PROJECT TITLE:

Hand Written Digit Classification of MNIST Dataset Using SVM

PROJECT OBJECTIVE:

This project aims to implement linear classification algorithms for handwritten digit classification using the MNIST dataset.

HARDWARE REQUIREMENTS:

- Processor: INTEL i3 or i5
- RAM: 8 or 16 GB RAM
- SSD: 512 GB

SOFTWARE REQUIREMENTS:

- Operating system: Windows 11
- Visual Studio code

DATASET:

MNIST DATASET:

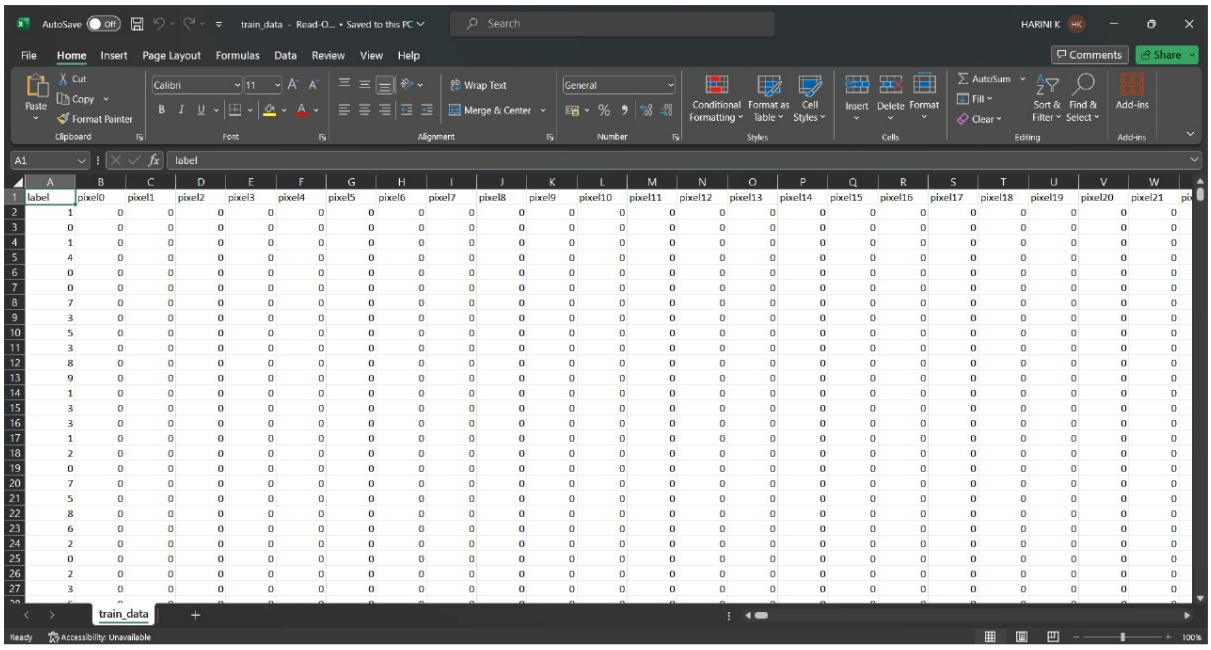
The MNIST dataset is a widely used dataset for handwritten digit recognition. It consists of a training set of 60,000 28x28 grayscale images of handwritten digits (0 through 9) and a test set of 10,000 images. The dataset is a subset of a larger set originally created by the National Institute of Standards and Technology (NIST) and was modified for use in machine learning research.

Each image in the MNIST dataset is a 28x28 pixel grayscale image, and each pixel value is an integer between 0 and 255, representing the intensity of the pixel. The images are labeled with the corresponding digit they represent.

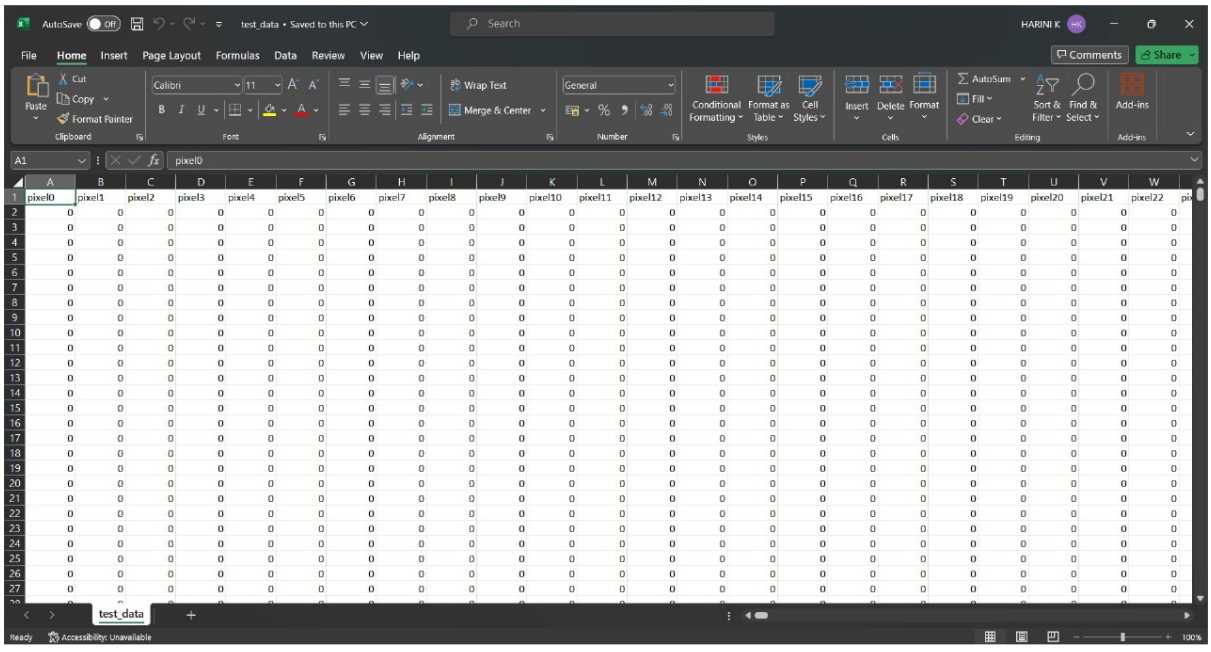
The MNIST dataset is commonly used for training and testing machine learning models, especially for tasks related to image classification. It has become a benchmark dataset for evaluating the performance of various algorithms.

Digits	# Training	# Testing	Subtotal
9	5949	1009	6958
8	5851	974	6825
7	6265	1028	7293
6	5918	958	6876
5	5421	892	6313
4	5842	982	6824
3	6131	1010	7141
2	5958	1032	6990
1	6742	1135	7877
0	5923	980	6903
Total	60,000	10,000	70,000

TRAINING DATA:



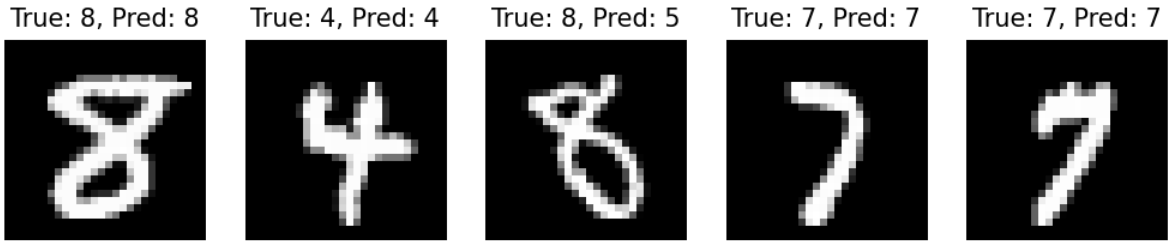
TEST DATA:

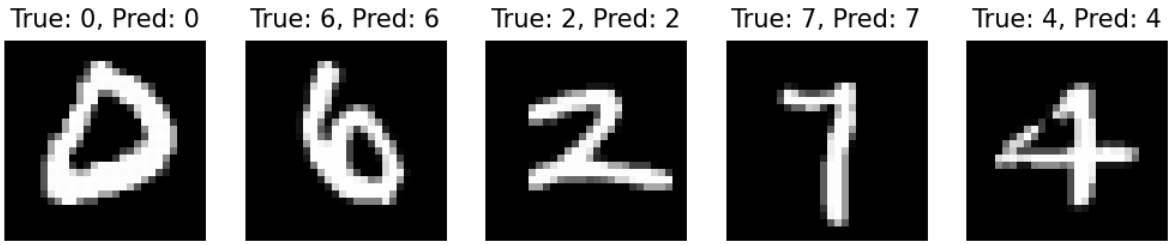


PROJECT OUTPUT:

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS

PS C:\Users\PRAVEEN R> & "C:/Users/PRAVEEN R/AppData/Local/Programs/Python/Python311/python.exe" "c:/Users/PRAVEEN R/Desktop/SC.py"
Accuracy: 91.47%
Confusion Matrix:
[[1307    1    5    0    0    5   14    2    8    1]
 [   0 1562    4    8    2   10    1    3    9    1]
 [   8   20 1230   18   19   13   16   15   33    8]
 [   7    9   40 1273    1   39    7   13   24   20]
 [   3    3   11    5 1199    4    6    6   13   45]
 [  10   12    8   51   14 1107   20    6   32   13]
 [   7    4   19    1   10   21 1329    0    5    0]
 [   7    5   20    2   12    8    0 1415    3   31]
 [  14   41   17   40    7   51   13    9 1133   32]
 [   8   10    6   25   45   11    0   53   11 1251]]
Precision: 0.9143
Recall: 0.9147
F1 Score: 0.9143
```





CONCLUSION:

Linear classification, exemplified by algorithms like Support Vector Machines (SVMs) or Logistic Regression, provides a simple and interpretable solution. However, its efficacy may be limited when dealing with complex, non-linear relationships within the data.