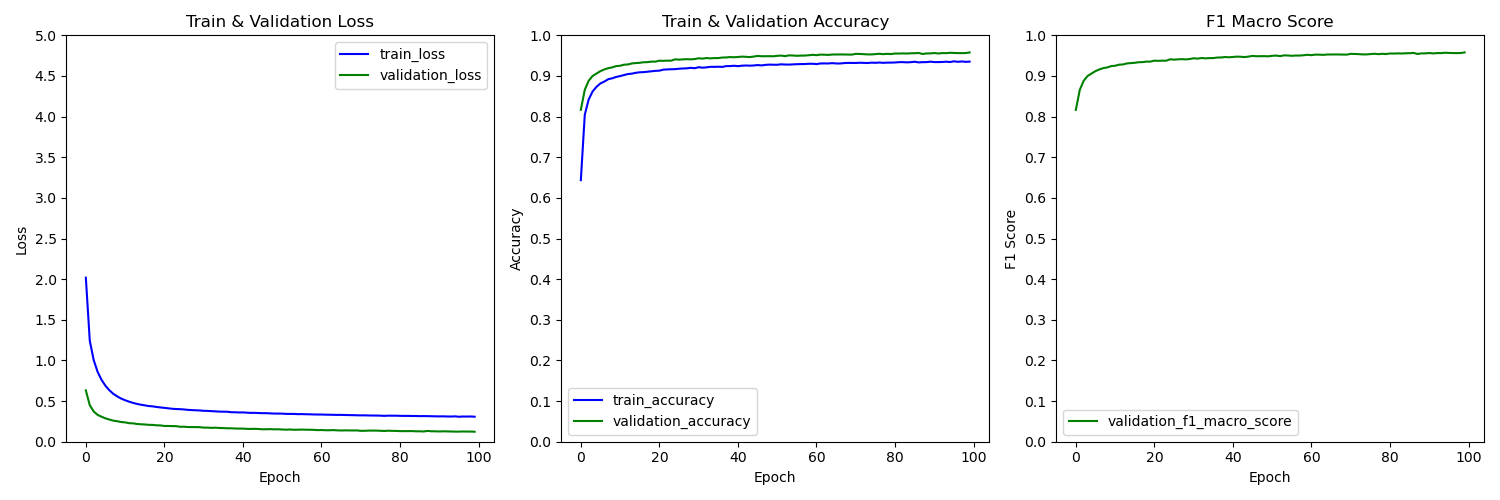
CSE 472: Offline 3 Report

Student ID: 1805115

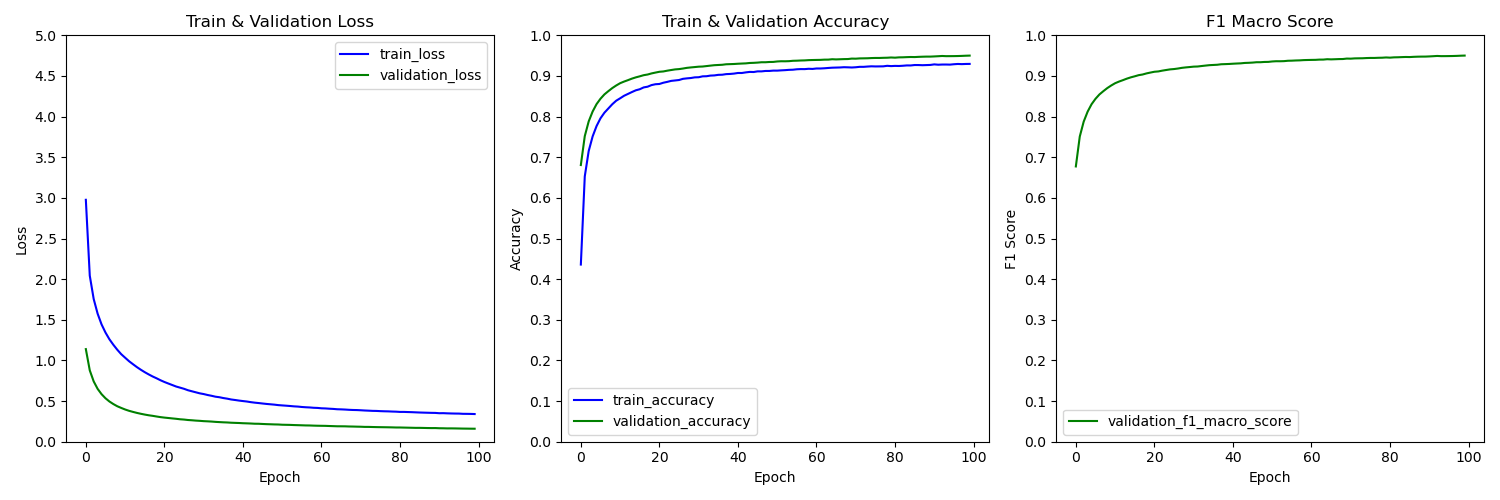
**Model 1: Dense1 (784, 784) -> ReLU -> Dropout (0.3) -> Dense2(784, 256) -> ReLU -> Dropout(0.2) -> Dense(256, 26) -> Softmax**

**Minibatch size: 624**

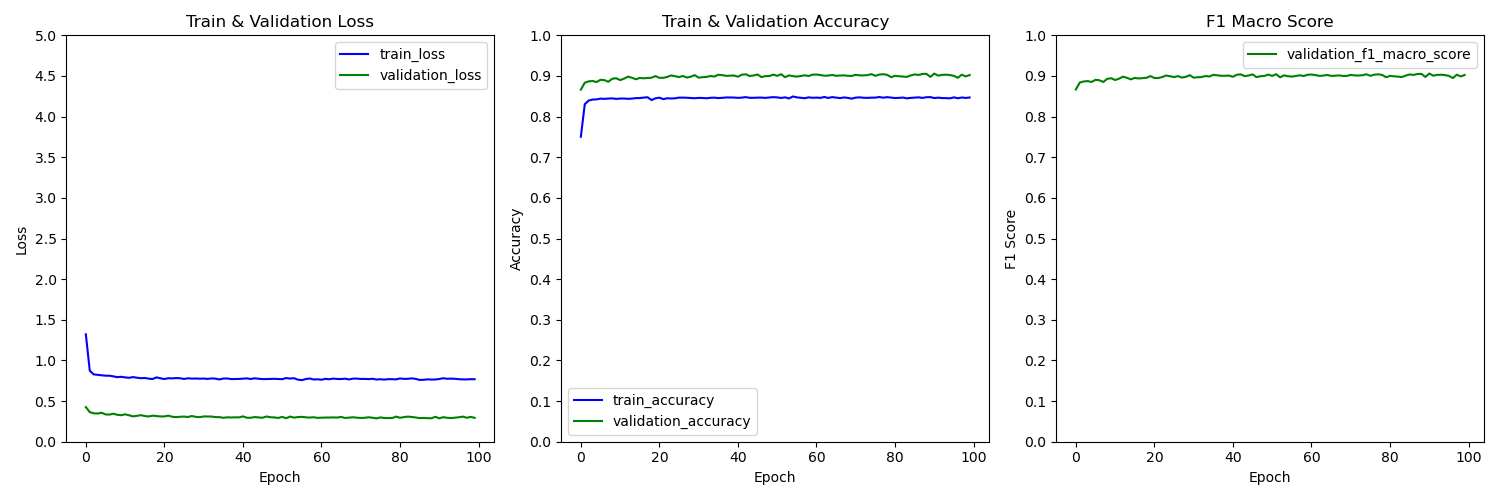
**Optimizer: Adam, learning rate decay: 5e-7**



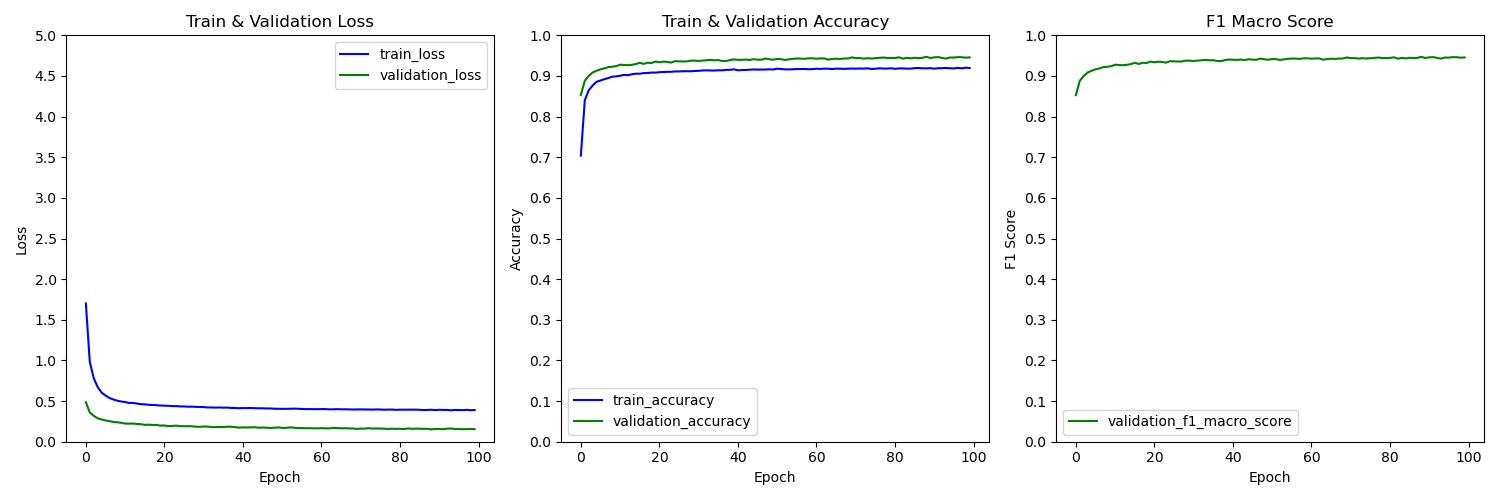
**Learning rate = 0.0005**

****

**Learning rate = 0.0001**

****

**Learning rate = 0.005**

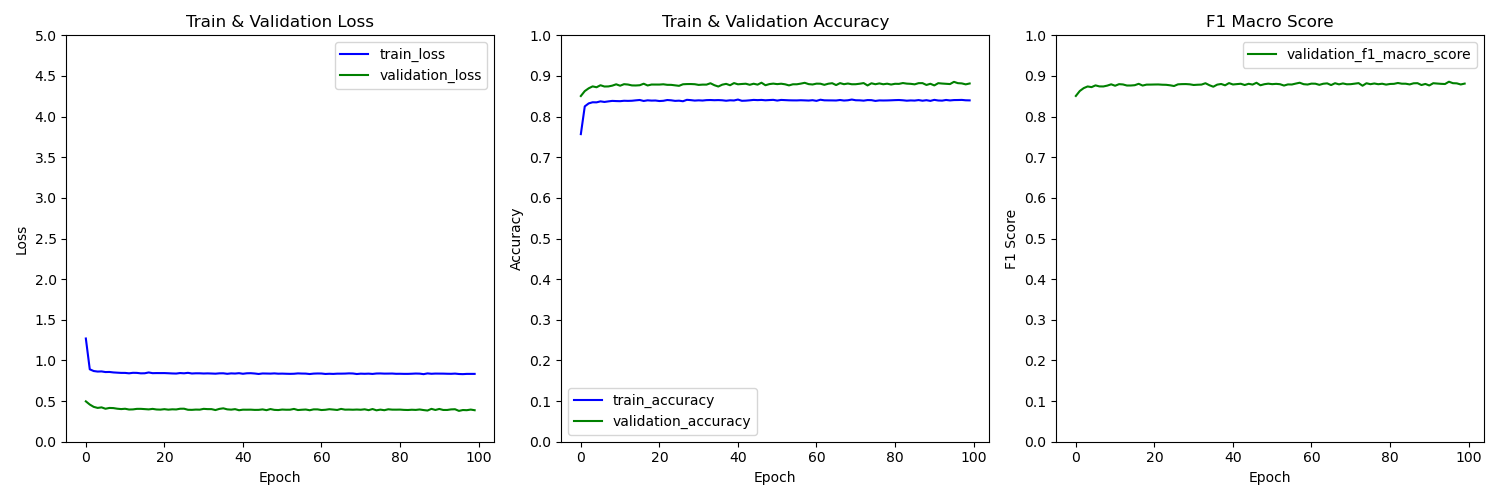
****

**Learning rate = 0.001**

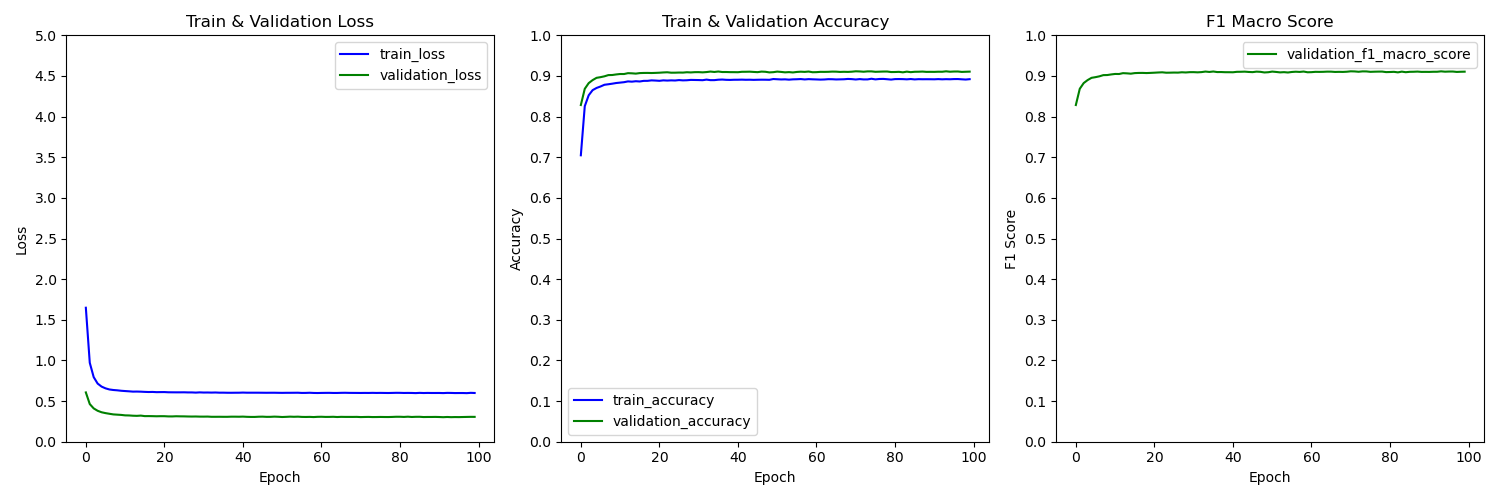
**Model 2: Dense1 (784, 1024) -> ReLU -> Dropout (0.4) -> Dense2(1024, 26) -> Softmax**

**Minibatch size: 624**

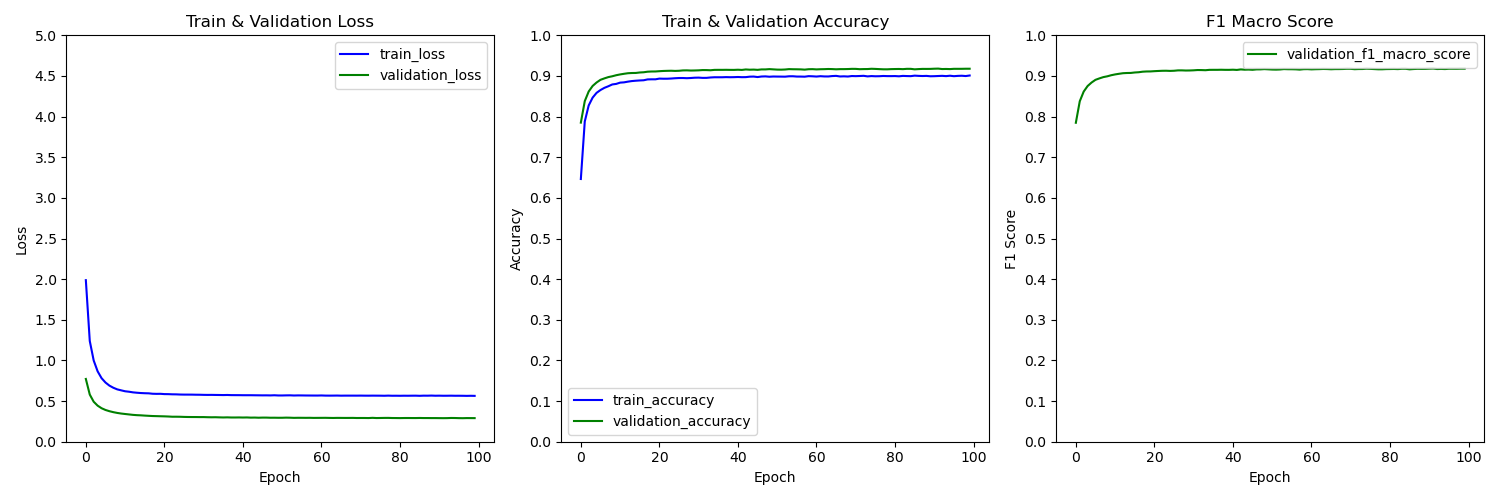
**Optimizer: Adam, learning rate decay: 5e-7**

****

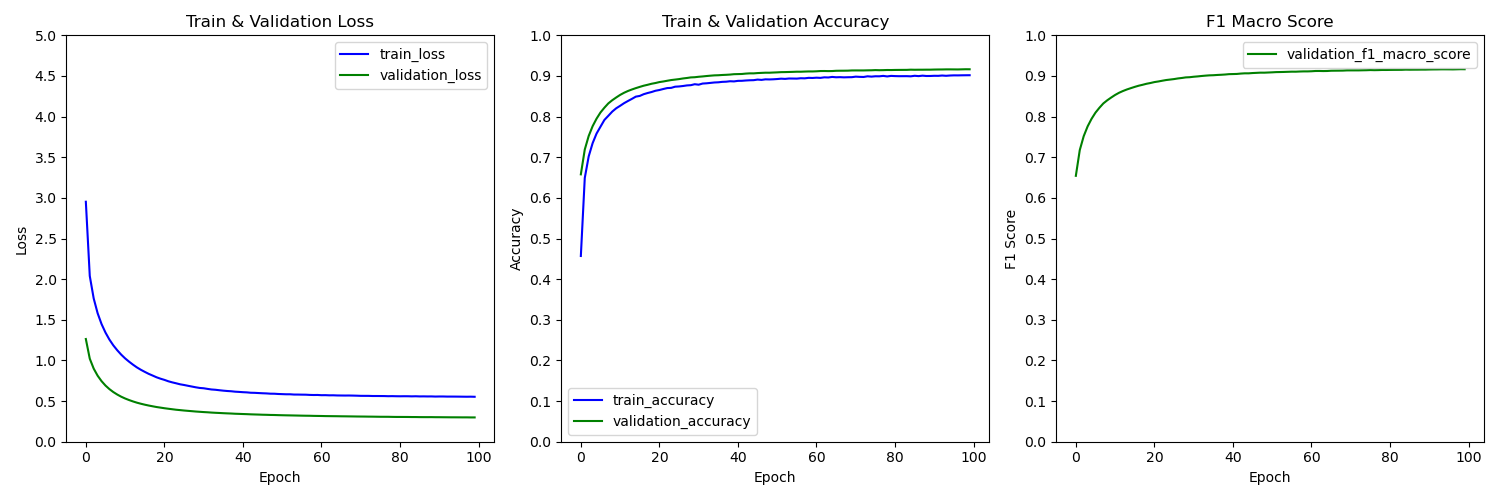
**Learning Rate = 0.005**

****

**Learning Rate = 0.001**

****

**Learning Rate = 0.0005**

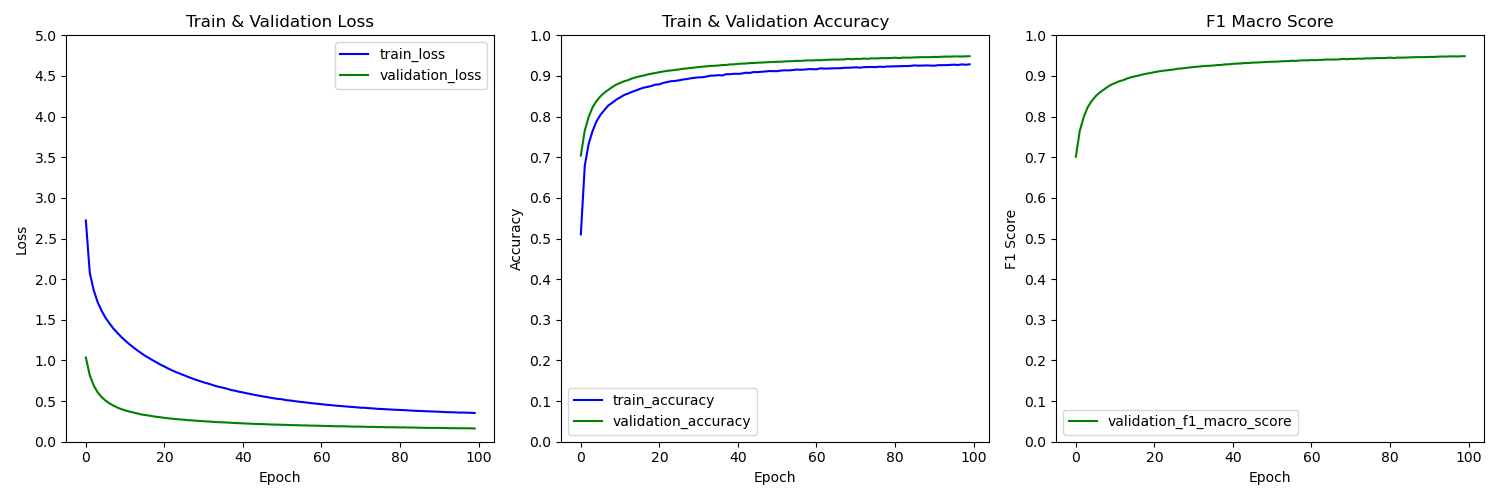
****

**Learning Rate = 0.0001**

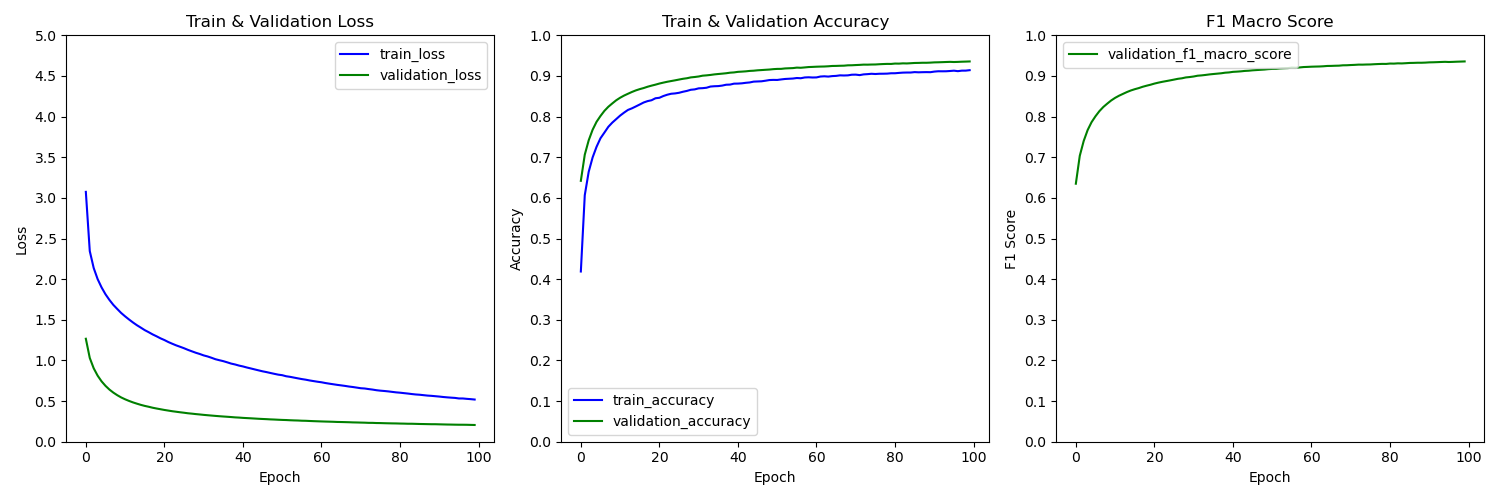
**Model 1: Dense1 (784, 784) -> ReLU -> Dropout (0.3) -> Dense2(784, 256) -> ReLU -> Dropout(0.2) -> Dense(256, 26) -> Softmax**

**Minibatch size: 624**

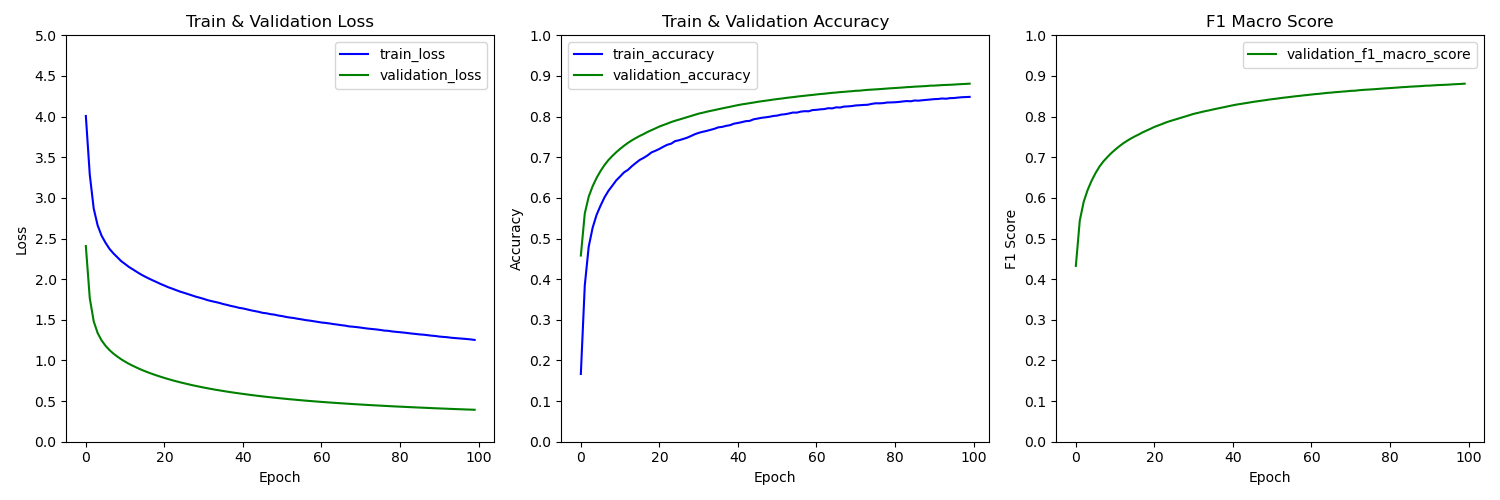
**Optimizer: Gradient Descent , learning rate decay: 5e-7**

****

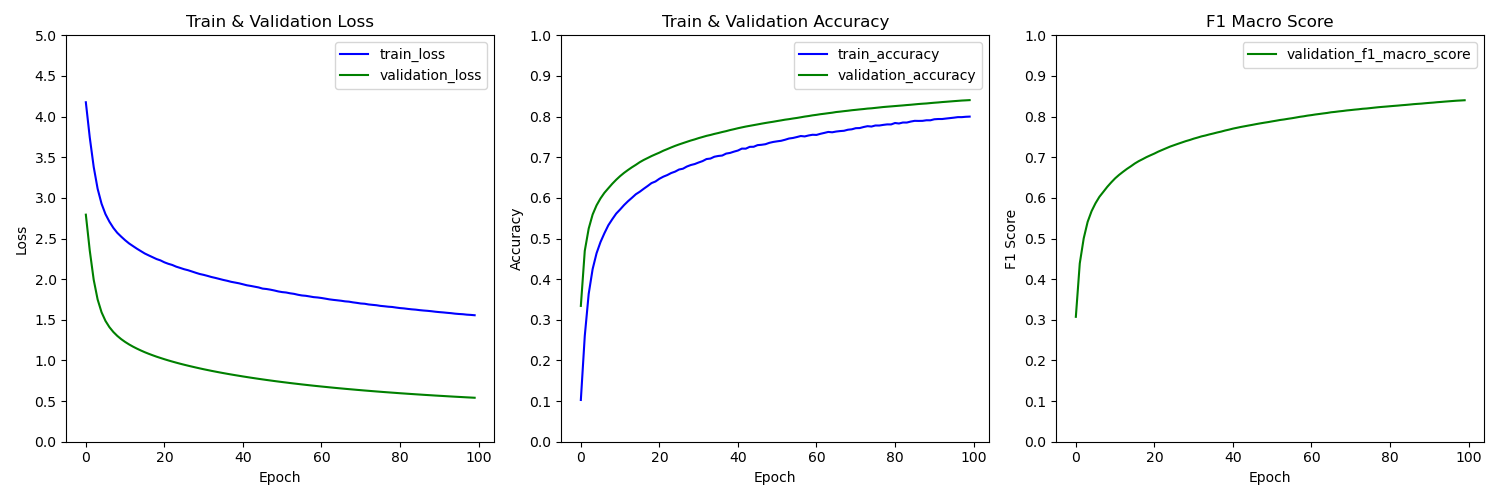
**Learning Rate = 0.1**

****

**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:**

