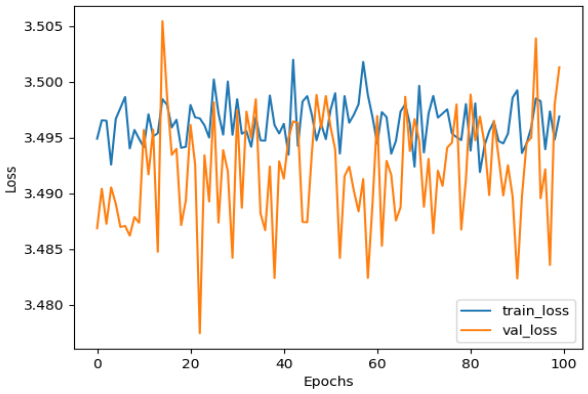
CSE 472: Offline 3 report

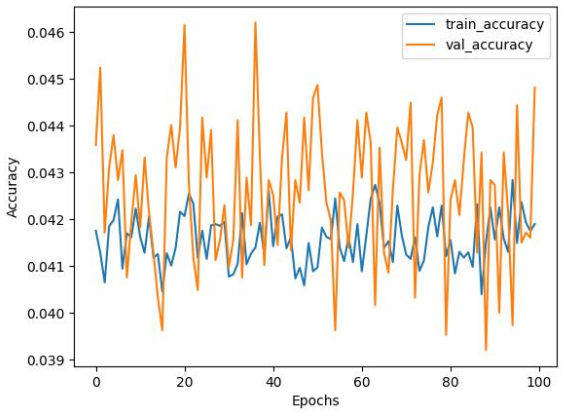
Student id: 1805092

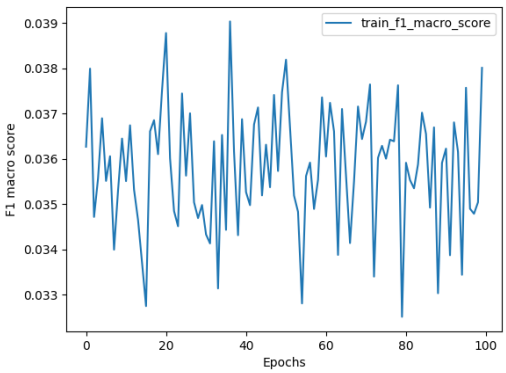
Model 1: Dense1 (784, 100) -> ReLU -> Dropout (0.3) -> Dense2(100, 26) -> Softmax

Minibatch size: 221

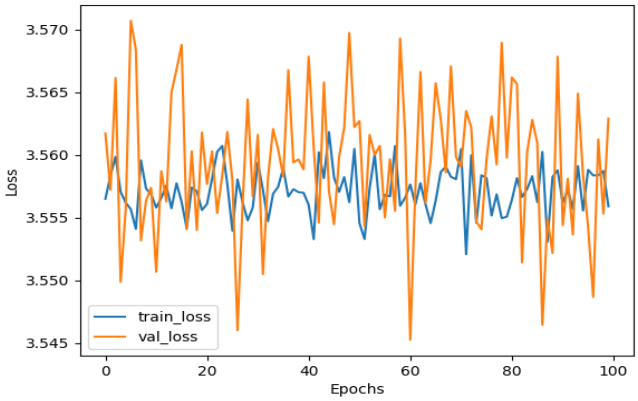
Optimizer: Adam, learning rate decay: 5e-7

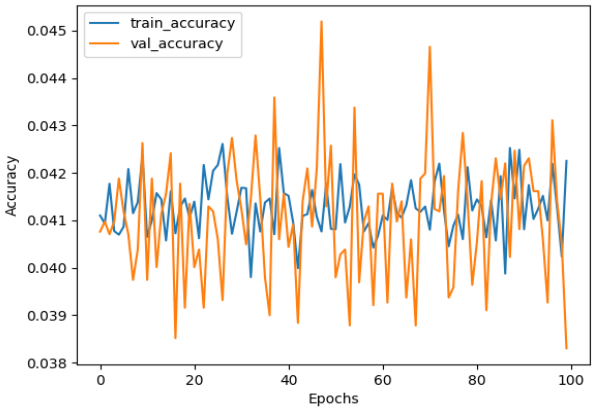


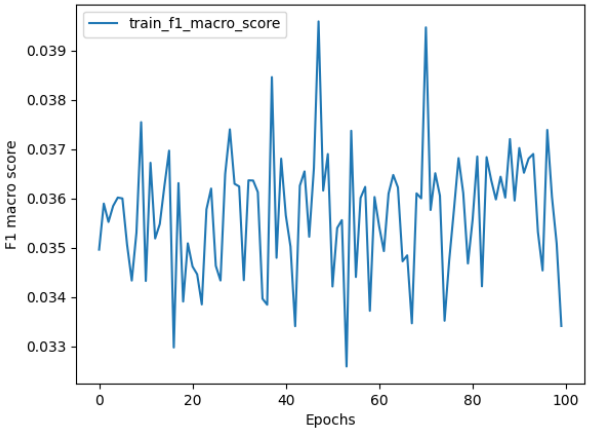




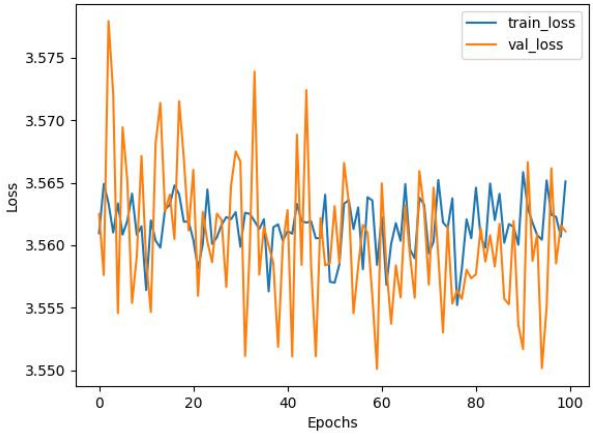
learning rate: 0.05

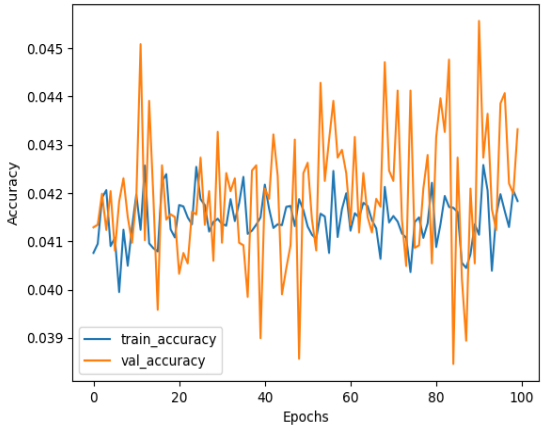


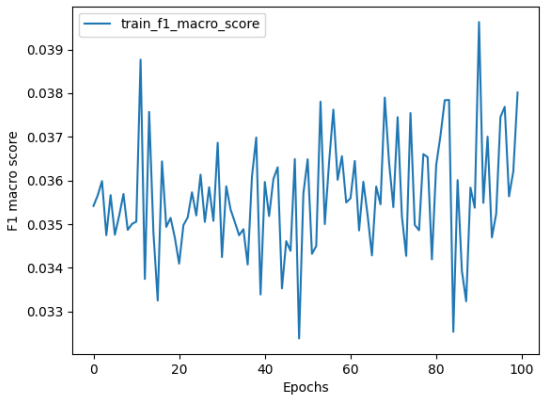




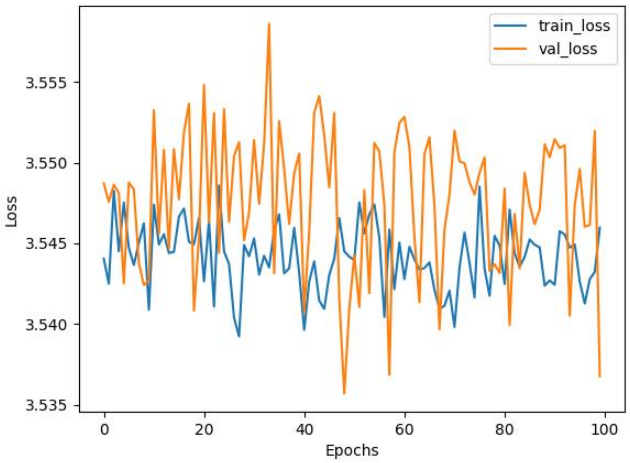
Learning rate : 0.01

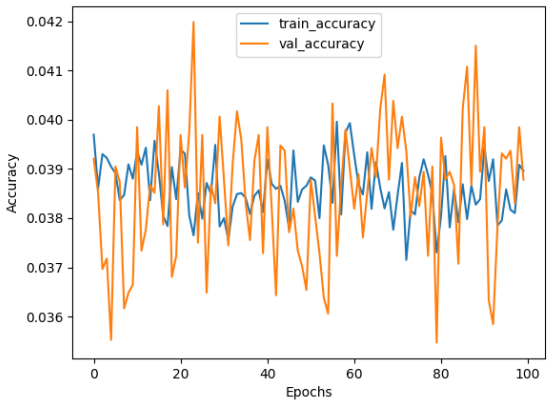


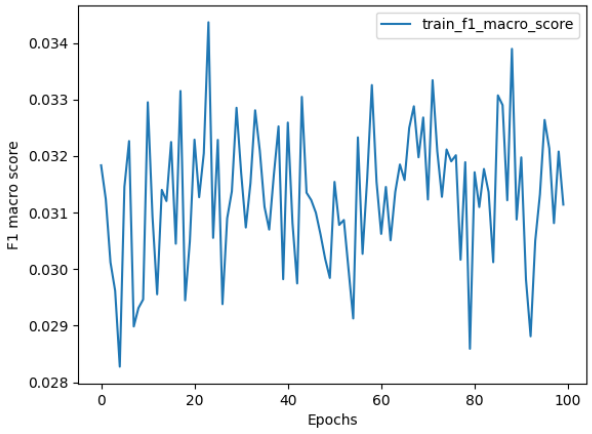




Learning rate: 0.0005





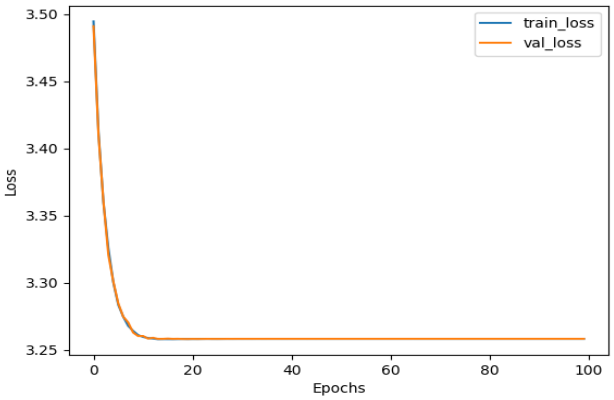


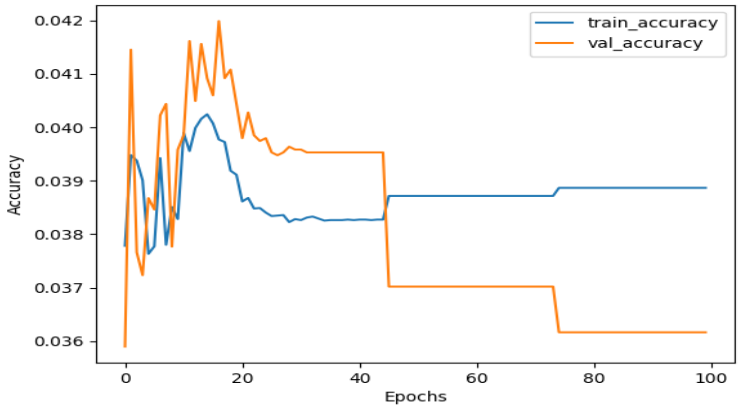
Learning rate = 0.0001

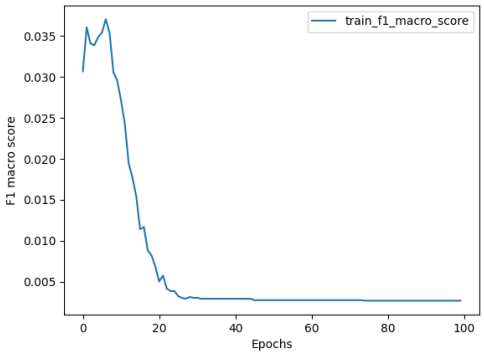
Model 2: Dense1 (784, 1248) -> ReLU -> Dropout (0.3) -> Dense2(1248, 26) -> Softmax

Minibatch size: randomly chosen from 512

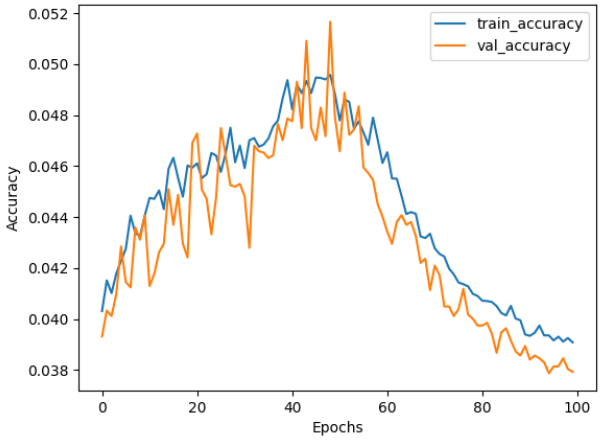
Optimizer: Stochastic Gradient Descent

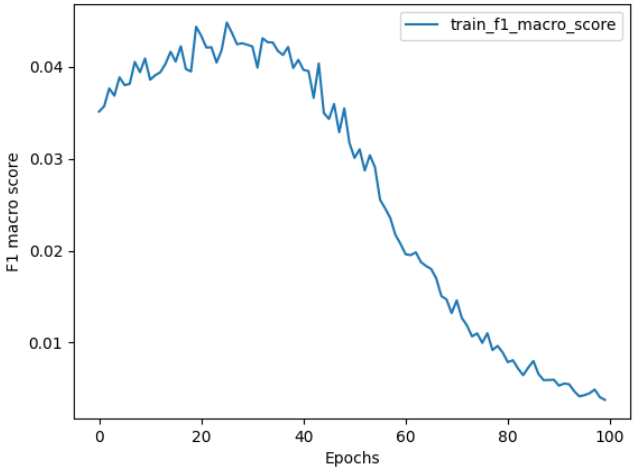


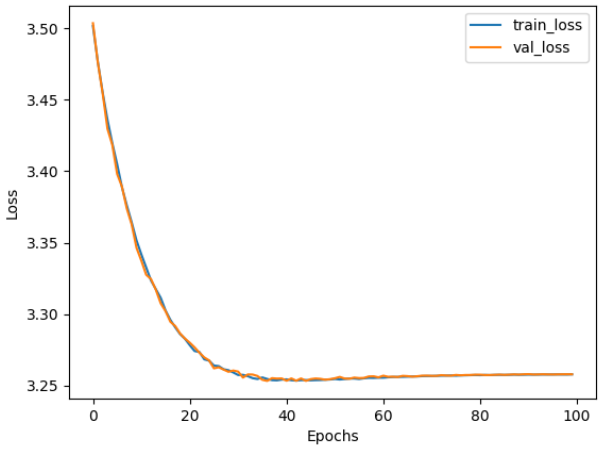




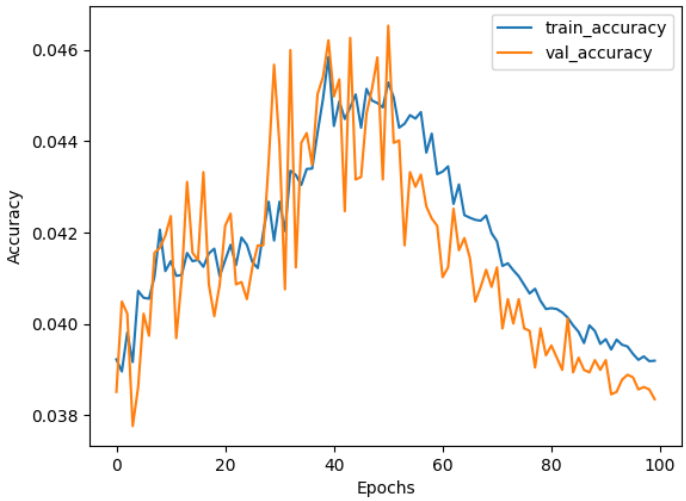
Learning rate =0.005

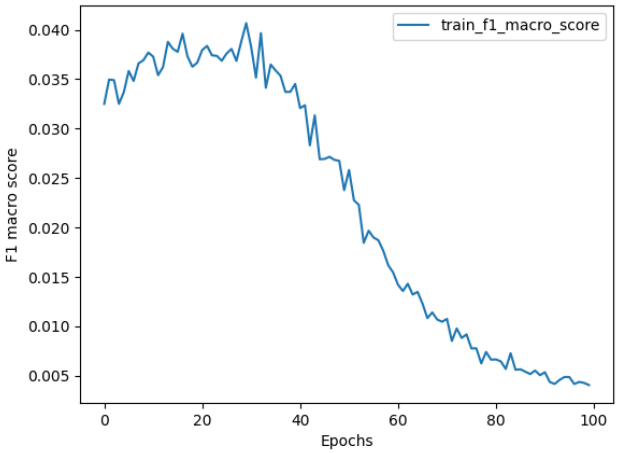


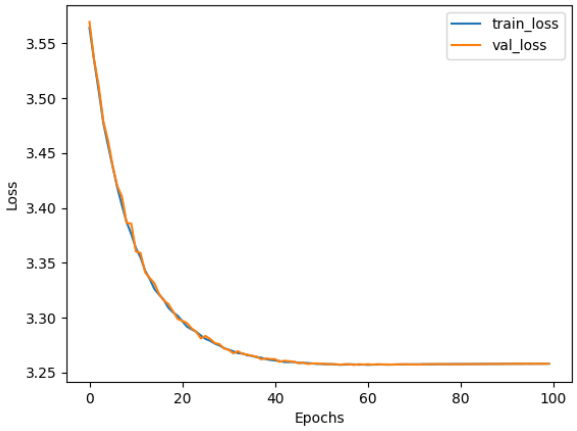




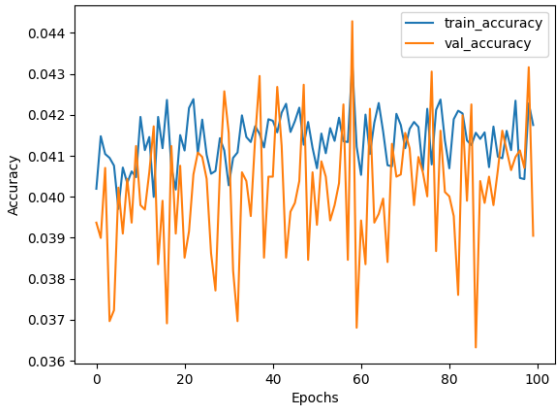
Learning rate = 0.001

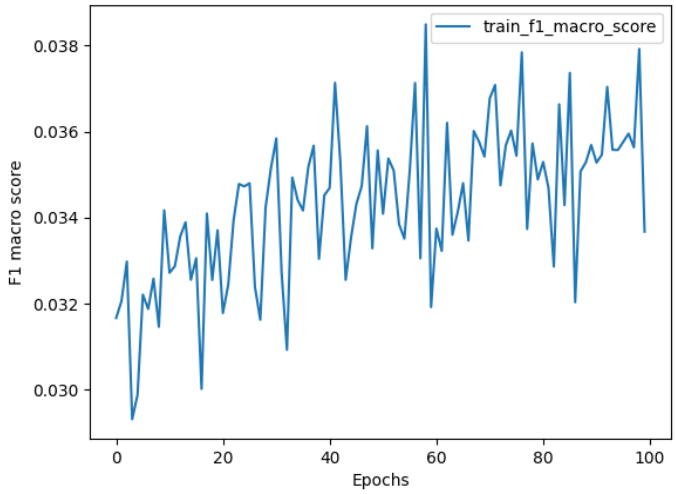


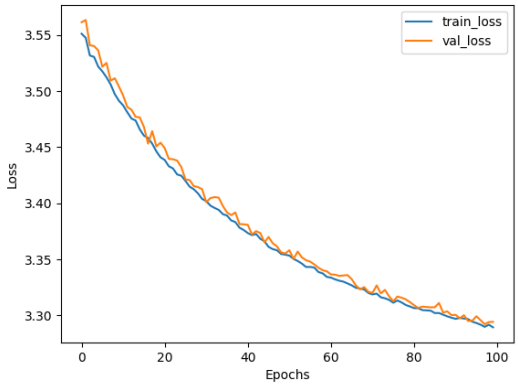




Learning rate = 0.0005





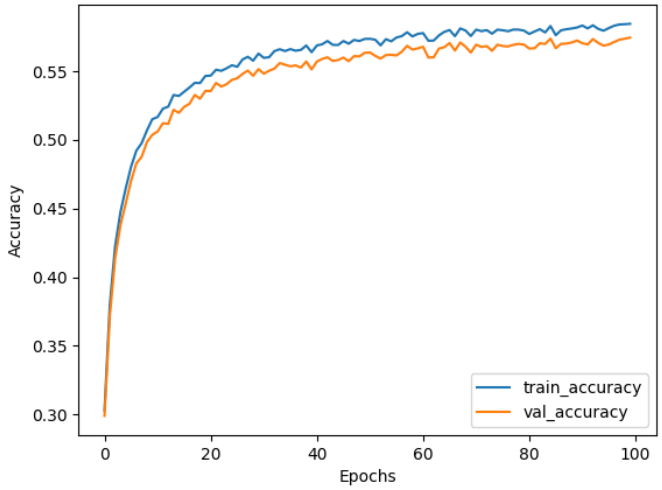


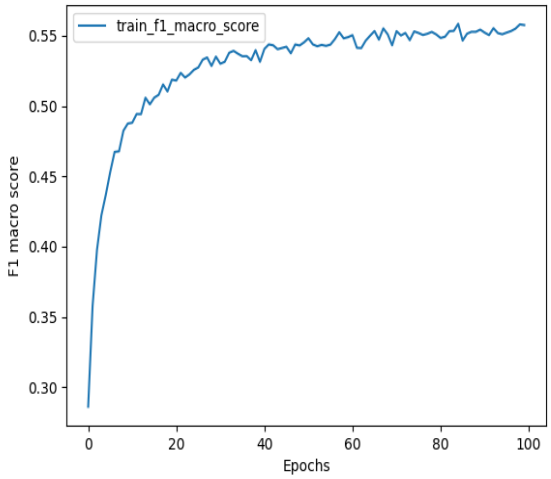
Learning rate= 0.0001

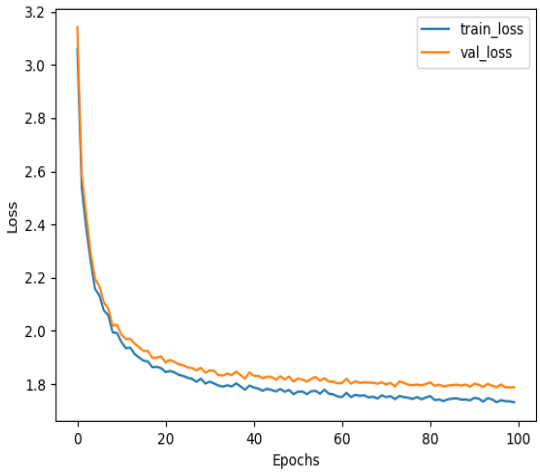
Model 3: Dense1 (784, 100) -> Dense2(100, 26)

Minibatch size: 221

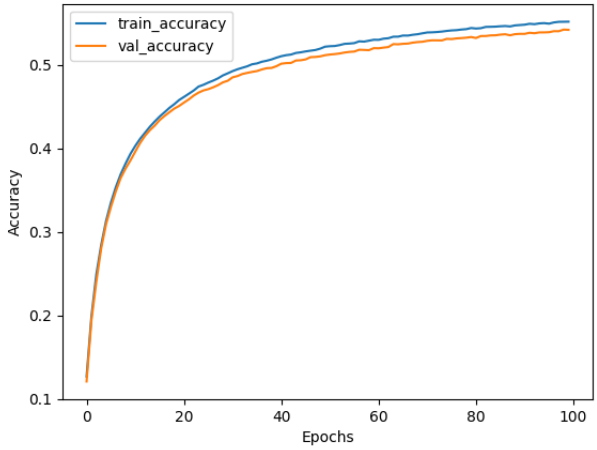
Optimizer: Adam, learning rate decay: 5e-7

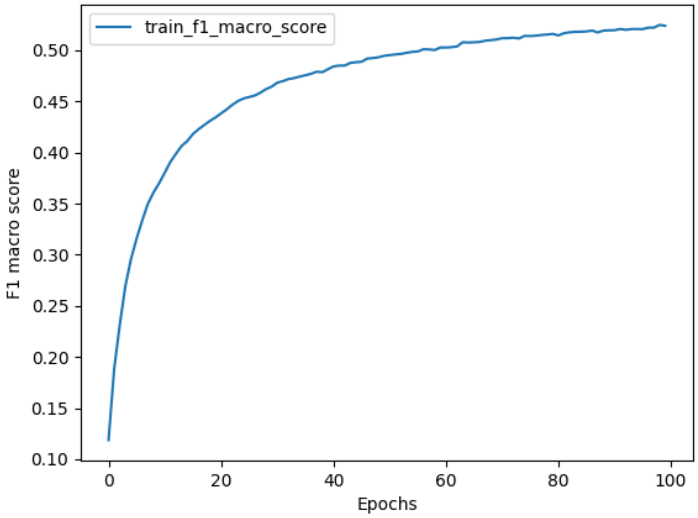


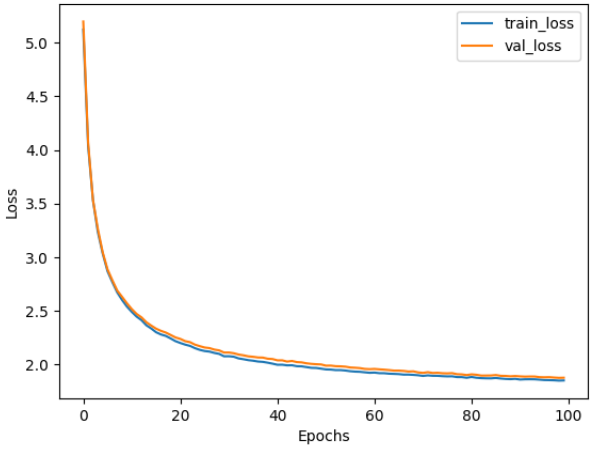




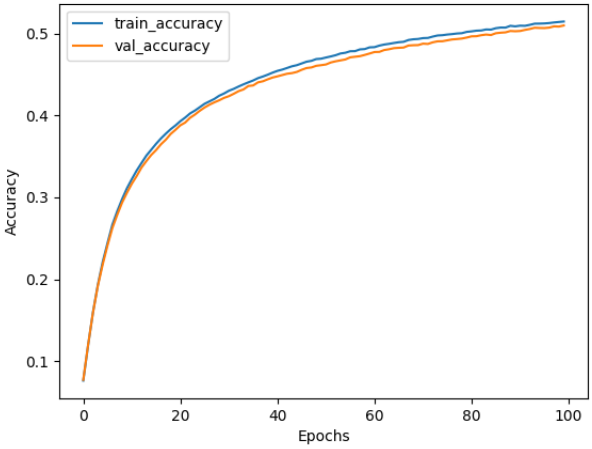
Learning rate= 0.005

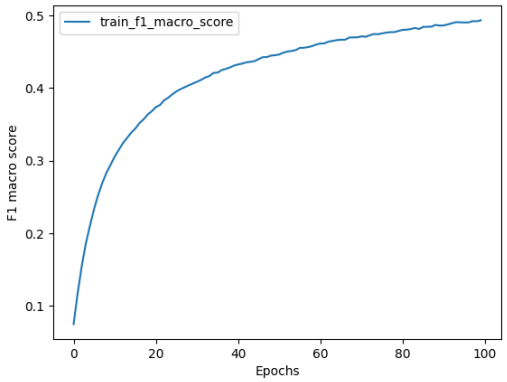


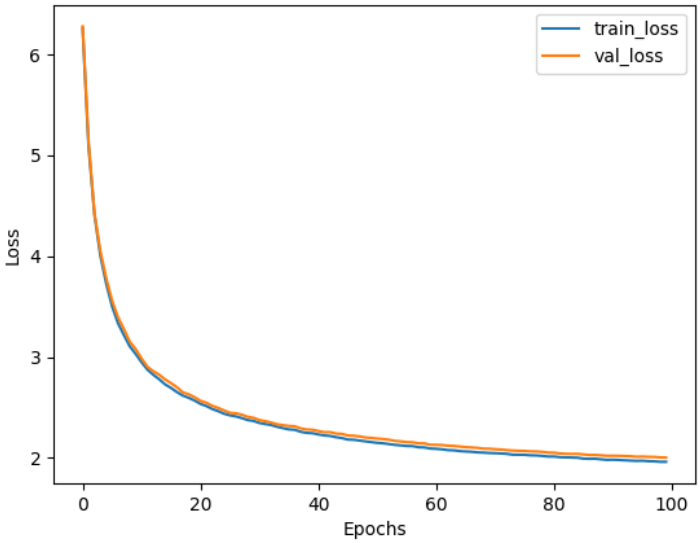




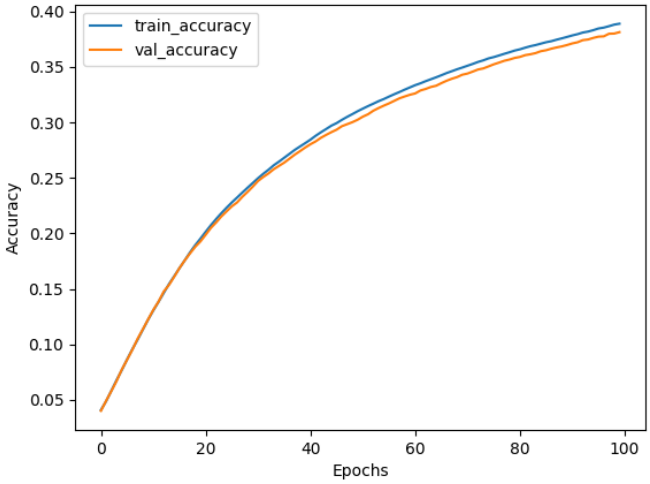
Learning rate = 0.001

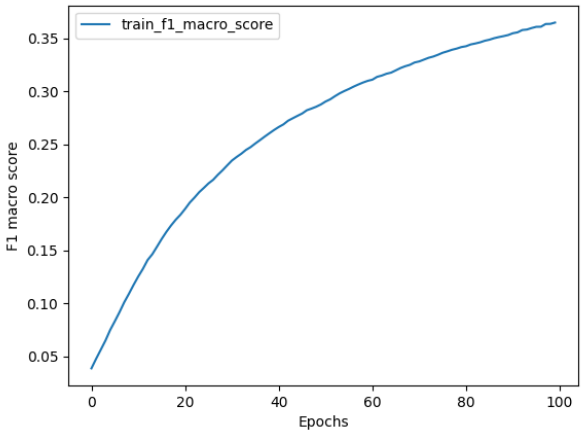


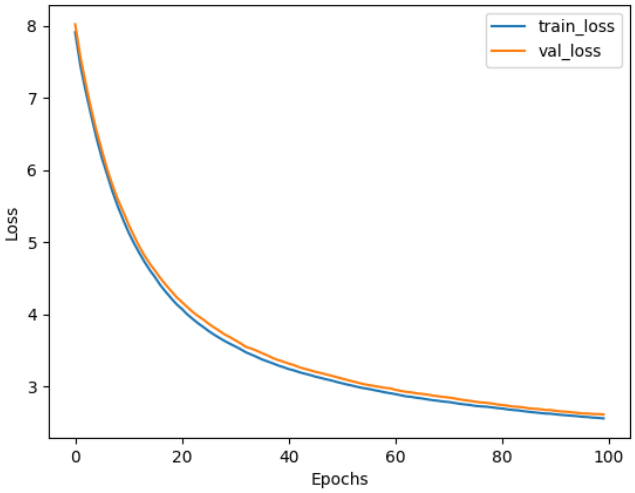




Learning rate = 0.0005







Learning rate= 0.0001

Best model: Model 3

Dense1 (784, 100) -> Dense2(100, 26)

Minibatch size: 221

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

Accuracy on test set: 55.7%