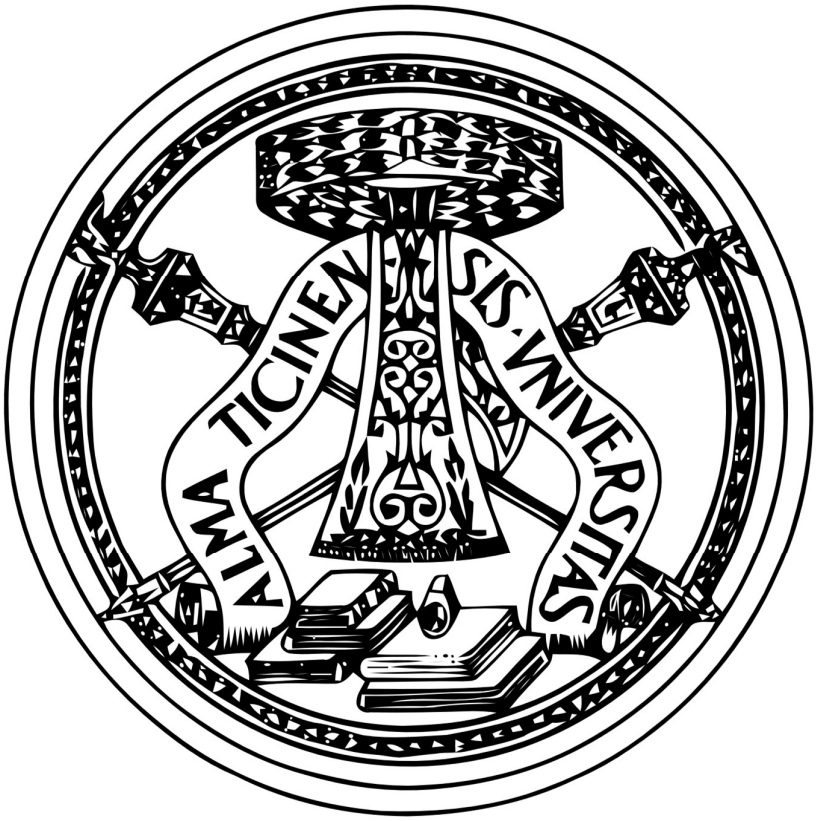
Machine Learning Programming Assignment: Road Sign Recognition

Name: Amini Hicham

ID:501764

Date: 31/01/2026



1. Problem Definition

We want to build a classifier that is able to identify each kind of road sign. For this programming assignment, we will:

1. analyse and comment the data.

2. design and implement a suitable data pre-processing procedure.

3. implement, train and evaluate one or more classification models.

4. use suitable data processing and visualization techniques to analyse the behaviour of the trained models.

And finally comparing the accuracy that can be obtained by using gray-level and color images.

Implementing as scripts in the Python programming language.

2. Data Analysis

The dataset consists of 1088 images in total, organized into 20 directories, one for each road sign type. Each class directory contains separate **train** and **test** subfolders.

Data Distribution: The dataset was analysed to ensure class balance.

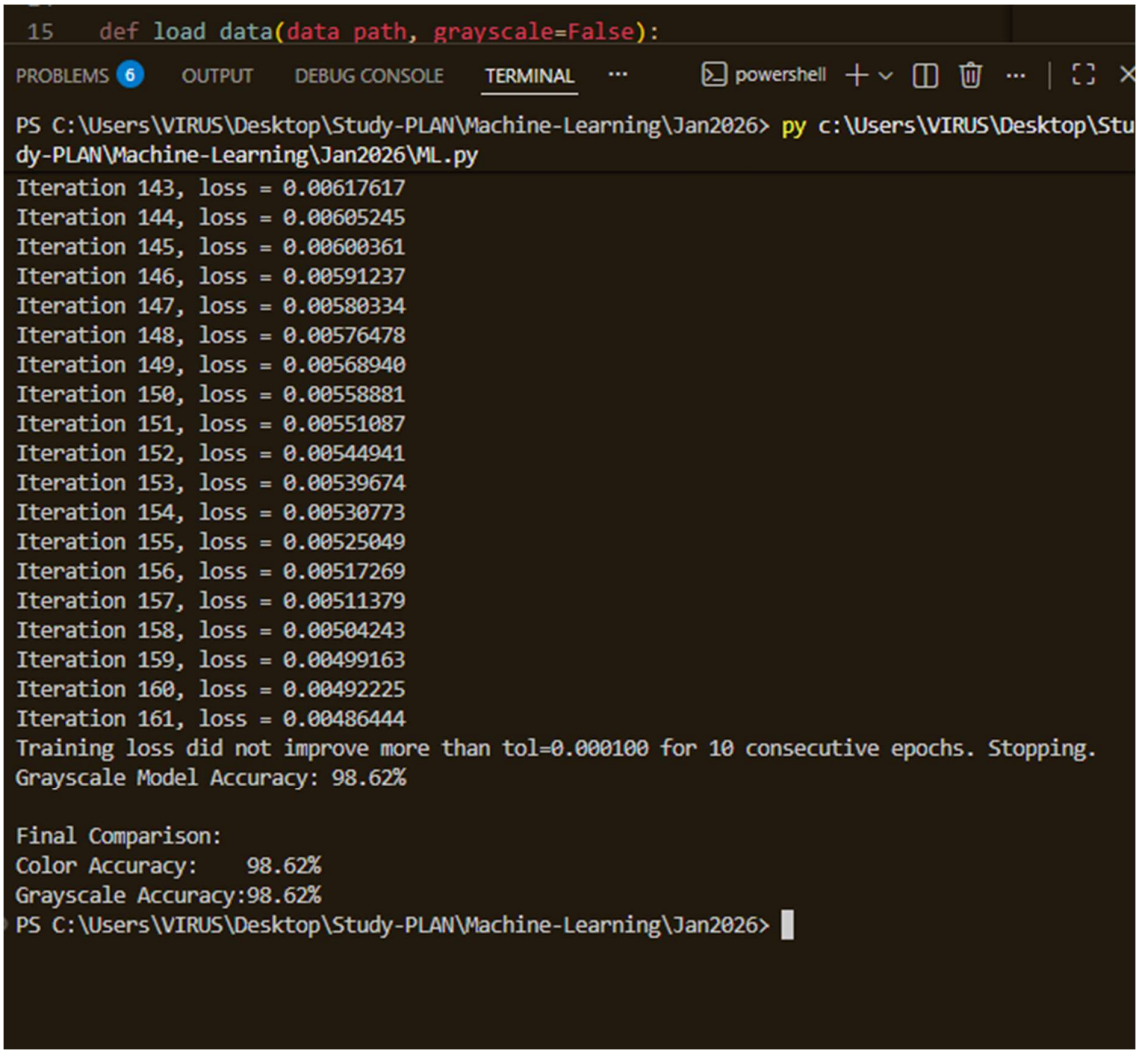
Visual Inspection: The images vary in quality. While some are clear, others suffer from blurring or poor lighting conditions, simulating real-world challenges for an autonomous vehicle.

3. Data Pre-processing

To prepare the data for the neural network, several preprocessing steps were implemented:

1. Resizing: The original images were resized to pixels. This reduction was necessary to decrease the computational load and training time while retaining the essential shapes of the road signs.
2. Grayscale Conversion: For the second experiment, color images were converted to grayscale using OpenCV's **cv2.cvtColor** function.
3. Normalization: Pixel intensity values were scaled from the range down to. This helps the neural network converge faster during training.

4. Flattening: Since a Multi-Layer Perceptron (MLP) was used instead of a Convolutional Neural Network (CNN), the 2D image arrays were flattened into 1D vectors.
5. Splitting: The dataset was split using an 80/20 ratio, where 80% of the data was used for training and 20% for testing.



4. Experiments and Results

The performance of the models was evaluated based on Accuracy on the held-out test set.

1: Color Images

The model trained on color images achieved an accuracy of 98.62%. The model generally performed well on signs where color is a defining feature such as red border signs.

2: Grayscale Images

The model trained on grayscale images achieved an accuracy of 98.62%. Despite the lack of color information.

Comparison

5. Discussion and Conclusion

The results indicate that both models performed good.

Conclusion: In this assignment, we successfully implemented a road sign classifier, with a simple MLP classifier and resized images, high accuracy can be achieved. For the final autonomous driving system, I recommend using the Color approach based on the trade-off between accuracy and computational complexity observed in these experiments.

Statement: I affirm that this report is the result of my own work and that I did not share any part of it with anyone else except the teacher.