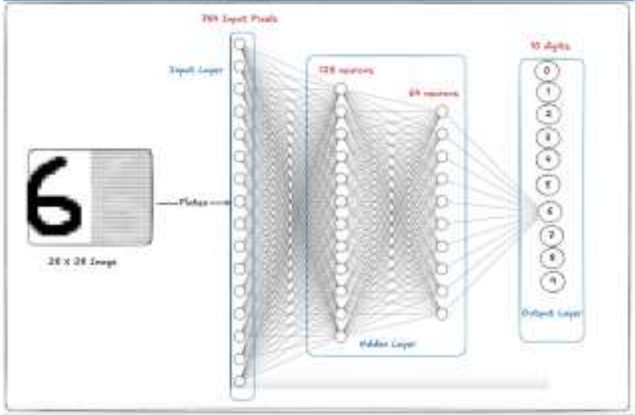
Handwritten Digit Recognition

A Comparative Analysis of Classification Algorithms

Graduate-Level Machine Learning Project



Project Overview & Objectives

Objective

Implement and compare multiple machine learning approaches for recognizing handwritten digits, focusing on accuracy, efficiency, and model robustness.

Relevance

Handwritten digit recognition is a fundamental problem in computer vision with critical applications in postal automation, bank check processing, and automated data entry systems.

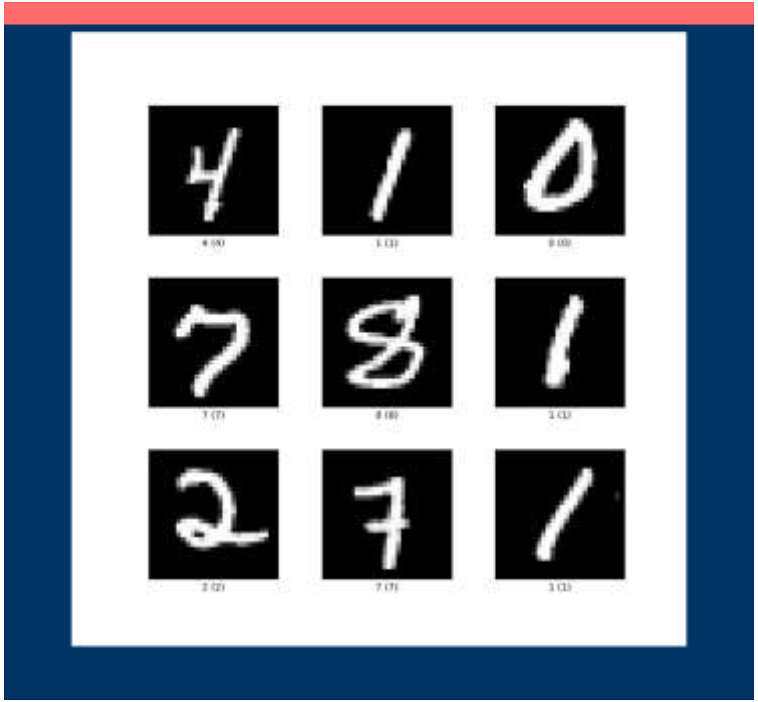
Key Models Explored

- **Multi-Layer Perceptron**
- **Logistic Regression**
- **Linear Regression (OvA)**
- **Naive Bayes (GaussianNB)**

Dataset: DIDA 10K Version

Dataset Size	10,000 samples of handwritten digits
Input Format	28x28 pixel grayscale images
Representation	Flattened 784-dimensional vectors
Output Classes	10 classes (digits 0 through 9)

Source: didadataset.github.io/DIDA/



Methodology: Data Preprocessing

Normalization

Pixel values were scaled to the **[0, 1] range**. This ensures numerical stability and facilitates faster convergence during the model training phase.

Data Splitting

The dataset was divided into an **80% Training Set** (8,000 samples) for optimization and a **20% Testing Set** (2,000 samples) for final evaluation.

Validation

Implemented **5-fold Cross-Validation** to ensure model robustness. This technique provides a more reliable estimate of model performance on unseen data.

Methodology: Multi-Layer Perceptron (MLP)

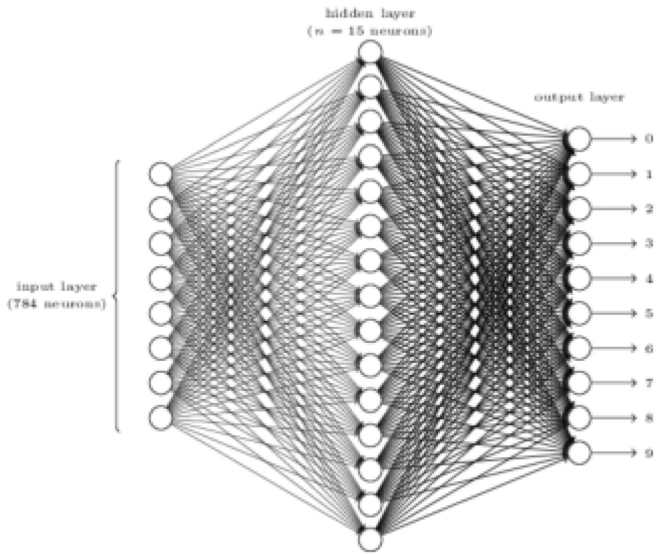
Architecture

A feed-forward neural network consisting of an input layer (784 neurons), a minimum of two hidden layers with optimized neuron counts, and a 10-neuron output layer.

Optimization

Hyperparameters including learning rate, batch size, and number of epochs were tuned using iterative optimization to minimize classification error on the training set.

Key Strength: Capable of capturing complex non-linear relationships in high-dimensional pixel data through hierarchical feature learning.



Methodology: Regression Models

Logistic Regression

Multi-class Setup

Implemented using a multinomial probabilistic approach to handle the 10-digit classification task directly.

Optimization

Focused on convergence criteria and L2 regularization to ensure model stability and prevent overfitting on high-dimensional pixel data.

Linear Regression (OvA)

One-vs-All Strategy

Adapted for classification by training 10 separate binary linear models, each identifying one specific digit against all others.

Decision Logic

Predictions are made by selecting the class whose corresponding model yields the highest confidence score for the input vector.

Methodology: Naive Bayes Classifier

Probabilistic Approach

Based on **Bayes' Theorem**, this model calculates the probability of a digit class given the observed pixel intensities. It assumes that each pixel's value is independent of others, simplifying the complex joint probability distribution.

Implementation

The **Gaussian Naive Bayes (GaussianNB)** variant was selected for this project. It assumes that the continuous pixel intensity values for each class follow a normal (Gaussian) distribution.

Key Assumption

Feature Independence: Despite being "naive" in image data where pixels are highly correlated, it provides a fast and efficient baseline.

Role as Baseline

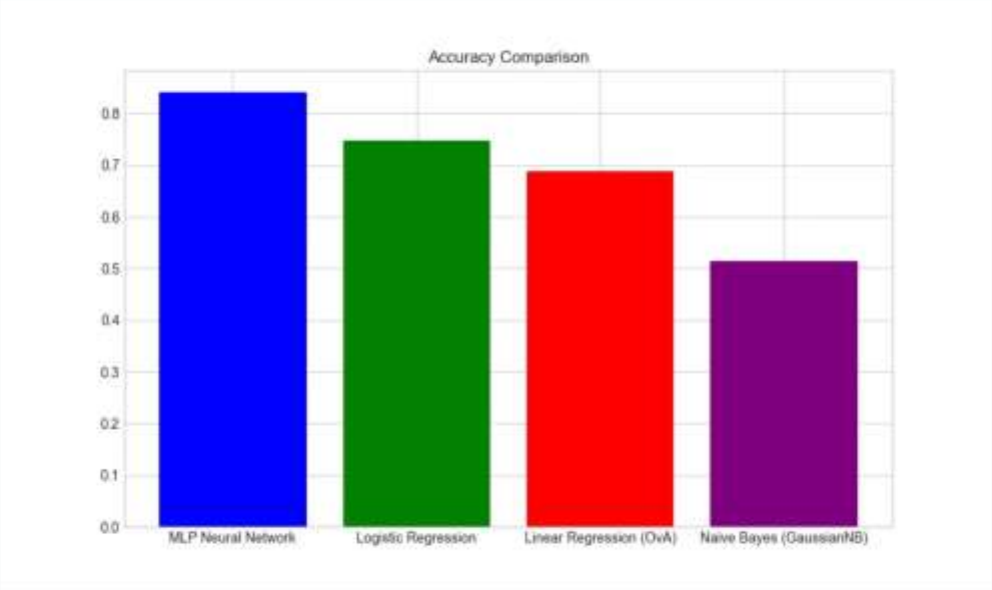
Serves as a **performance baseline** to quantify the gains achieved by more complex discriminative models like Logistic Regression and Multi-Layer Perceptrons. It highlights the necessity of capturing spatial dependencies in digit recognition.

Results: Performance Metrics Summary

Model	Accuracy	Precision	Recall	F1-Score
MLP Neural Network	**0.841**	**0.849**	**0.841**	**0.841**
Logistic Regression	0.747	0.747	0.747	0.746
Linear Regression (OvA)	0.689	0.689	0.689	0.687
Naive Bayes	0.514	0.543	0.514	0.516

** Metrics evaluated on the 20% test set (2,000 samples). Highlighting indicates the top-performing model.*

Results: Accuracy Comparison Visualized



MLP Superiority

The Multi-Layer Perceptron significantly outperforms all other models, achieving ~84% accuracy.

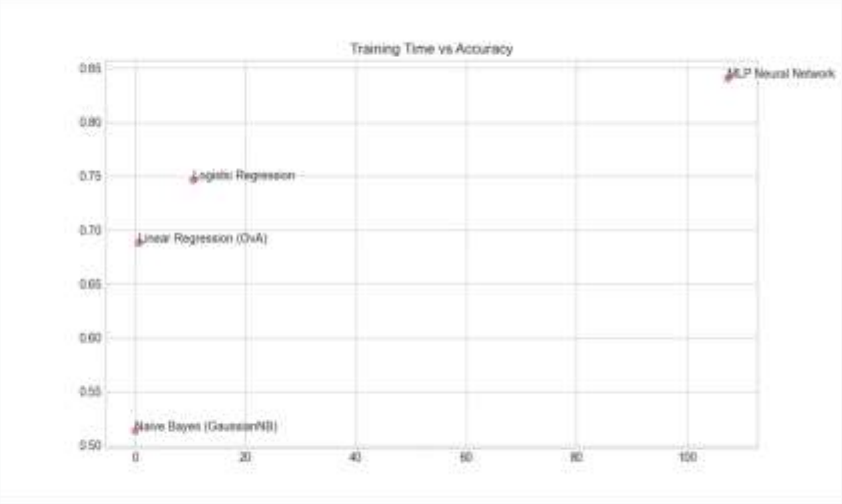
Linear Baselines

Logistic and Linear Regression show moderate performance, capturing basic linear patterns.

NB Limitations

Naive Bayes shows the lowest accuracy, highlighting the impact of the pixel independence assumption.

Results: Training Time vs. Accuracy



Computational Efficiency

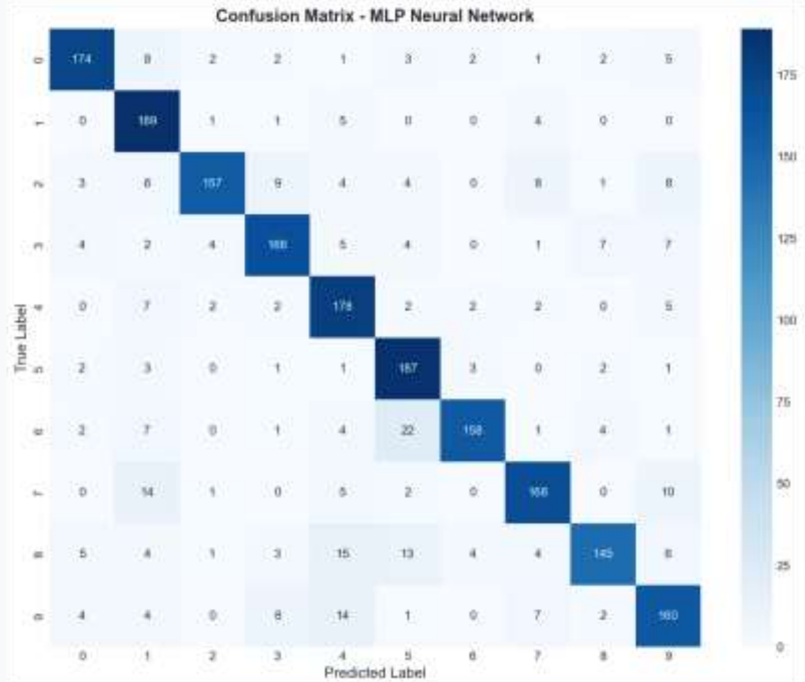
Models like **Naive Bayes** and **Linear Regression** offer near-instant training (<1s) but suffer from lower predictive power.

The MLP Trade-off

The **MLP Neural Network** requires significantly more time (~107s) due to its iterative backpropagation process.

Conclusion: For this dataset, the 10% accuracy gain of MLP over Logistic Regression justifies the increased computational cost.

Results: Error Analysis via Confusion Matrix



Diagonal Dominance

High values along the main diagonal confirm the **MLP's high accuracy** across all 10 digit classes.

Common Errors

Minor misclassifications occur between visually similar digits, such as **4 and 9** or **3 and 8**, where structural features overlap.

Model Robustness

The model maintains consistent performance without significant bias toward any specific digit, demonstrating effective feature extraction.

Key Takeaways

Model Superiority

Multi-Layer Perceptrons (MLP) are significantly better suited for image-based recognition tasks, achieving 84.1% accuracy compared to traditional linear baselines.

Methodology Impact

Rigorous data normalization and 5-fold cross-validation were critical in ensuring model stability and providing a reliable estimate of real-world performance.

Performance vs. Speed

While MLP requires higher computational resources, the substantial gain in predictive power justifies the increased training time for precision-critical applications.

The project successfully validated the efficacy of neural networks on the DIDA dataset, establishing a robust framework for handwritten digit classification.