

Problem2

Report

Network configuration: 4 layers and 300 neurons

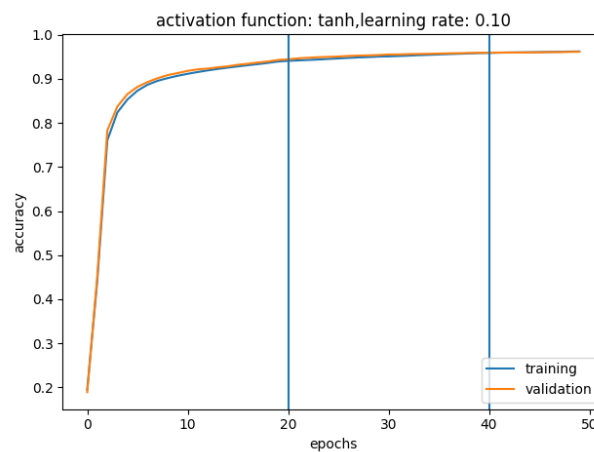
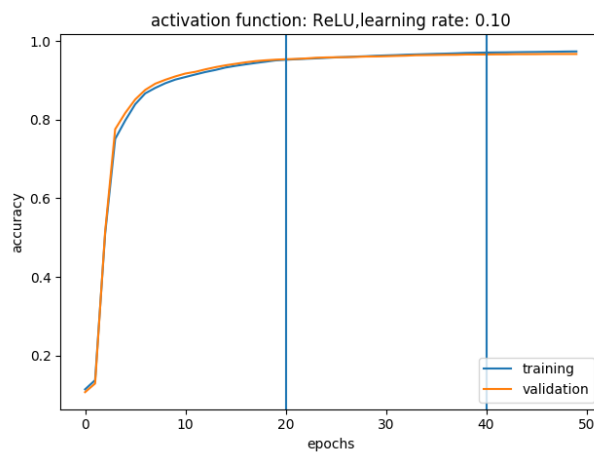
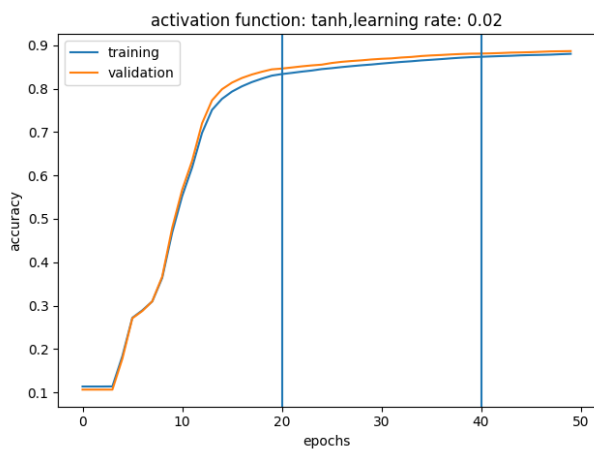
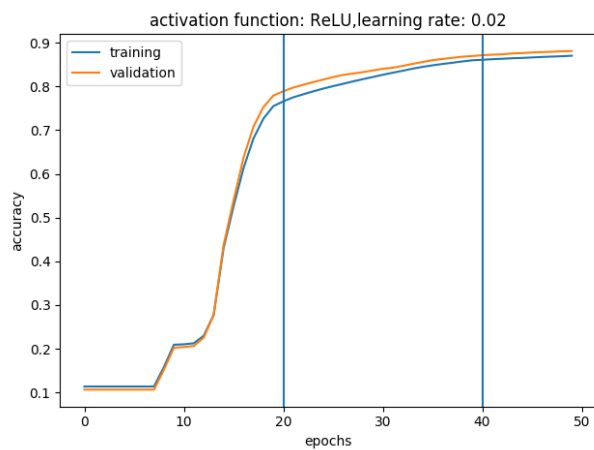
Batch size: 500

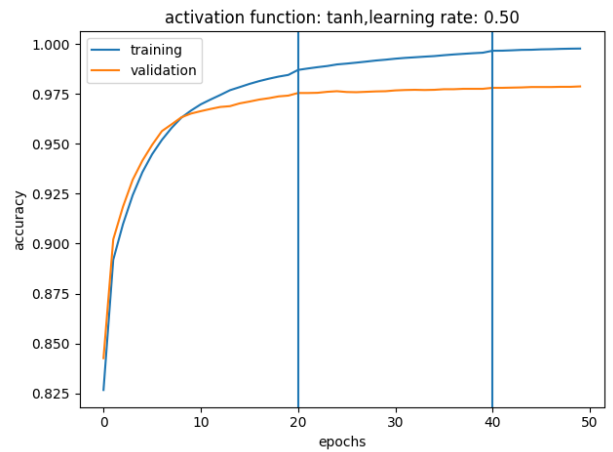
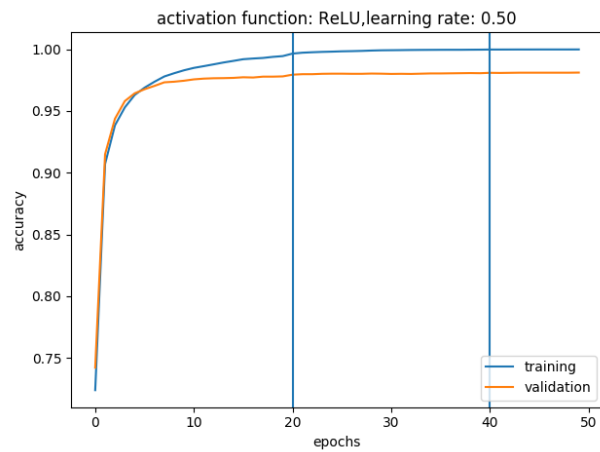
Initial learning rate: 0.02, 0.1 and 0.5

Initialize parameters: random number

```
W1 = 0.01 * np.random.randn(200, 784)
W2 = 0.01 * np.random.randn(100, 200)
W3 = 0.01 * np.random.randn(10, 100)
b1 = 0.01 * np.random.randn(200)
b2 = 0.01 * np.random.randn(100)
b3 = 0.01 * np.random.randn(10)
```

6 curves:





Final test accuracy:

activation fuction: relu, learning rate: 0.50, test_accuracy: 0.981