Report on CNN Offline Assignment

Submitted By:
Tanzim Hossain Romel
Student ID: 1705069

Architectures Used:

1. Architecture 1:

```
cnn = NeuralNetwork(
        input_dim=(28, 28, 1),
        layers=[
            Conv(5, 1, 8, activation=relu),
            Pool(2, 2),
            Conv(3, 1, 16, activation=relu),
            Pool(2, 2),
            Conv(3, 1, 32, activation=relu),
            Pool(2, 2),
            Flatten(),
            FullyConnected(64, relu),
            FullyConnected(10, softmax),
        ],
        optimizer=gradient_descent,
        cost_function=softmax_cross_entropy
    )
```

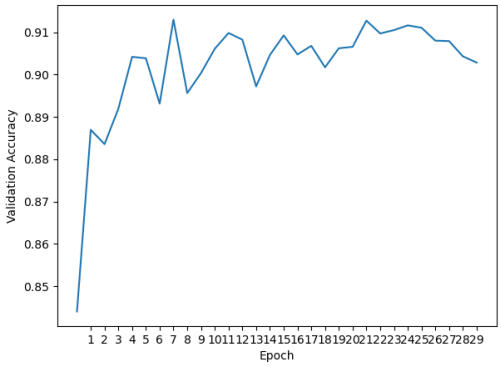
2. Architecture 2:

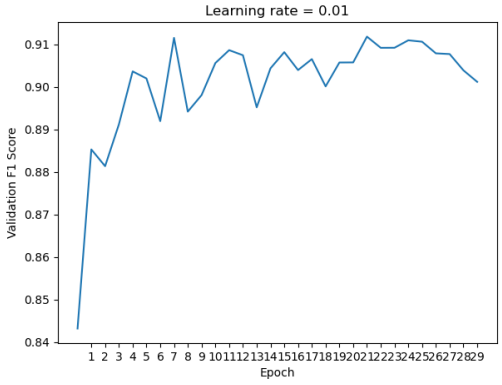
```
cnn = NeuralNetwork(
        input_dim=(28, 28, 1),
        layers=[
            Conv(5, 1, 6, activation=relu),
            Pool(2, 2),
            Conv(5, 1, 16, activation=relu),
            Pool(2, 2),
```

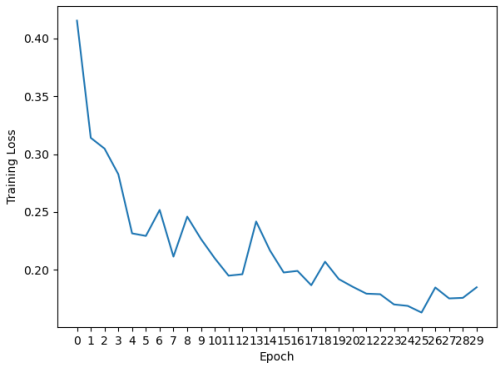
```
            Flatten(),
            FullyConnected(120, relu),
            FullyConnected(84, relu),
            FullyConnected(10, softmax),
        ],
        optimizer=gradient_descent,
        cost_function=softmax_cross_entropy
    )
```

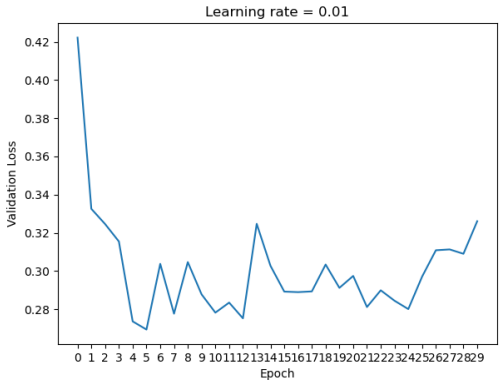
Plots:

1. Architecture 1, Learning Rate = 0.01, Batch Size = 64

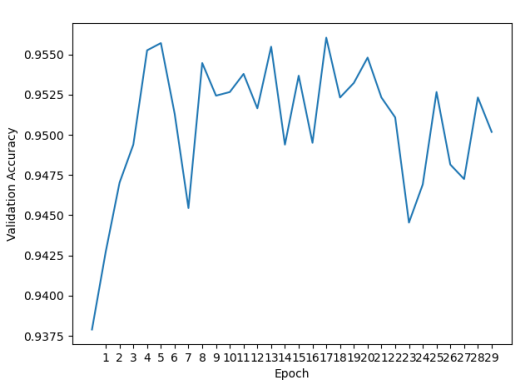


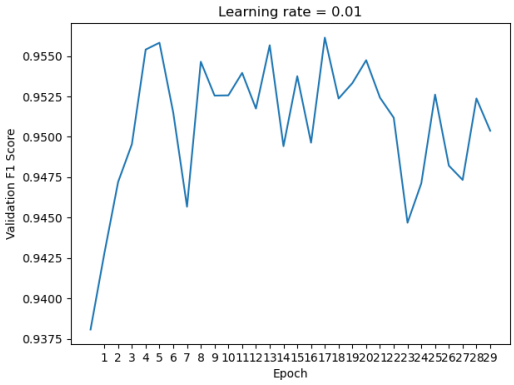


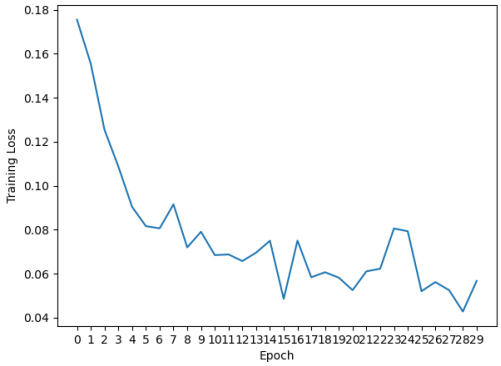


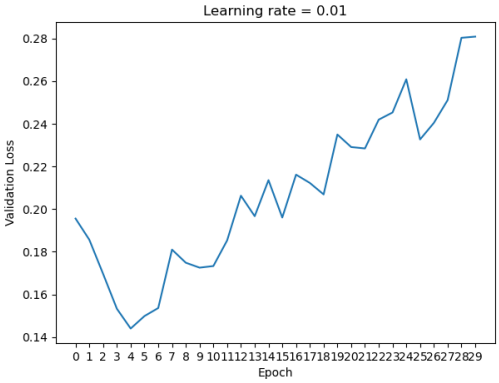


2. Architecture 2, Learning Rate = 0.01, Batch Size: 200

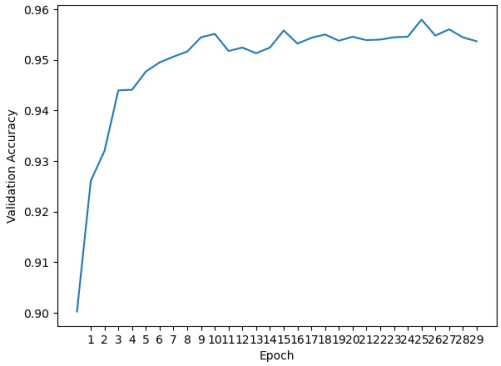


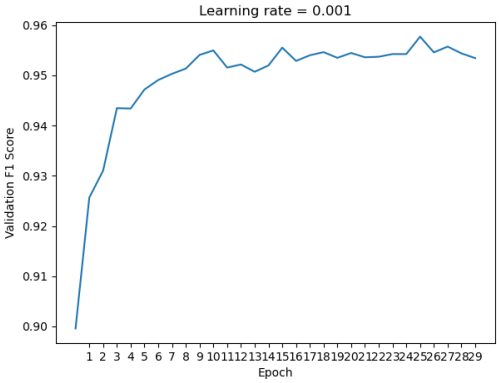


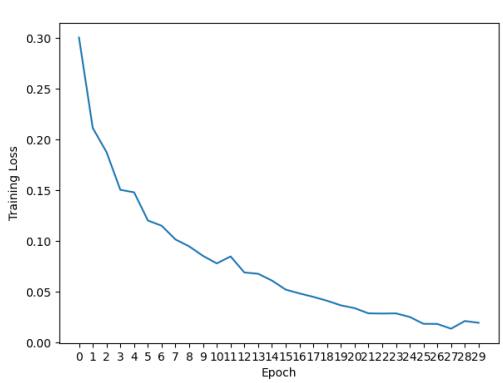


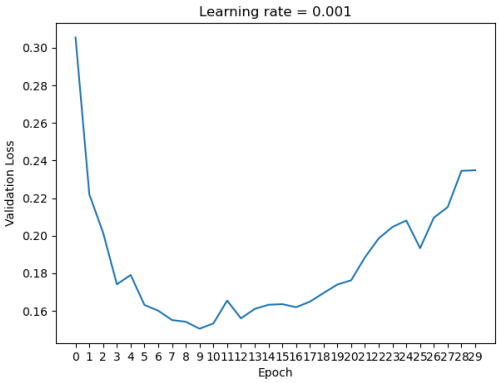


3. Architecture 2, Learning Rate = 0.001

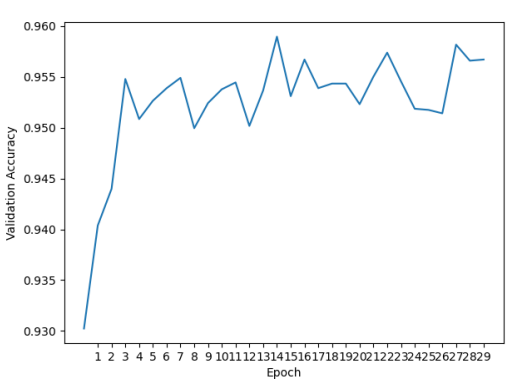


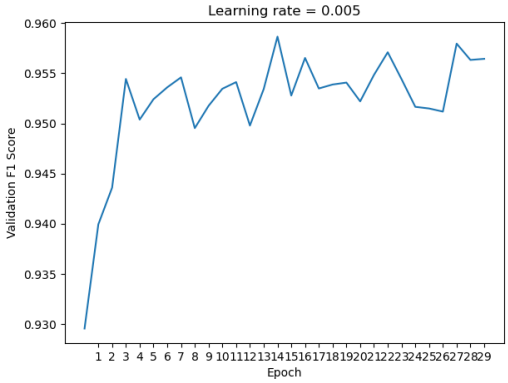


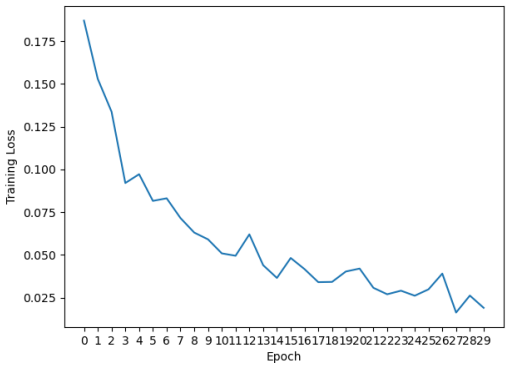


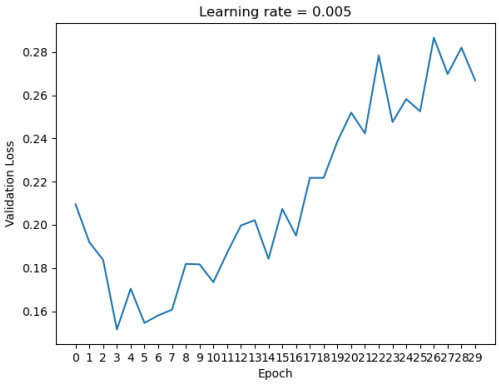


4. Architecture 2, Learning Rate = 0.005









Best Model:

Model 3 performed best among all others with:

Test accuracy = 0.9587458745874588 (95.87%)
Macro F1 Score = 0.9585968004380184