Handwritten Digit Recognition Using Machine Learning

Krishna Balaji
Arizona State University

April 22, 2025

Abstract

This project investigates the task of handwritten digit recognition using the MNIST dataset. Various machine learning models including Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and a simple Neural Network were implemented and evaluated. Performance was measured using accuracy, precision, recall, F1-score, and confusion matrices. The results indicate that the Neural Network achieved the highest accuracy, with ensemble methods such as Random Forest and KNN producing competitive performance.

Contents

1 Introduction

Handwritten digit recognition is a classic problem in machine learning that serves as an ideal benchmark for model comparison. The MNIST dataset, containing 60,000 training images and 10,000 testing images of 28×28 pixel grayscale digits, has been widely used for this purpose.

The main objectives of this project are:

- To implement and compare several machine learning models including both traditional models and a neural network.
- To evaluate model performance using metrics such as accuracy, precision, recall, F1-score, and confusion matrices.
- To identify the most effective approach for this classification problem.

2 Methodology

2.1 Data Preprocessing

The MNIST dataset is composed of grayscale images that represent digits from 0 to 9. The following preprocessing steps were performed:

1. **Normalization:** Each pixel's intensity was scaled from the range [0, 255] to [0, 1].
2. **Flattening:** For traditional machine learning models, the 28×28 images were reshaped into 784-dimensional vectors.

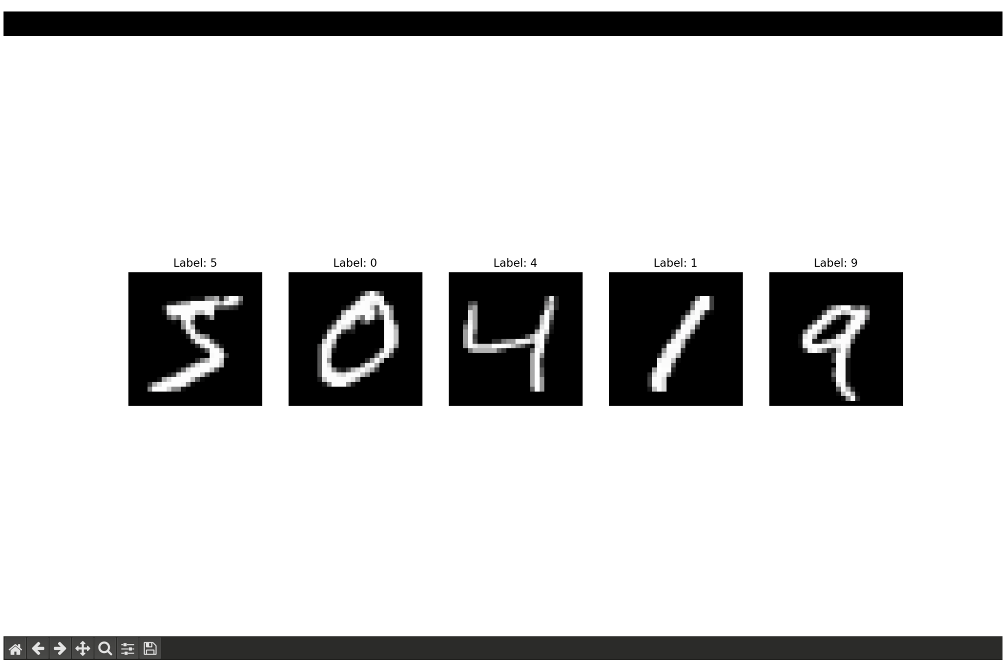

Figure 1: Sample images from the MNIST dataset.

2.2 Model Descriptions

The following models were implemented:

2.2.1 Logistic Regression

A linear classifier applied to the flattened input data.

2.2.2 Decision Tree

A tree-based classifier which splits the data based on feature values; may overfit if not constrained.

2.2.3 Random Forest

An ensemble method that aggregates predictions from multiple decision trees to improve overall performance.

2.2.4 K-Nearest Neighbors (KNN)

A non-parametric classifier that assigns a label based on the majority vote among its k nearest neighbors.

2.2.5 Neural Network

A simple feed-forward neural network with:

- A `Flatten` layer converting each 28×28 image into a vector.
- One hidden Dense layer with 128 neurons and ReLU activation.
- An output Dense layer with 10 neurons and softmax activation.

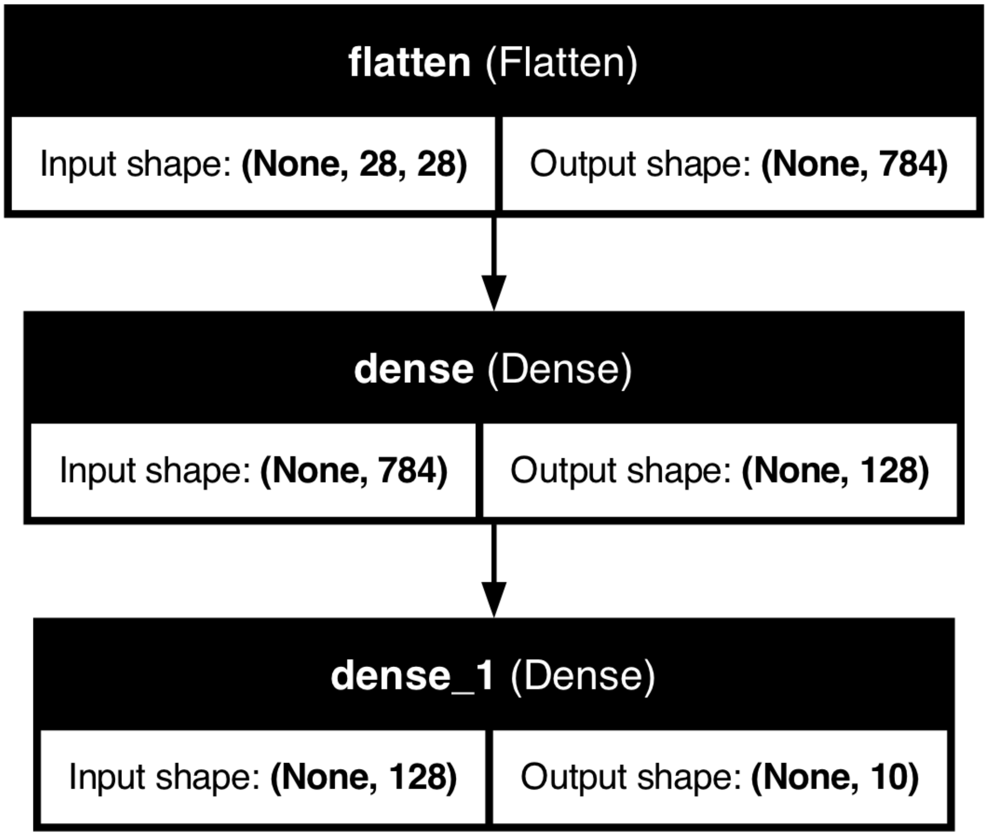

Figure 2: Neural Network Architecture.

2.3 Training and Evaluation

All models were trained using the MNIST training set and evaluated on the testing set. The evaluation metrics include:

- **Accuracy**
- **Precision, Recall, and F1-Score** (provided in classification reports)
- **Confusion Matrix** (visualized as heatmaps)

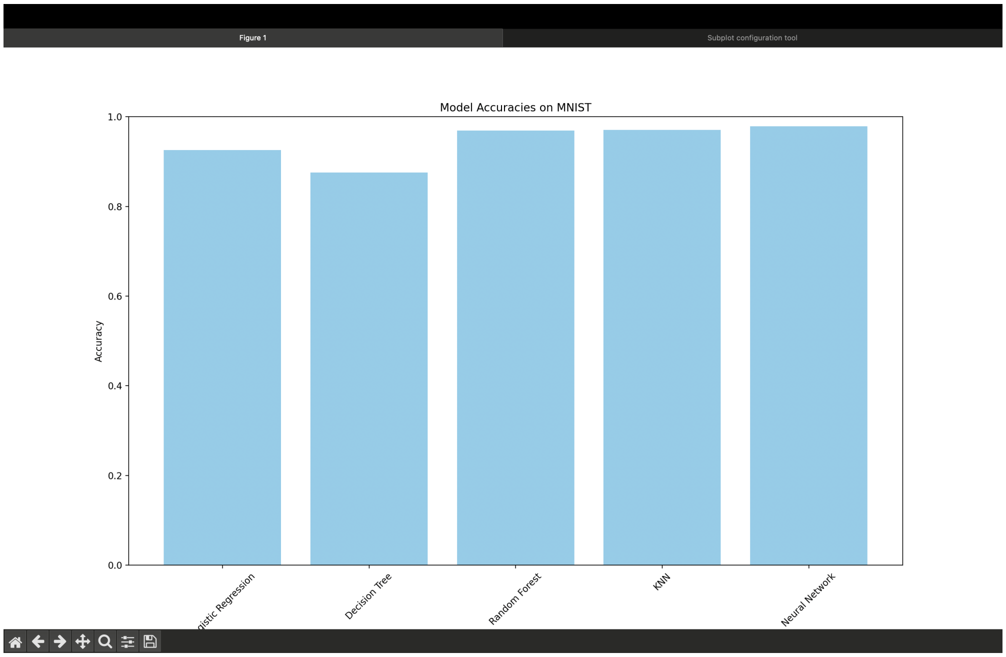

Figure 3: Comparison of model accuracies on the MNIST dataset.

3 Results

Table 1 summarizes the performance of the implemented models.

Model	**Accuracy**
Logistic Regression	92.6%
Decision Tree	87.6%
Random Forest	96.9%
K-Nearest Neighbors	97.1%
Neural Network	97.8%

Table 1: Model accuracies on the MNIST dataset.

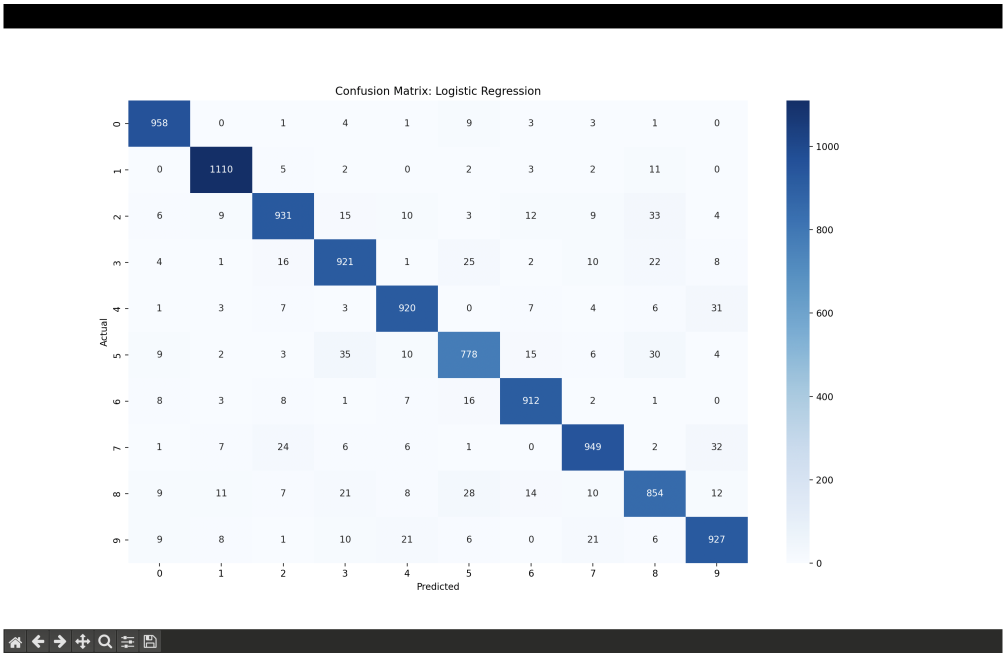

Figure 4: Confusion Matrix for Logistic Regression.

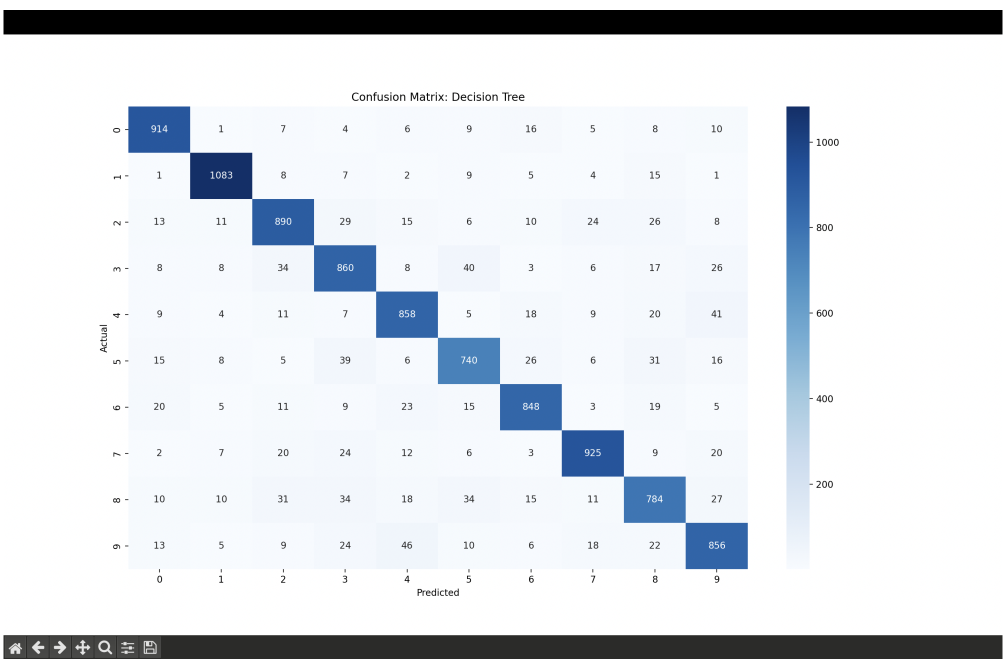

Figure 5: Confusion Matrix for Decision Tree.

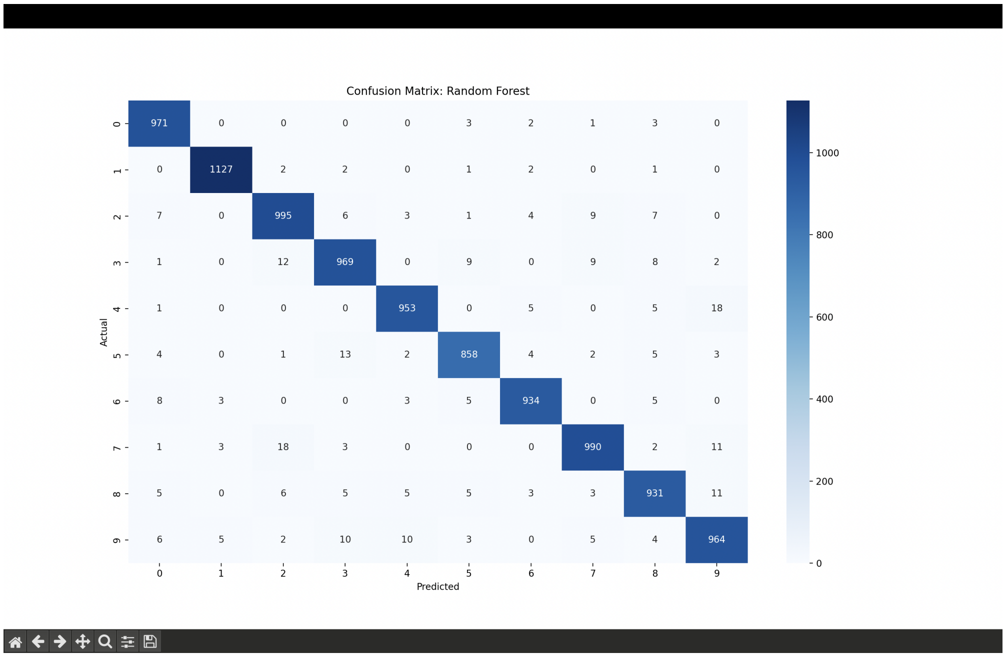

Figure 6: Confusion Matrix for Random Forest.

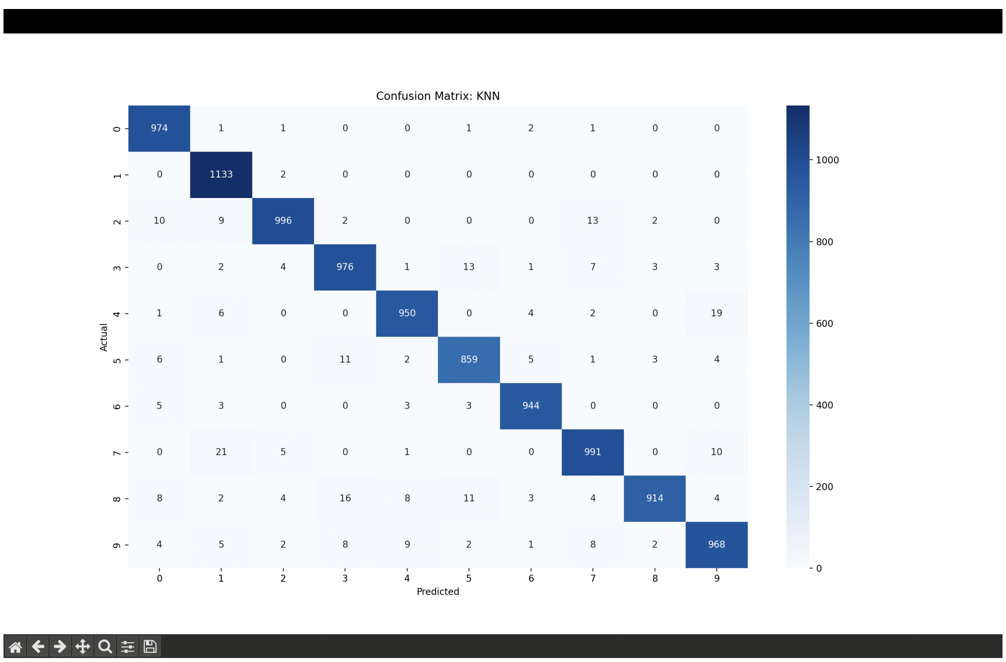

Figure 7: Confusion Matrix for K-Nearest Neighbors.

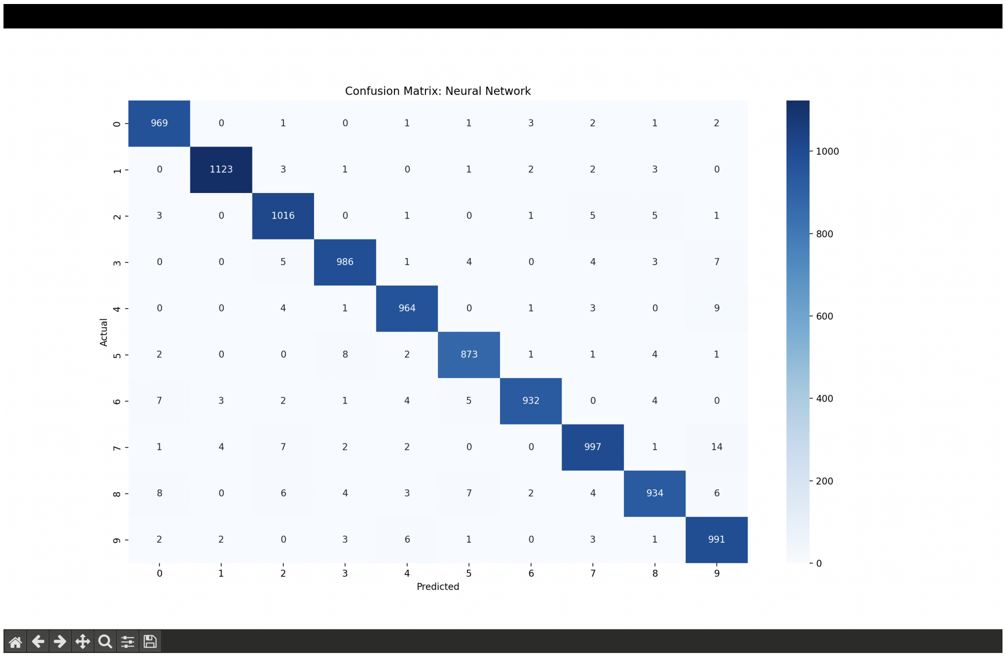

Figure 8: Confusion Matrix for the Neural Network.

4 Discussion

The experimental results indicate that:

- The **Neural Network** achieved the highest accuracy (97.8%), indicating its superior ability to model non-linear relationships in the image data.
- Both the **Random Forest** and **K-Nearest Neighbors** achieved accuracies of approximately 97%, suggesting that ensemble methods and non-parametric approaches can effectively handle this classification task.
- The **Logistic Regression** model, while effective, was outperformed by the non-linear models.
- The **Decision Tree** exhibited the lowest performance (87.6%), likely due to overfitting.

5 Conclusion

This study implemented and compared several machine learning approaches for the task of handwritten digit recognition. The Neural Network outperformed traditional models, demonstrating its capability to capture complex patterns in image data. While Random Forest and KNN also offered robust performance, the Neural Network's higher accuracy suggests that deep learning methods are particularly suited for this type of image classification task. Future work could explore deeper architectures such as Convolutional Neural Networks (CNNs) to further improve performance.

6 References

1. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.

2. MNIST Dataset. `http://yann.lecun.com/exdb/mnist/`

3. Scikit-learn documentation: `https://scikit-learn.org/`

4. TensorFlow/Keras documentation: `https://www.tensorflow.org/guide/keras/`