```
Running the code is intuitive as I have added markdown cells.
Model 1:
layers = [200,50,10]:
      learning rate = 0.001,
            accuracy, F1 for validation set: 88.32%, 0.8819
            accuracy, F1 for Test set: 88.10%, 0.8809
      learning rate = 0.002,
            accuracy, F1 for validation set: 87.3%, 0.87
            accuracy, F1 for Test set: 87.28%%, 0.87
      learning rate = 0.0035,
            accuracy, F1 for validation set: 87.00%, 0.8681
            accuracy, F1 for Test set: 86.09%, 0.8603
      learning rate = 0.005,
            accuracy, F1 for validation set: 84.58%, 0.8406
            accuracy, F1 for Test set: 84.92%, 0.8501
layers = [100,10]:
      learning rate = 0.001,
            accuracy, F1 for validation set: 88.33%, 0.8830
            accuracy, F1 for Test set: 87.49%, 0.8754
      learning rate = 0.002,
            accuracy, F1 for validation set: 86.78%, 0.8674
            accuracy, F1 for Test set: 86.44%, 0.8648
      learning rate = 0.0035,
            accuracy, F1 for validation set: 85.43%, 0.8528
            accuracy, F1 for Test set: 85.57%, 0.8550
      learning rate = 0.005,
            accuracy, F1 for validation set: 84.58%, 0.8445
            accuracy, F1 for Test set: 84.94%, 0.8489
layers = [256,128,32,10]:
      learning rate = 0.001,
            accuracy, F1 for validation set: 88.58%, 0.8842
            accuracy, F1 for Test set: 88.20%, 0.8815
      learning rate = 0.002,
            accuracy, F1 for validation set: 87.68%, 0.8762
            accuracy, F1 for Test set: 87.30%, 0.8732
      learning rate = 0.0035,
            accuracy, F1 for validation set: 86.18%, 0.8612
            accuracy, F1 for Test set: 86.36%, 0.8641
      learning rate = 0.005,
            accuracy, F1 for validation set: 84.98%, 0.8490
            accuracy, F1 for Test set: 84.73%, 0.8472
```

All results are reproducible. Running time ratio (approx.) 1:2:3. Best results highlighted