ML LAB END-SEMESTER

Prabhav Gupta DA24B018

November 15, 2025

(a) Summary of Models Used & System Architecture

In this project, I tried many different machine learning models on the MNIST digit dataset. Steps taken to find the best model:

- First I loaded the MNIST CSV files and defined metrics.

- Sometimes applied PCA to reduce the number of features.

- Trained many different models like Softmax, Decision Trees, Random Forests, Boosting, and KNN.

- Then compared their performance using Accuracy and F1 score.

- Finally, combined the best models into an ensemble.

All Models Used

- **Softmax Regression:** This model tries to guess the class by checking which one has the highest score.

- **Improved Softmax:** Same as Softmax Regression but trained in smaller batches with regularization to avoid overfitting. It makes the accuracy better.

- **PCA + Softmax:** PCA reduces the size of the data by keeping only the most important patterns. Then Softmax is trained on this smaller data, which makes it faster.

- **Decision Tree:** This model keeps splitting until it decides which number the digit looks like.

- **Random Forest:** This is just many decision trees working together as a forest. Each tree predicts a value and the one with highest number of predicted values is chosen as the final prediction.

- **Gradient Boosting:** This model builds trees one after another. Each new tree tries to fix the mistakes made by the previous trees.

- **PCA + KNN:** This model checks which training digits look most similar to the new digit by the help of its closest neighbours.

- **Ensembles:** Here, I combine multiple models so that if one model makes a mistake, the others fixes it and this combined model gave the highest accuracy.

(b) Hyperparameter Tuning and Results (Summarized Tables)

In this part, I tested many models with different hyperparameters. Below are summarized tables including only the main results. We compare performance using only Accuracy and F1 score.

Softmax Regression (Basic)

Learning Rate	Epochs	Accuracy	F1 Score
0.05	100	0.8350	0.8318
0.05	200	0.8590	0.8569
0.05	400	0.8725	0.8706
0.1	100	0.8595	0.8574
0.1	200	0.8725	0.8706
0.1	400	0.8845	0.8832
0.2	100	0.8725	0.8706
0.2	200	0.8845	0.8832
0.2	400	**0.8925**	**0.8911**

Improved Softmax

LR	Epochs	Batch	L2	Accuracy	F1 Score
0.2	1500	128	0.0001	0.8855	0.8844
0.2	1500	128	0.0010	**0.9025**	**0.9013**
0.2	1500	256	0.0010	0.8975	0.8963
0.3	1500	128	0.0010	0.8995	0.8978
0.3	1500	256	0.0010	0.9040	0.9030

PCA + Softmax

PCA	LR	L2	Batch	Accuracy
40	0.1	0.001	128	0.9132
40	0.2	0.001	256	0.9120
40	0.3	0.001	128	0.9108
60	0.1	0.001	128	0.9208
60	0.2	0.001	256	**0.9220**
60	0.3	0.001	256	0.9184

Decision Tree

Max Depth	Min Split	Accuracy	F1 Score
5	2	0.7500	0.7490
10	2	**0.8030**	**0.8002**
15	2	0.8030	0.8002

Random Forest

Trees	Depth	Accuracy	F1 Score
3	8	0.8340	0.8313
5	8	0.8680	0.8662
8	10	**0.8880**	**0.8867**

Gradient Boosting

PCA	Depth	Estimators	Accuracy	F1 Score
30	4	4	0.6475	0.6301
35	5	3	**0.6875**	**0.6726**
40	5	3	0.6950	0.6800

PCA + KNN

PCA	K	Accuracy	F1 Score
40	4	0.9508	0.9502
60	3	0.9468	0.9455
60	5	**0.9556**	**0.9554**

Ensembles

Ensemble Type	Accuracy
Softmax + DT + RF + GB (Stacking)	0.9124
Softmax + KNN (Weighted)	0.9568
5-Model Weighted Ensemble	**0.9596**

(c) Steps Taken to Optimize Performance & Run Time

I used many optimization techniques to make the models run faster:

- PCA reduced the feature size significantly.

- Softmax was improved using batches and L2 regularization.

- Boosting was trained on a reduced PCA subset.

- Ensembles used the best hyperparameters after hyperparameter tuning so as a result it gave the best performance.

Overall evaluation on the basis of accuracy:

- Softmax Basic: around 0.89

- Improved Softmax: around 0.90

- PCA + SoftmaxPolynomial: around 0.92

- Decision Tree: around 0.80

- Random Forest: around 0.88

- Boosting: around 0.69

- KNN (PCA 60): around 0.95

- Best Ensemble: around **0.9596**

(d) Thoughts & Observations

This project was mainly about trying different models and finding the best model by improving accuracy.

- PCA helped a lot. It made almost every model faster.

- KNN surprisingly became the best model after PCA.

- Decision Trees did not work very well.

- Softmax improved a lot after adding regularization.

- Boosting gave the worst results as compared to other models.

- The ensemble of multiple models gave the highest accuracy.

- I observed that combining models is often better than relying on a single model.