Assignment 3:
Function Approximation with Neural Network and Backpropagation

1905095 - Md Raihan Sobhan

November 8, 2024

1 Introduction

In this assignment, we had to implement a Feed-Forward Neural Network (FNN) from scratch and apply our FNN to classify apparel.

2 Instructions for Running the Code

To reproduce the results in this report, follow these steps:

1. Set up a Python virtual environment and activate it:

   ```
   python3 -m venv myenv
   source myenv/bin/activate
   ```

2. Install the required packages:

   ```
   pip install --upgrade pip
   pip install numpy matplotlib scikit-learn pillow
   pip install torch torchvision torchaudio --index-url https://download.pytorch.org/w
   ```

3. Select Kernel - myenv (Python 3.12.5) on VS Code.

4. Run the main Jupyter Notebook File 1905095.ipynb by executing: Run All on VS Code.

   This will train the neural network on the FashionMNIST dataset, perform evaluation, and save the best model in pickle.

3 Training Results: Learning Rate and Model Comparisons

We experimented with 12 different configurations, combining four learning rates (`0.005`, `0.001`, `0.0005`, `0.0001`) and three hidden layer sizes (`64`, `128`, `256`). For each combination, we recorded training loss, validation loss, accuracy, and Macro-F1 score over epochs.

3.1 Learning Rate and Model Performance

Each combination of learning rate and hidden layer size is summarized below. We present graphs showing loss, accuracy, and Macro-F1 score over epochs for each combination, followed by the confusion matrix.

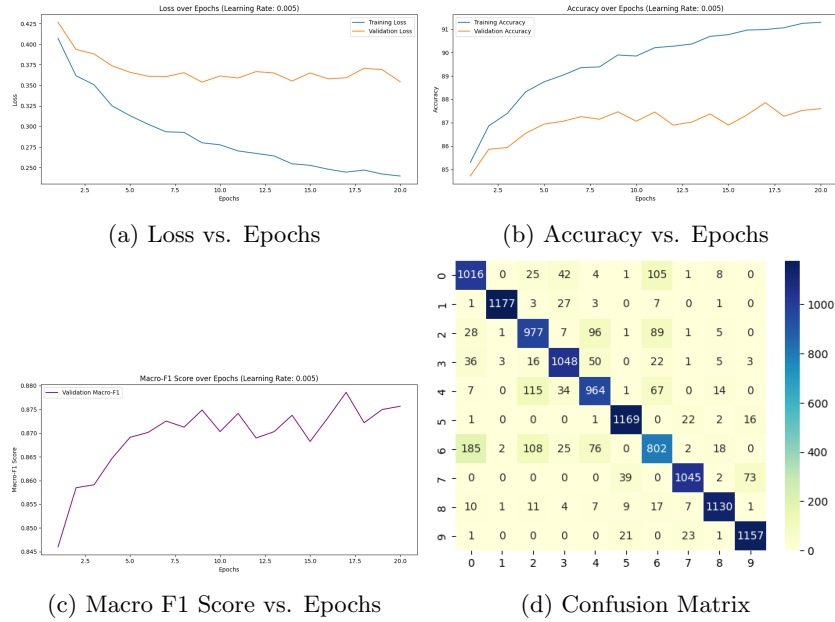### 3.1.1 Learning Rate: 0.005, Hidden Layer Size: 64

(a) Loss vs. Epochs

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 1: Metrics for Learning Rate 0.005, Hidden Layer Size 64
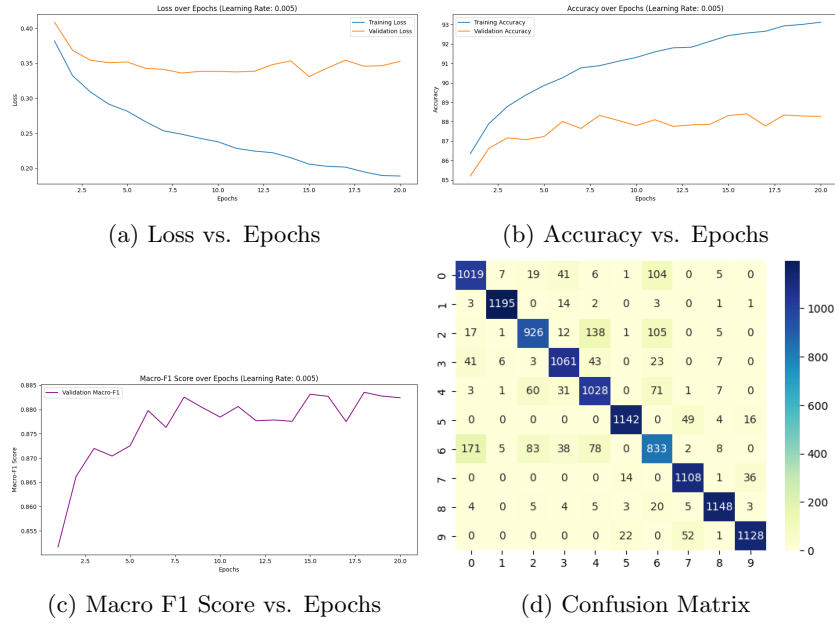
3.1.2 Learning Rate: 0.005, Hidden Layer Size: 128

(a) Loss vs. Epochs

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 2: Metrics for Learning Rate 0.005, Hidden Layer Size 128

4

3.1.3 Learning Rate: 0.005, Hidden Layer Size: 256

(a) Loss vs. Epochs
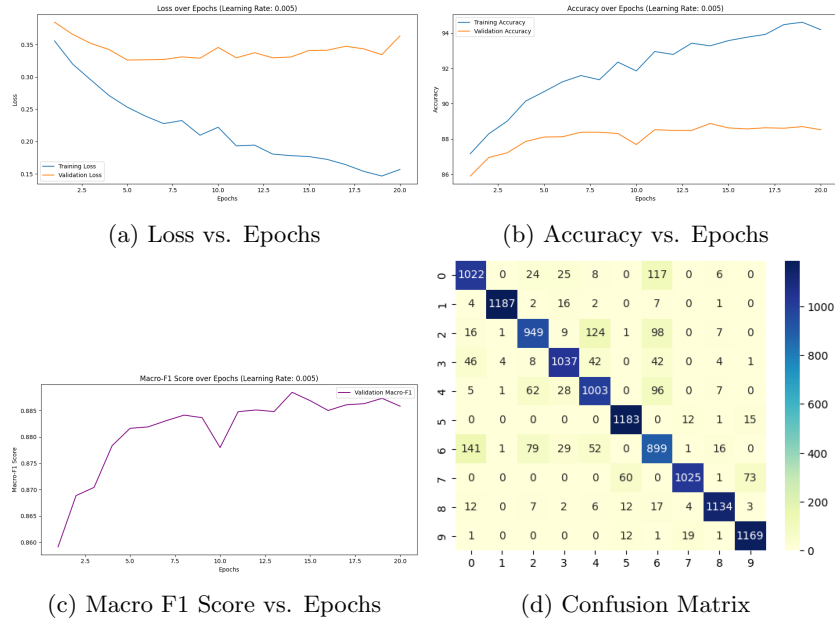

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 3: Metrics for Learning Rate 0.005, Hidden Layer Size 256
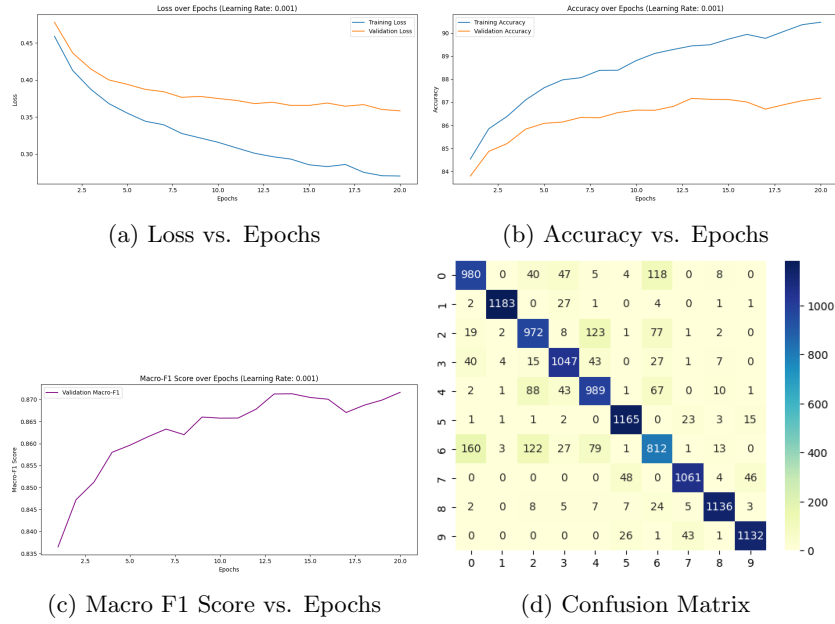
3.1.4 Learning Rate: 0.001, Hidden Layer Size: 64

(a) Loss vs. Epochs

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 4: Metrics for Learning Rate 0.001, Hidden Layer Size 64

3.1.5   Learning Rate: 0.001, Hidden Layer Size: 128

(a) Loss vs. Epochs
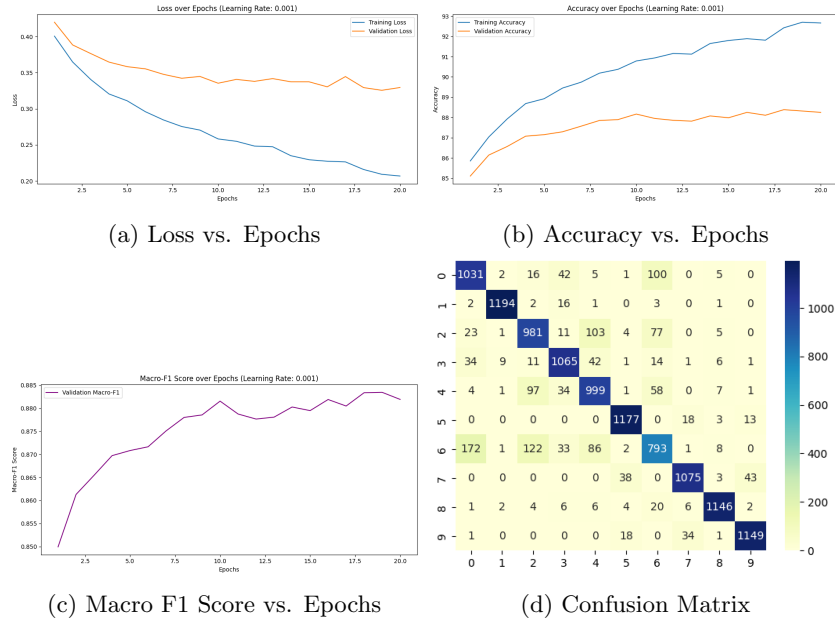

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 5: Metrics for Learning Rate 0.001, Hidden Layer Size 128

3.1.6 Learning Rate: 0.001, Hidden Layer Size: 256

(a) Loss vs. Epochs
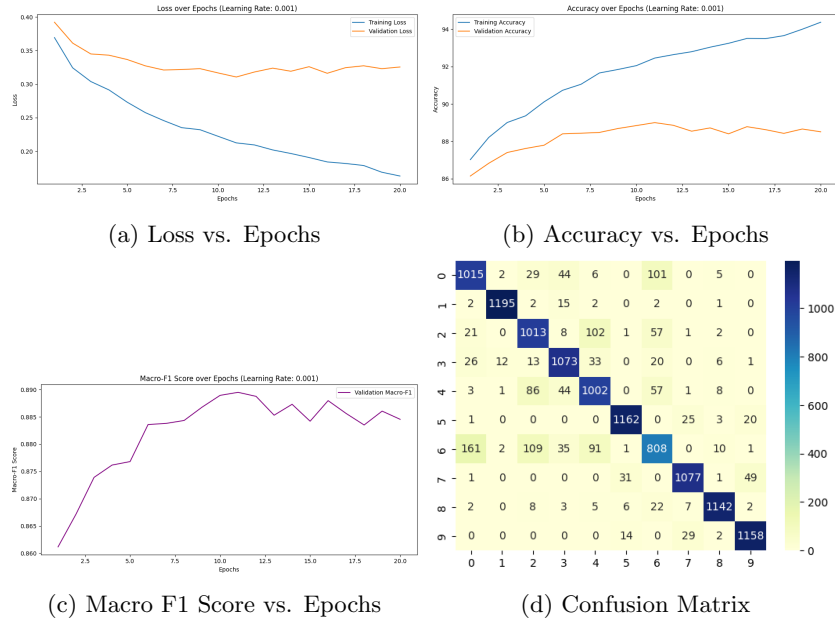
(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 6: Metrics for Learning Rate 0.001, Hidden Layer Size 256

3.1.7 Learning Rate: 0.0005, Hidden Layer Size: 64

(a) Loss vs. Epochs
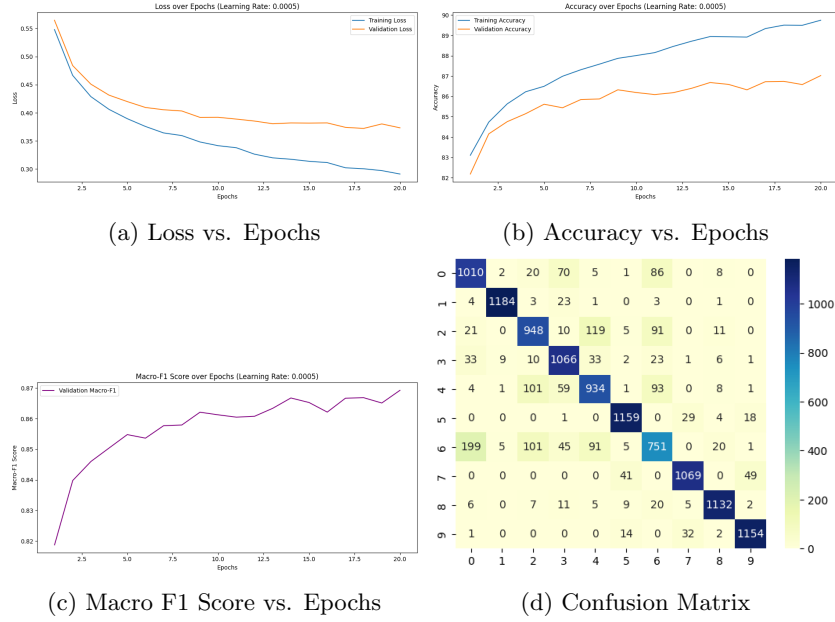
(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 7: Metrics for Learning Rate 0.0005, Hidden Layer Size 64

3.1.8 Learning Rate: 0.0005, Hidden Layer Size: 128

(a) Loss vs. Epochs
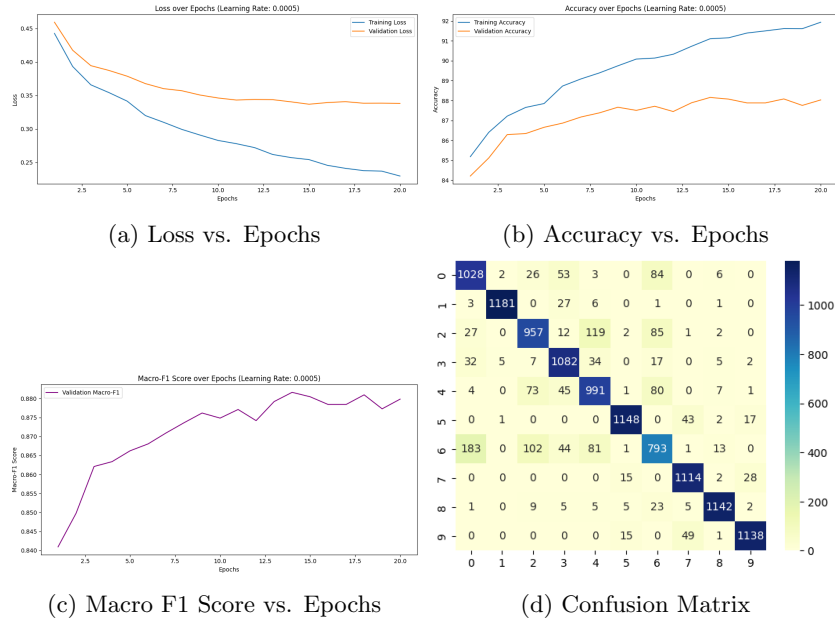
(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 8: Metrics for Learning Rate 0.0005, Hidden Layer Size 128

3.1.9   Learning Rate: 0.0005, Hidden Layer Size: 256

(a) Loss vs. Epochs
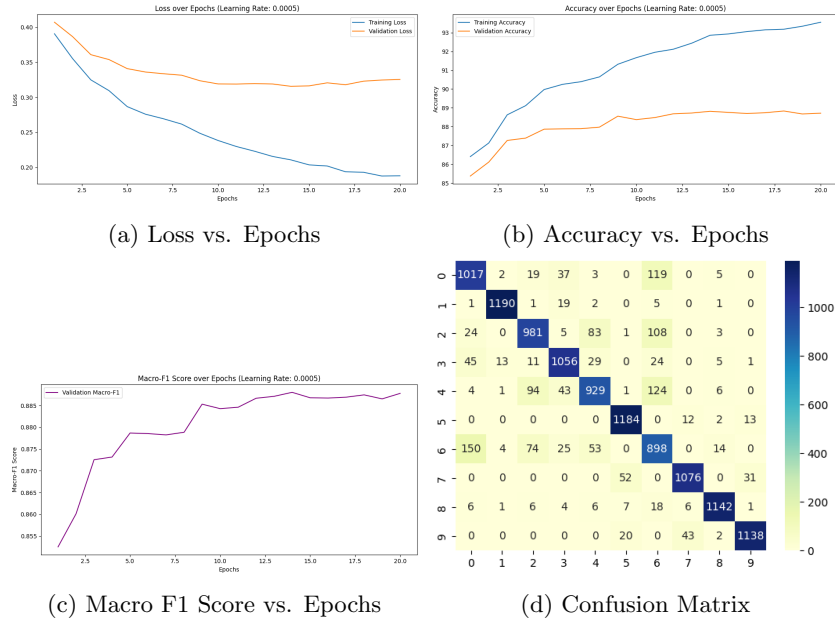

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 9: Metrics for Learning Rate 0.0005, Hidden Layer Size 256

3.1.10 Learning Rate: 0.0001, Hidden Layer Size: 64

(a) Loss vs. Epochs

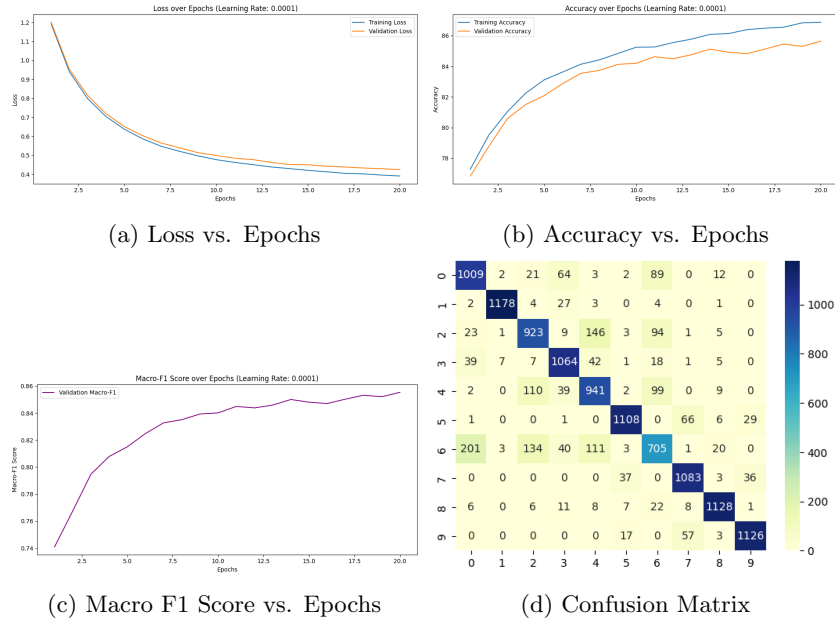(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 10: Metrics for Learning Rate 0.0001, Hidden Layer Size 64

3.1.11   Learning Rate: 0.0001, Hidden Layer Size: 128

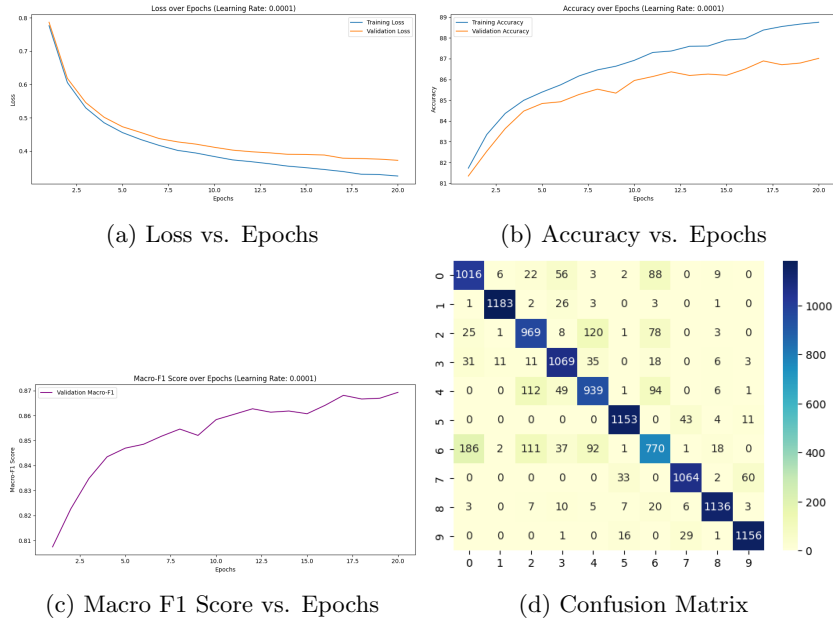(a) Loss vs. Epochs

(b) Accuracy vs. Epochs

(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 11: Metrics for Learning Rate 0.0001, Hidden Layer Size 128

3.1.12 Learning Rate: 0.0001, Hidden Layer Size: 256

(a) Loss vs. Epochs

(b) Accuracy vs. Epochs

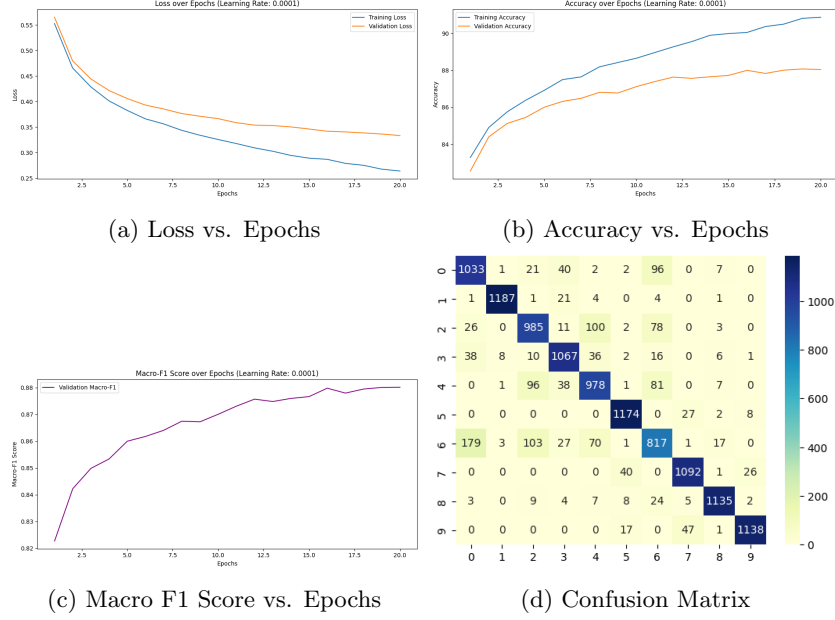(c) Macro F1 Score vs. Epochs

(d) Confusion Matrix

Figure 12: Metrics for Learning Rate 0.0001, Hidden Layer Size 256

4 Comparison and Selection of the Best Model

From the experiments, we identified the best model based on the highest validation Macro-F1 score. Table 1 summarizes the final metrics for each configuration.

Table 1: Validation Metrics for Each Model Configuration

LR	Hidden Layer	Train Loss	Val Loss	Train Acc	Val Acc	Val Macro-F1
0.005	64	0.2343	0.3603	91.35%	87.17%	0.8714
0.005	128	0.1874	0.3494	93.11%	87.94%	0.8791
0.005	256	0.1558	0.3437	94.22%	88.52%	0.8857
0.001	64	0.2693	0.3616	90.41%	87.04%	0.8701
0.001	128	0.2085	0.3382	92.61%	88.36%	0.8837
0.001	256	0.1606	0.3254	94.39%	88.93%	**0.8890**
0.0005	64	0.2938	0.3706	89.76%	87.07%	0.8705
0.0005	128	0.2292	0.3380	92.09%	88.15%	0.8812
0.0005	256	0.1864	0.3245	93.35%	88.38%	0.8835
0.0001	64	0.3901	0.4288	86.90%	85.32%	0.8526
0.0001	128	0.3194	0.3709	88.99%	87.20%	0.8714
0.0001	256	0.2649	0.3361	90.92%	88.08%	0.8798

5 Test Performance of the Best Model

The model with the highest validation Macro-F1 score was chosen as the best model (LR = 0.001 and Hidden Layer Size 256). We evaluated it on an independent test set and recorded the following metrics:

- **Test Accuracy**: 89.07

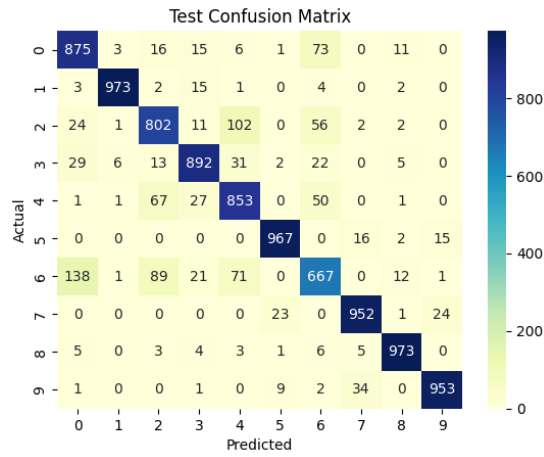- **Test Macro-F1 Score**: 0.8900029201513565

Figure 13: Confusion Matrix for Best Model on Test Set

6 Conclusion

This report presents the results of training a neural network on the FashionM-NIST dataset, evaluating performance with 12 configurations of learning rate and hidden layer size. The optimal model achieved a validation Macro-F1 score of 0.89 and also achieved a strong test accuracy of 89.07