**Experiment 3 Description**

**1. Model Description:**

* This is a fully connected feedforward neural network.
* Input Layer has 784 neurons (28×28 pixel images flattened).

**Hidden Layers**:

* Layer 1: 128 neurons with sigmoid activation.
* Layer 2: 64 neurons with sigmoid activation.

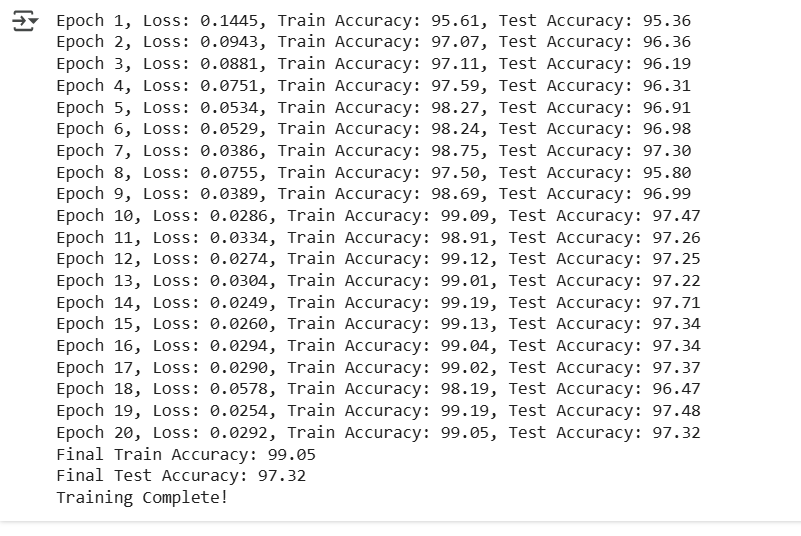
**Output Layer**:

* 10 neurons (0-9)
* Loss Function: Sparse Softmax Cross-Entropy.
* Optimizer: Adam with a learning rate of 0.01.
* Mini-batch Gradient Descent: Batch size of 100.
* Epochs: 20.

**2. Code Description:**

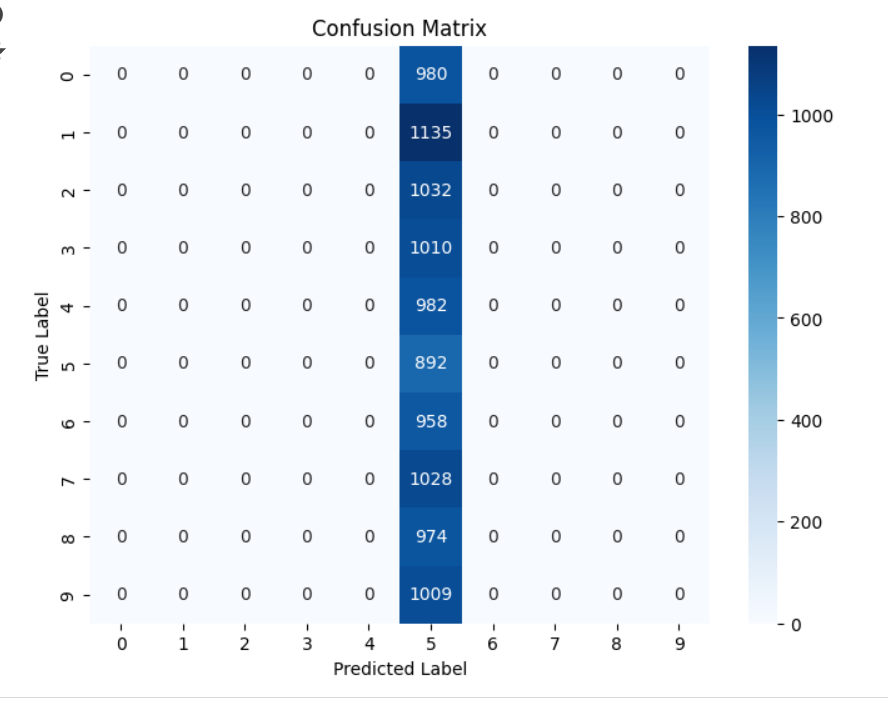
* Flattens images from 28×28 → 784 pixels.
* Normalizes pixel values to [0,1]. (dividing it by 255)
* Adam Optimizer for faster convergence.
* Using sparse softmax cross-entropy, which allows to directly use integer labels instead of applying one-hot encoding.
* Compares predictions with actual labels.
* Computes mean accuracy.
* Mini-batch training with batch size 100.
* Evaluates accuracy and loss after every epoch.

**3. Output:**

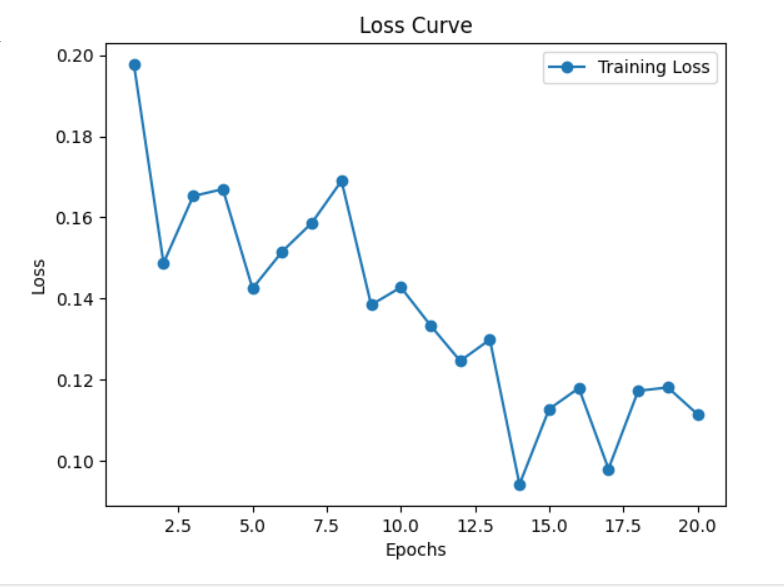


**4. Performance Evaluation:**

* Confusion matrix



* Loss Function:



**5. Improvements:**

* Use ReLU Activation:
* Use Convolutional Neural Network (CNN):