OPEN ENDED LAB

Classification of MNIST Handwritten Digits Using Machine Learning

Objective

The goal of this open-ended lab is to allow students to experiment with different classification models and compare their performance on the MNIST dataset. Students will:

- Understand the structure of the dataset.
- Train various machine learning models.
- Evaluate the models and analyze their performance.
- Document their findings in a detailed report.

Dataset Description

The dataset consists of handwritten digits represented as 28x28 pixel grayscale images. Each image is labeled with a corresponding digit (0-9). The dataset has been preprocessed as follows:

- Flattened images into a 1D vector of 784 features (28x28 = 784).
- Split into training and testing sets.
- Stored in CSV files (`mnist_train.csv` and `mnist_test.csv`).

Tasks

- Load `mnist_train.csv` and `mnist_test.csv` into a DataFrame.
- Train different classification models.
- Evaluate the model performance using appropriate metrices.
- Do hyperparameter tuning to improve the performance of the models.
- Submit the report in the following format
    - **Introduction** (Overview of dataset)
    - **Methodology** (How the dataset was prepared, models used, hyperparameters)
    - **Results** (Tables, graphs, comparison of model performance)
    - **Discussion** (Which model performed best and why?)
    - **Conclusion** (Summary of findings)

Why Is This Lab Open-Ended?

This lab is open-ended because students have the freedom to choose their models, experiment with hyperparameters, and explore techniques to improve accuracy. Rather than following a rigid set of instructions, students are encouraged to think critically, compare results, and justify their choices in a written report.