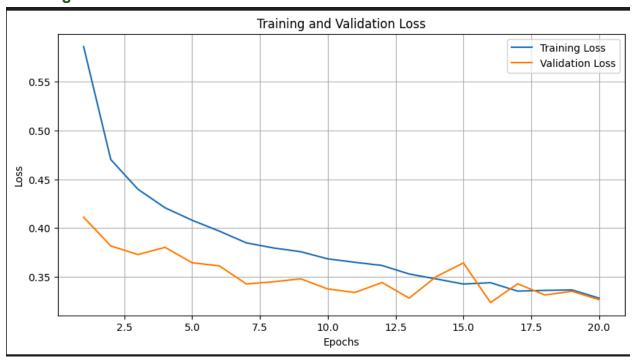
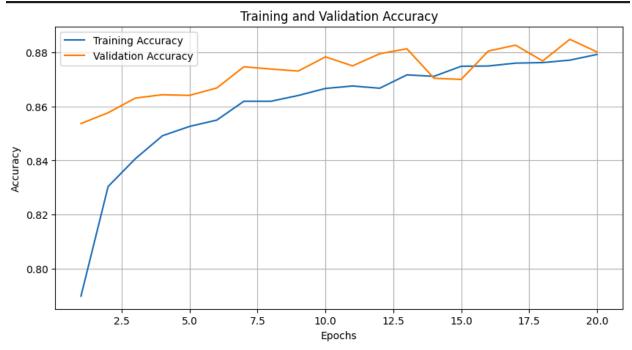
Models: <u>Model1 (One Layer, Hidden Dim = 128):</u>

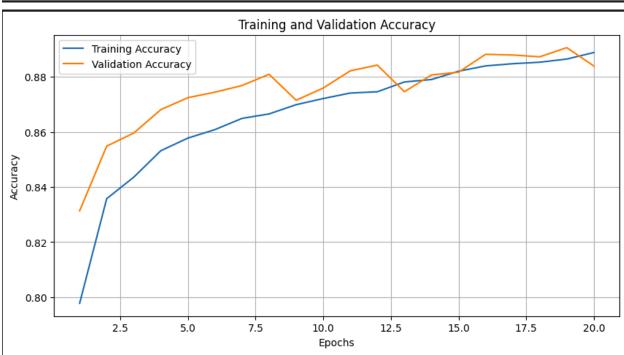
Learning rate 0.01:



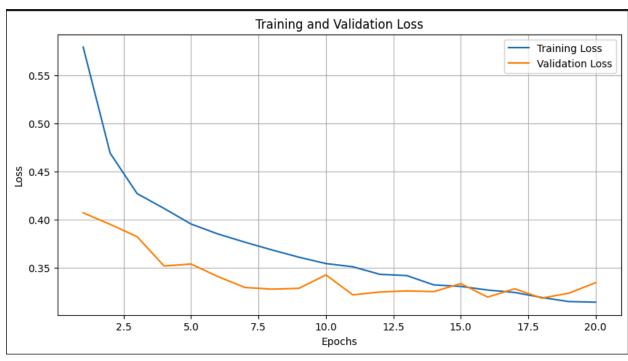


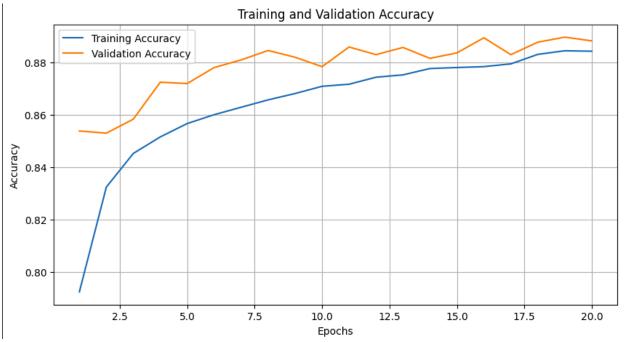
Learning rate 0.001:





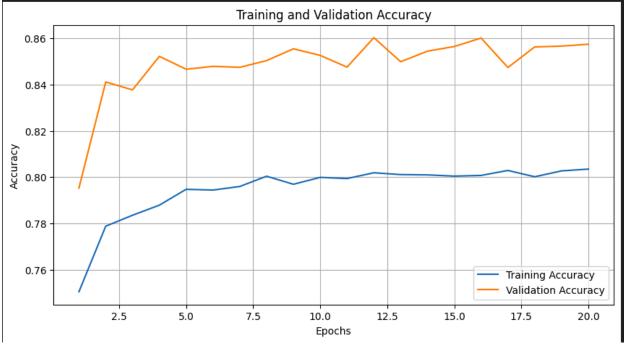
Learning rate 0.005:





Learning rate 0.05:

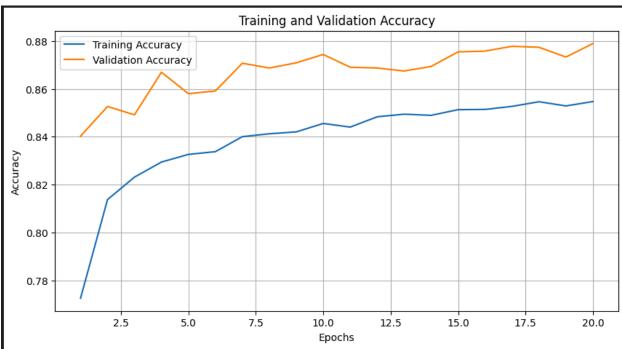




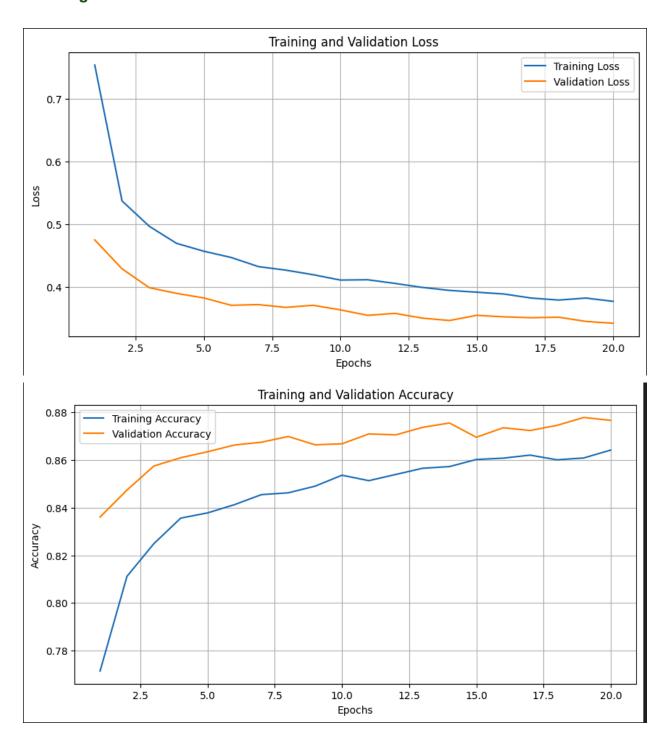
Model2 (One Layer, Hidden Dim = 64):

Learning rate 0.01:

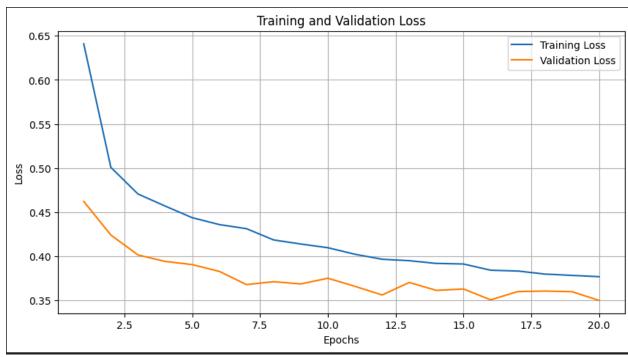


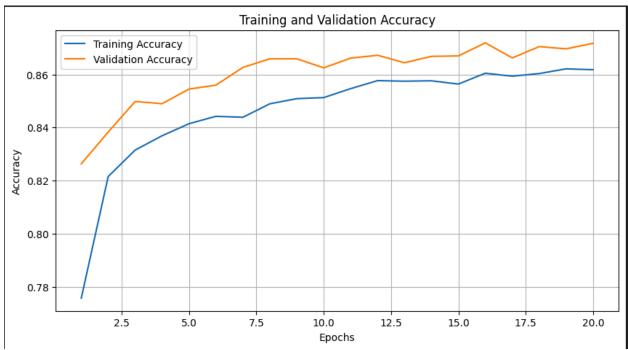


Learning rate 0.001:



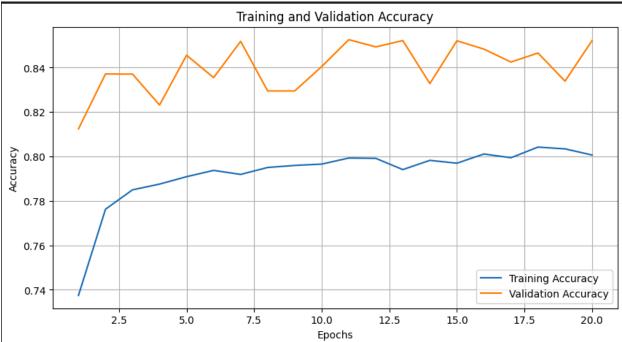
Learning rate 0.005:





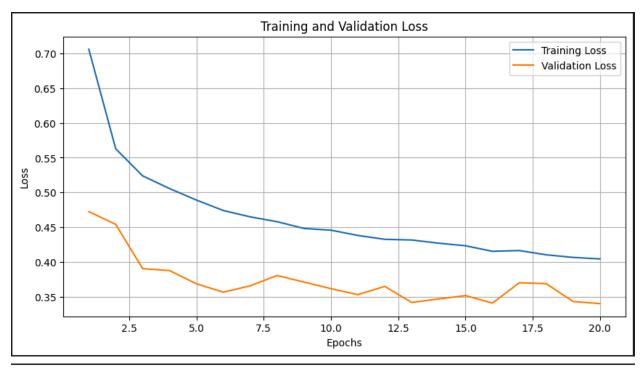
Learning rate 0.05:

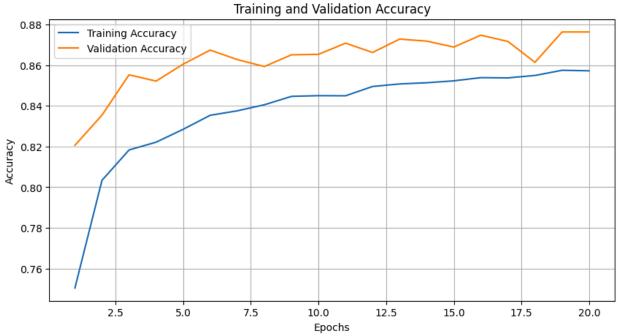




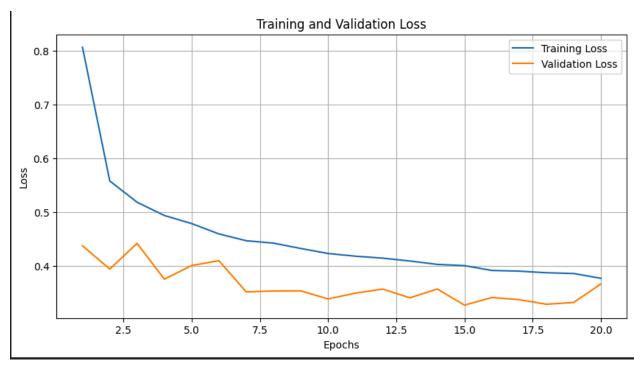
<u>Model3 (Two Layers, First Layer Hidden Dim = 128, Second Layer Hidden Dim = 64):</u>

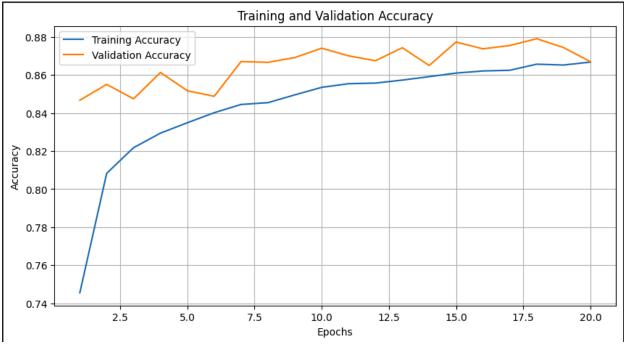
Learning rate 0.01:





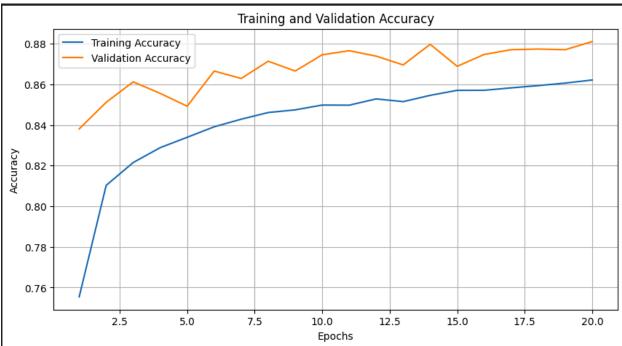
Learning rate 0.001:





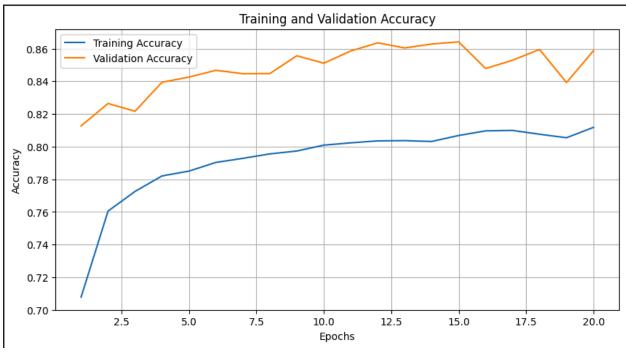
Learning rate 0.005:





Learning rate 0.05:





Validation Macro F1 Score: 0.8592589538216586

Best Model: Model1 with learning rate 0.005
Test Accuracy for best model: 0.8131

Test Macro F1 Score for best model: 0.8102594034053245