

Neural Network Assignment Report

Name: Jayant Parmar

Roll No: g25ait1072

1. Overview

This report presents the results of a hyperparameter sweep over hidden layer sizes (8, 16, 32, 64) and learning rates (0.1, 0.01, 0.001) for a single-hidden-layer MLP implemented from scratch. The dataset used is the Optical Recognition of Handwritten Digits dataset. **2. Hyperparameter Results**

hidden=8:

lr=0.100 → acc=0.9167

lr=0.010 → acc=0.3667

lr=0.001 → acc=0.1694

hidden=16:

lr=0.100 → acc=0.9583

lr=0.010 → acc=0.6333

lr=0.001 → acc=0.2389

hidden=32:

lr=0.100 → acc=0.9528

lr=0.010 → acc=0.8472

lr=0.001 → acc=0.2250

hidden=64:

lr=0.100 → acc=0.9556

lr=0.010 → acc=0.8806

lr=0.001 → acc=0.4000

Best Configuration: hidden_size=16, learning_rate=0.1

3. Best Model Training Summary (600 Epochs)

Epoch 1: loss=2.3009, val_acc=0.1778

Epoch 100: loss=0.1305, val_acc=0.9611

Epoch 200: loss=0.0692, val_acc=0.9722

Epoch 300: loss=0.0435, val_acc=0.9694

Epoch 400: loss=0.0301, val_acc=0.9694

Epoch 500: loss=0.0223, val_acc=0.9667

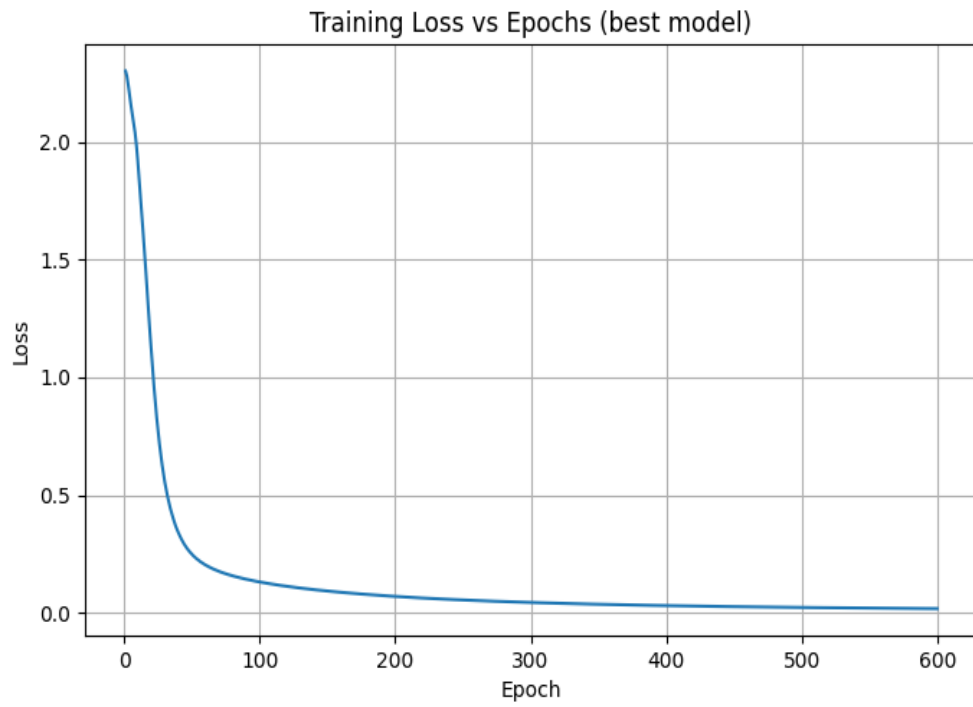
Epoch 600: loss=0.0173, val_acc=0.9667

Final Test Accuracy: 0.9667

4. Analysis

- Loss decreases smoothly and stabilizes after ~200 epochs.
- Validation accuracy peaks early and remains stable, indicating no overfitting.
- A hidden size of 16 gives the best balance of model capacity and generalization.

Training Loss Curve



Test Accuracy Curve

