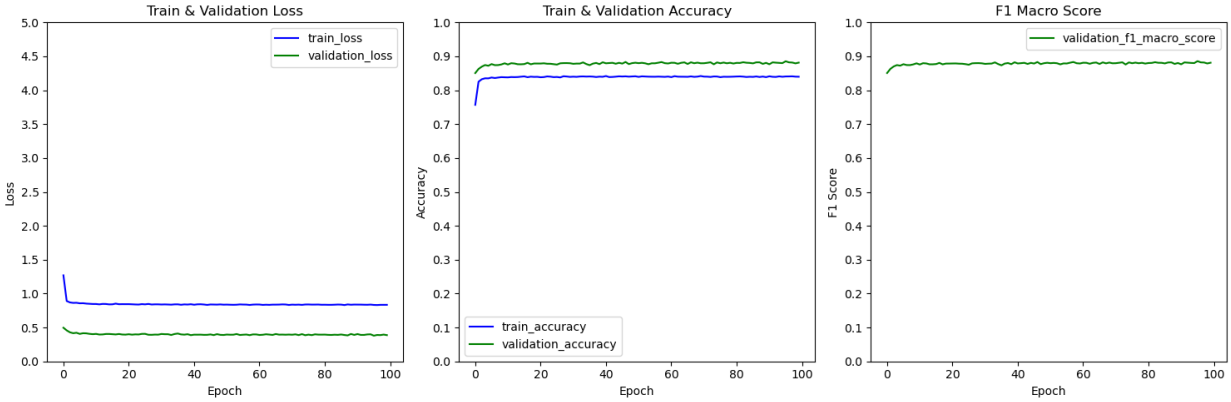


Learning rate = 0.001

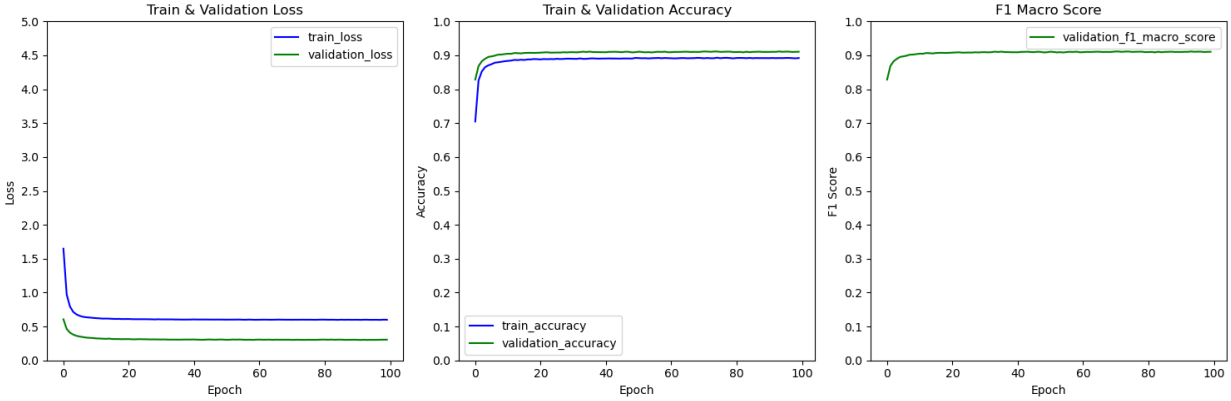
Model 2: Dense1 (784, 1024) -> ReLU -> Dropout (0.4) -> Dense2(1024, 26) -> Softmax

Minibatch size: 624

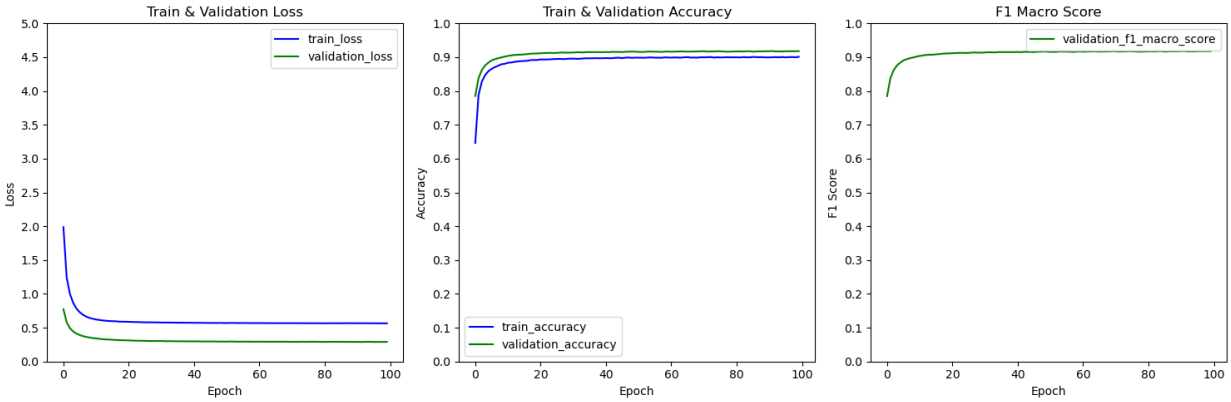
Optimizer: Adam, learning rate decay: 5e-7



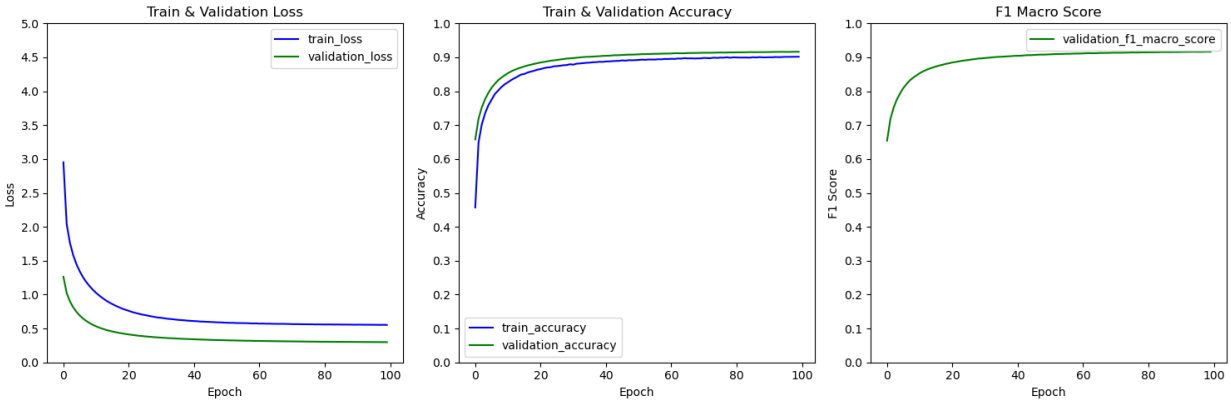
Learning Rate = 0.005



Learning Rate = 0.001



Learning Rate = 0.0005

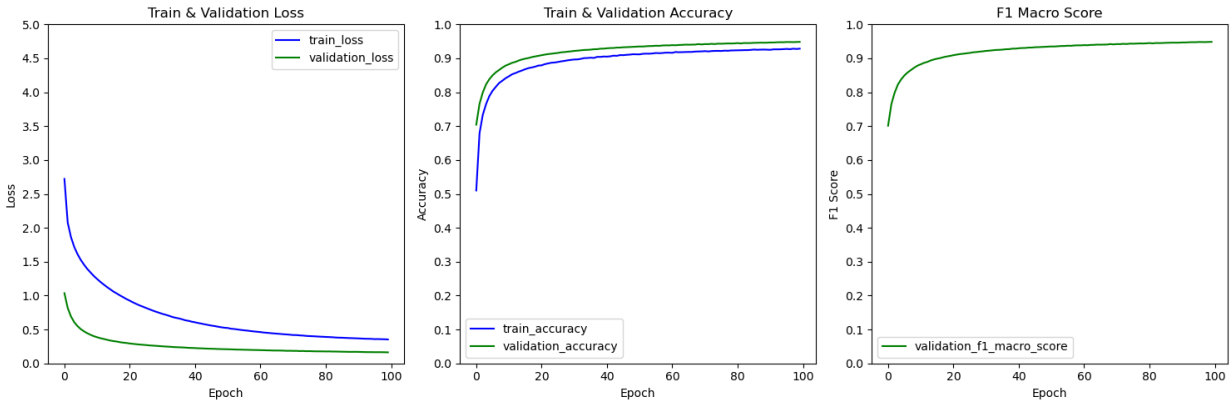


Learning Rate = 0.0001

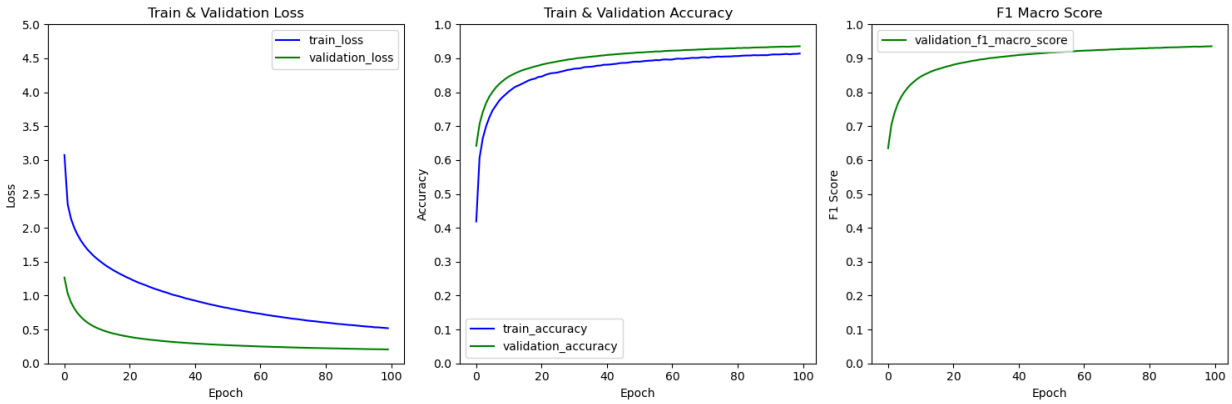
Model 1: Dense1 (784, 784) -> ReLU -> Dropout (0.3) -> Dense2(784, 256) -> ReLU -> Dropout(0.2) -> Dense(256, 26) -> Softmax

Minibatch size: 624

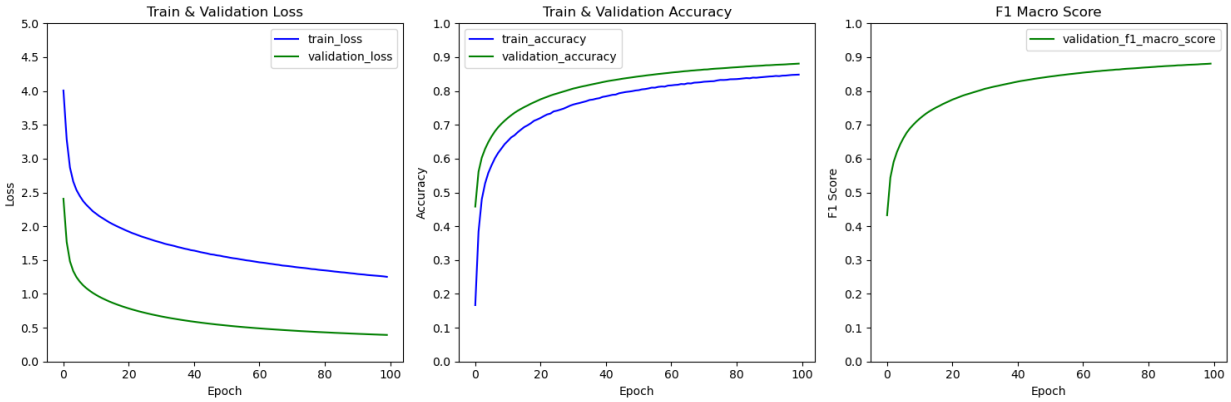
Optimizer: Gradient Descent , learning rate decay: 5e-7



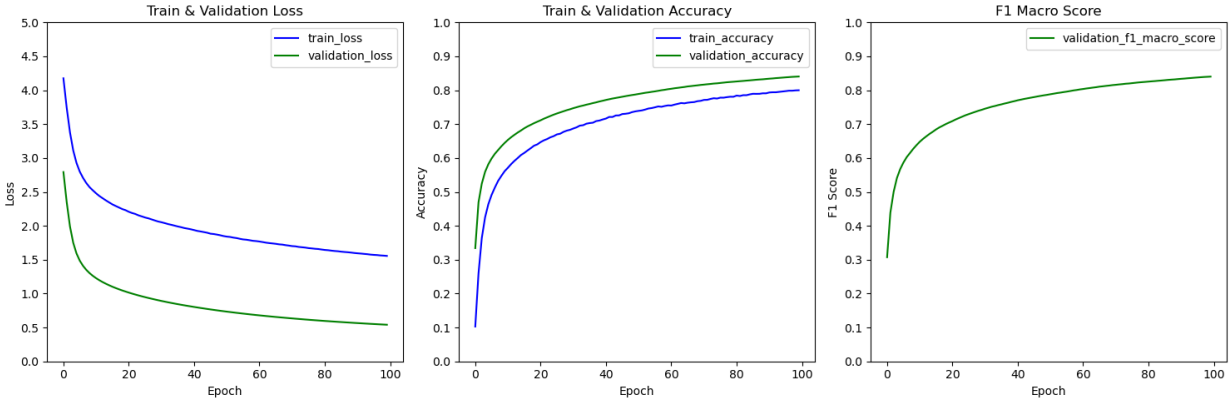
Learning Rate = 0.1



Learning Rate = 0.05



Learning Rate = 0.001



Learning Rate = 0.005

The Best Model

Dense1 (784, 784) -> ReLU -> Dropout (0.3) -> Dense2(784, 256) -> ReLU -> Dropout(0.2) -> Dense(256, 26) -> Softmax

Minibatch size: 624

Optimizer: Adam, learning rate decay: 5e-7

Learning Rate = 0.0005

Independent test accuracy: 92.68%, **f1 macro score**: 92.69%

Independent test loss: 22.5%

Confusion Matrix report:

label	TP:	TN:	FP:	FN:
0	736	19930	70	64
1	757	19974	26	43
2	743	19978	22	57
3	756	19929	71	44
4	763	19951	49	37
5	766	19962	38	34
6	638	19903	97	162
7	750	19954	46	50
8	623	19748	252	177
9	742	19967	33	58
10	749	19972	28	51
11	591	19812	188	209
12	783	19978	22	17
13	763	19944	56	37
14	786	19946	54	14
15	781	19964	36	19
16	674	19864	136	126
17	751	19968	32	49
18	779	19978	22	21
19	773	19948	52	27
20	750	19955	45	50
21	749	19958	42	51
22	768	19990	10	32
23	769	19967	33	31
24	758	19963	37	42
25	780	19975	25	20