Handwritten Character Recognition Using Machine Learning

1. Introduction

Handwritten character recognition is a crucial area in machine learning and computer vision. The goal of this project is to develop a machine learning model that can recognize handwritten English alphabets using a dataset of labeled handwritten characters.

2. Problem Statement

Recognizing handwritten text is a challenging task due to variations in writing styles. This project aims to develop an accurate classification model using machine learning techniques to identify handwritten alphabets from a given dataset.

3. Objectives

- To preprocess and analyze a dataset of handwritten characters.
- To train a machine learning model for character classification.
- To evaluate the model's accuracy and performance using various metrics.

4. Proposed System

The system utilizes a dataset of handwritten characters, processes the data, and applies a machine learning model to classify the characters. The model is trained using a **Random Forest Classifier**, and its performance is assessed based on accuracy and classification reports.

5. Software & Hardware Requirements

Software Requirements:

- Python
- Jupyter Notebook / Google Colab
- Scikit-Learn
- Pandas, NumPy, Matplotlib, Plotly
- Machine Learning Algorithms (Random Forest Classifier)

Hardware Requirements:

- Computer with at least 4GB RAM
- High-performance processor (i5 or above recommended)

6. Methodology

1. **Data Collection**: The dataset is obtained from **A_Z Handwritten Data.csv**, containing labeled character samples.
2. **Data Preprocessing**: Handling missing values, normalizing data, and visualizing character distributions.
3. **Feature Selection**: Extracting relevant features for model training.
4. **Model Training**: Using a **Random Forest Classifier** to classify the characters.
5. **Testing & Evaluation**: Measuring accuracy, precision, recall, and generating a classification report.

7. Results & Conclusion

The trained model successfully predicts handwritten characters with a high accuracy score. The classification report provides insights into the model's performance across different characters.

	precision	recall	f1-score	support
0	0.99	1.00	0.99	4637
1	0.99	0.99	0.99	2787
2	1.00	1.00	1.00	6175
accuracy			0.99	13599
macro avg	0.99	0.99	0.99	13599
weighted avg	0.99	0.99	0.99	13599