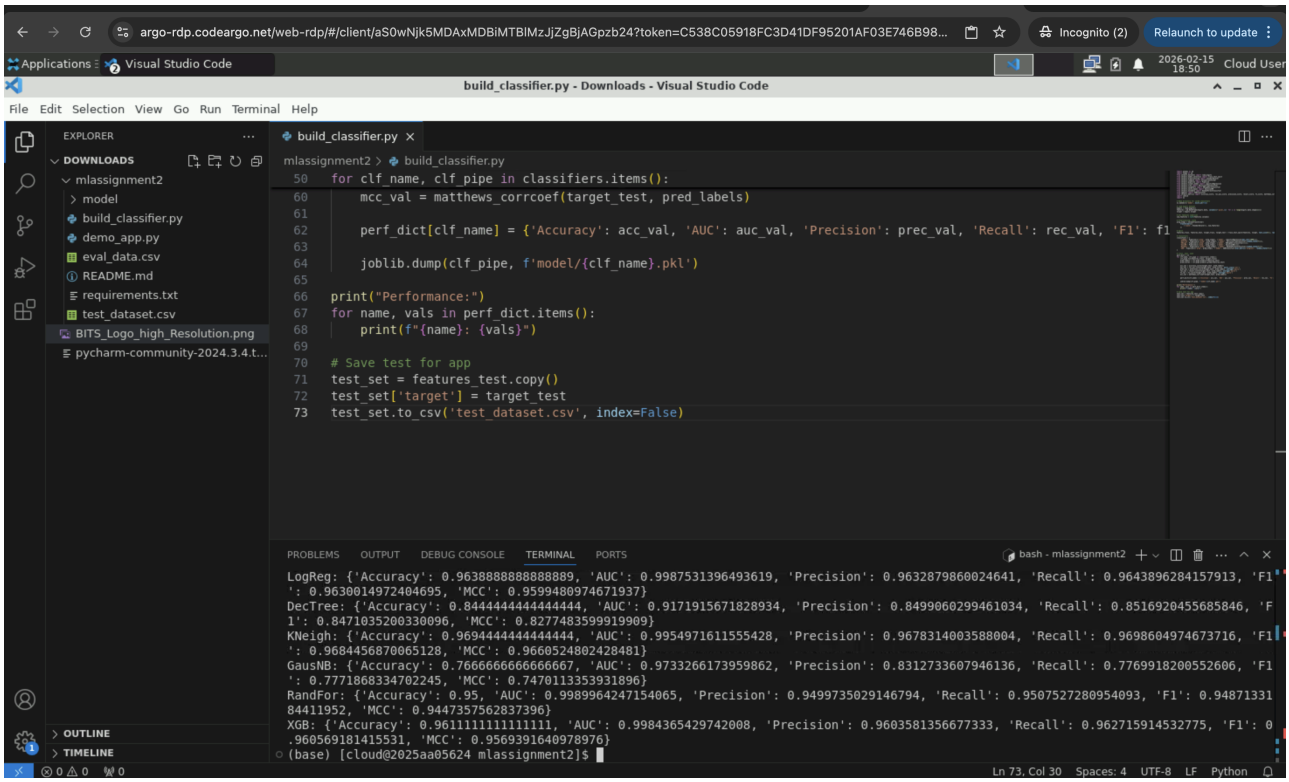
Assignment 2 Submission

Mutyala Sri Lakshmi Akhila (2025aa05624)

1. GitHub Repository: https://github.com/2025aa05624-glitch/mlassignment2.git/
2. Live Streamlit App: https://mlassignment2-lgqfm7shbfurkdzqlcneks.streamlit.app/
3. Screenshot



Classification App for Digits

a. Problem Statement

Goal: Classify images of handwritten digits (0-9) with ML classifiers, evaluate them, and build a Streamlit app for interactive testing.

b. Dataset Description

Digits dataset: 1,797 entries, 64 numeric features (pixels). Classes: 10 digits. Balanced, no misses. From scikit-learn/UCI.

c. Models Used

Implemented:

1. Logistic Regression
2. Decision Tree
3. K-Nearest Neighbors

4. Gaussian Naive Bayes
5. Random Forest Ensemble
6. XGBoost Ensemble

Metrics Comparison

Classifier	Accuracy	AUC	Precision	Recall	F1	MCC
LogReg	0.96	1.00	0.96	0.96	0.96	0.96
DecTree	0.84	0.92	0.85	0.85	0.85	0.83
KNeigh	0.97	1.00	0.97	0.97	0.97	0.97
GausNB	0.77	0.97	0.83	0.78	0.78	0.75
RandFor	0.95	1.00	0.95	0.95	0.95	0.94
XGB	0.96	1.00	0.96	0.96	0.96	0.95

Performance Notes

Classifier	Notes on Performance
LogisticRegression	Reliable, near-perfect on simple patterns.
DecisionTree	Average, tends to overfit pixel noise.
KNN	Excellent, suits distance-based image tasks.
GaussianNaiveBias	Poorest, assumptions don't fit data well.
RandonForest	Good, ensembles help consistency.
XGBoost	Strong, boosting captures details.

Setup

1. Git clone https://github.com/2025aa05624-glitch/mlassignment2.git
2. pip install -r requirements.txt
3. python build_classifiers.py

How to Use

- streamlit run demo_app.py
- Upload CSV (with 'test_dataset' from repo), choose model, see results.

App Link

https://mlassignment2-lgqfm7shbfurkdzqlcneks.streamlit.app/