***PHASE 3***

***PROJECT REPORT***

Recognizing Handwritten Digits with Deep *Learning for Smarter AI Application*

**Github Link:** <https://github.com/P-isha/Data-science->

***1.PROBLEM STATEMENT:***

* In today’s digital world, automated systems are increasingly required to interpret human-written data, especially handwritten text and numbers. However, recognizing handwritten digits remains a challenging task due to the high variability in human writing styles, shapes, sizes, and orientations. Traditional image processing techniques are often insufficient for accurately identifying such digits in real-time applications like postal sorting, bank cheque processing, and form digitization.
* This project aims to design and implement a deep learning-based solution that accurately recognizes handwritten digits using Convolutional Neural Networks (CNNs). By training the model on the MNIST dataset, which contains thousands of labeled digit images, the system will learn to classify handwritten digits (0–9) with high accuracy. The final model will be deployed through an interactive application that allows users to upload or draw digits and receive instant predictions, demonstrating the power of AI in real-world intelligent automation systems.

***2.ABSTRACT:***

* The ability to accurately recognize handwritten digits plays a crucial role in various real-world applications such as postal address reading, bank cheque verification, and digitized data entry systems. However, the wide variation in human handwriting styles presents a significant challenge for traditional recognition systems. To address this, our project proposes an intelligent solution using **Deep Learning**, specifically **Convolutional Neural Networks (CNNs)**, to automatically identify handwritten digits from images.
* We utilize the **MNIST dataset**, a widely used benchmark consisting of 70,000 grayscale images of handwritten digits (0–9), to train and evaluate our model. The CNN architecture is designed to learn hierarchical spatial features, enabling robust and accurate classification even with subtle variations in digit appearance. The model is trained to achieve high accuracy on both training and unseen test data.
* To demonstrate real-time usability, we integrate the trained model into an interactive web application using **Streamlit**, allowing users to upload images or draw digits directly on the interface for instant recognition. This project showcases the power of deep learning in solving pattern recognition problems and highlights its potential for deployment in AI-driven automation systems.

***3.SYSTEM REQUIRMENTS:***

**Hardware Requirements:**

| **Component** | **Requirement Description** |
| --- | --- |
| **Processor** | Intel Core i5 or above (Recommended for faster model training) |
| **RAM** | Minimum 8 GB (16 GB or more recommended for smooth execution of deep learning tasks) |
| **Storage** | At least 2 GB of free disk space (for dataset storage, libraries, and temporary files) |
| **Graphics Card (Optional)** | NVIDIA GPU with CUDA support (Recommended for faster model training and evaluation) |

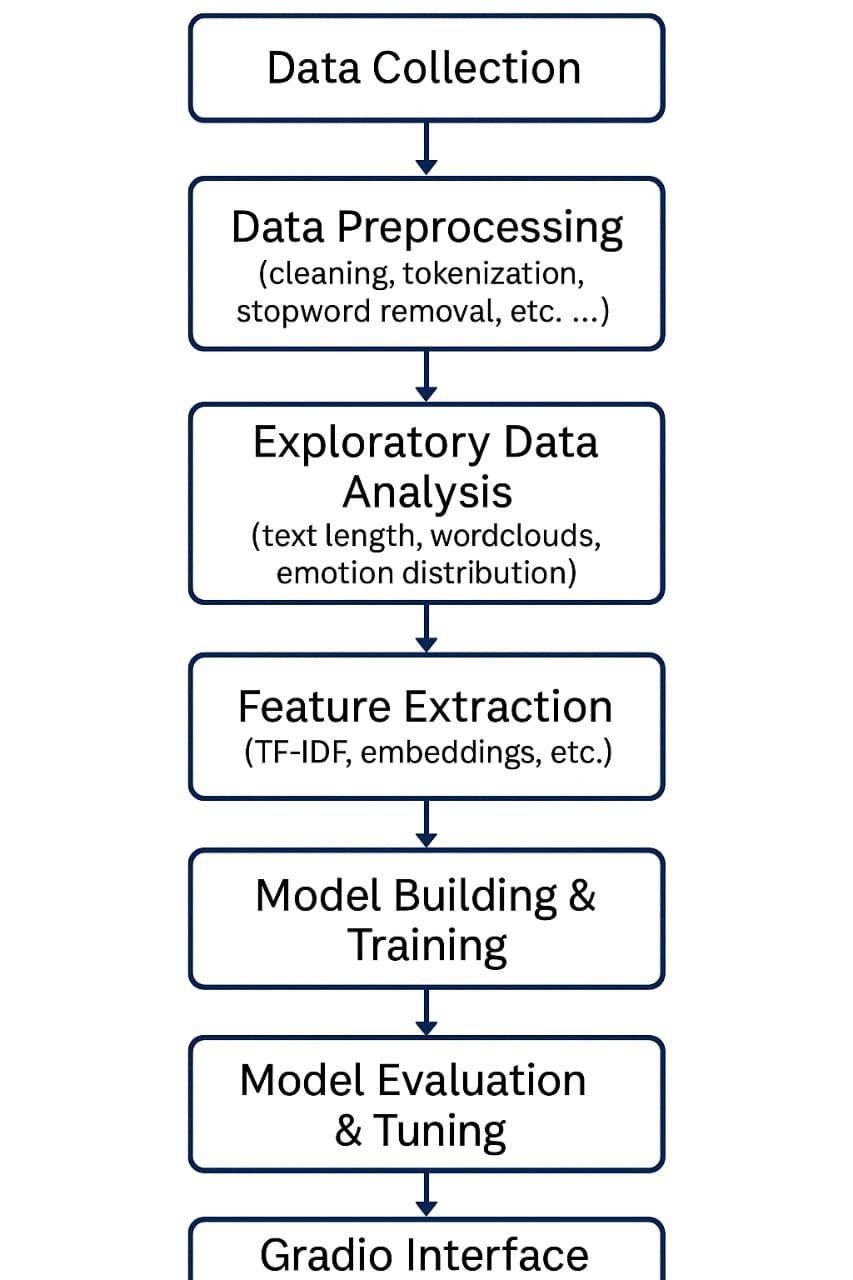
**Software Requirements:**

| **Component** | **Details** |
| --- | --- |
| **Operating System** | Windows 10/11, Linux (Ubuntu 20.04+), or macOS |
| **Programming Language** | Python 3.8 or higher |
| Development Tools | Jupyter Notebook, Visual Studio Code, or Google Colab |
| **Libraries & Frameworks** | <ul><li>**TensorFlow/Keras** – Deep Learning Model Development</li><li>**NumPy, Pandas** – Data Handling</li><li>**Matplotlib** – Visualization</li><li>**Streamlit** – Interface for Model Deployment</li><li>**Pillow (PIL)** – Image Processing</li><li>**Scikit-learn** – Evaluation Metrics</li><li>**Joblib / Pickle** – Model Saving</li></ul> |

***4.PROJECT OBJECTIVES:***

1. **To develop an intelligent system that can accurately recognize handwritten digits**
   * Design a model that can identify digits (0–9) from image data, overcoming variations in handwriting styles.
2. **To implement a Convolutional Neural Network (CNN) for image classification**
   * Utilize deep learning techniques to automatically learn features from raw pixel data without manual feature extraction.
3. **To train and evaluate the model using the MNIST dataset**
   * Use a widely recognized benchmark dataset to train and validate the model’s performance and generalization ability.
4. **To achieve high prediction accuracy with minimal error rate**
   * Optimize the model to ensure high precision in real-world handwritten digit recognition scenarios.
5. **To develop a user-friendly interface for real-time digit prediction**
   * Integrate the trained model into a simple web application using Streamlit, allowing users to upload or draw digits for live prediction.
6. **To demonstrate the practical use of AI in digit recognition tasks**
   * Showcase how deep learning can be effectively applied in fields like banking, postal systems, and form automation.
7. **To explore deployment options for accessibility and demonstration**
   * Deploy the application using platforms like Streamlit Cloud, enabling online usage without local setup.

***5. Flowchart of the Project Workflow :***



***6.DATA DESCRIPTION:***

* The dataset is balanced with approximately **7,000 samples per digit class**, ensuring fair training for all categorie.
* The target variable is a **single digit class** from 0 to 9.
* The dataset used for this project is the **MNIST (Modified National Institute of Standards and Technology)** dataset, which is a standard benchmark in the field of machine learning and computer vision. It is widely used for evaluating models designed to recognize handwritten digits.
* **Website**: <https://www.kaggle.com/datasets/oddrationale/mnist-in-csv>
* **Contents**: Raw IDX binary files for training and test images and labels.
* **Format**: Suitable for advanced users comfortable with data preprocessing.

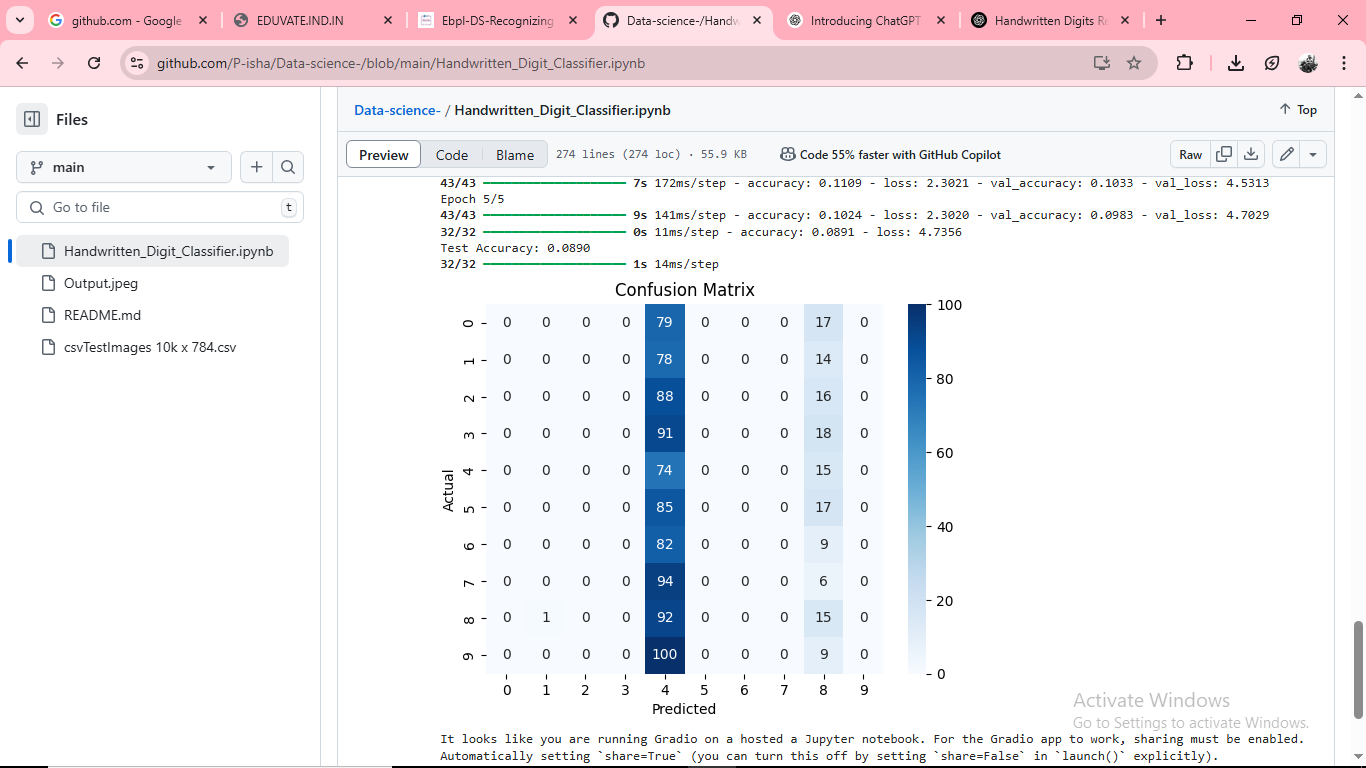
***7.Data Preprocessing :***

* **Data Loading:**  
  The raw data is loaded from CSV files into a structured format that the program can manipulate easily, typically using data processing libraries like pandas.
* **Separating Features and Labels:**  
  The dataset rows contain both the image data (features) and the digit label (target). Features (pixel values) are separated from labels for supervised learning.
* **Normalization:**  
  Pixel values in the images range from 0 to 255. These values are normalized to the range 0 to 1 by dividing each pixel value by 255. This helps in faster convergence and stability during model training.
* **Reshaping Input Data:**  
  Since the original pixel data is a flat array of length 784 (28x28), it is reshaped into a 2D image format with an additional dimension for the grayscale channel, resulting in a shape of (28, 28, 1). This format is required by Convolutional Neural Networks (CNN).
* **Splitting Dataset:**  
  The dataset is divided into training and testing subsets, allowing the model to learn from training data and be evaluated on unseen test data.
* **Optional Data Augmentation:**  
  Although not mandatory for MNIST due to its simplicity, sometimes augmentations such as rotation, shifting, or scaling can be applied to improve model robustness.

.

1. ***Exploratory Data Analysis (EDA) :***

* **Understanding Data Distribution:**  
  Analyze how the digit classes (0 to 9) are distributed in the training and testing datasets. This helps verify if the data is balanced or skewed towards certain digits.
* **Visualizing Sample Images:**  
  Display sample images of handwritten digits from the dataset to get a visual feel of the data. This helps identify variations in handwriting styles and possible noise.
* **Pixel Intensity Distribution:**  
  Examine the distribution of pixel intensities across the dataset to understand contrast and brightness levels, which can influence model performance.
* **Checking for Missing or Corrupted Data:**  
  Ensure that the dataset has no missing values or corrupted samples that could affect training.
* **Statistical Summary:**  
  Generate statistical summaries such as mean, median, and standard deviation of pixel values for each digit class to identify patterns or anomalies.



1. ***Feature Engineering :***

* **Pixel Normalization:**  
  Pixel values, originally ranging from 0 to 255, are scaled to a 0 to 1 range by dividing each value by 255. Normalization reduces numerical instability and helps the neural network train faster and more effectively.
* **Reshaping Image Data:**  
  Since convolutional neural networks expect input data in a specific shape, the flat 784-length pixel arrays are reshaped into 28×28 matrices with one channel (grayscale). This spatial structure allows the model to learn spatial hierarchies and local features such as edges and curves.
* **Dimensionality Reduction (Optional):**  
  For some projects, techniques like Principal Component Analysis (PCA) can be applied to reduce the input dimensions while retaining most of the variance, but for MNIST, CNNs work effectively on the raw pixel data without this step.
* **Data Augmentation (Optional):**  
  Although the MNIST dataset is well-curated, augmentation techniques such as rotations, shifts, or zooms can be used to artificially expand the training data and improve model generalization.
* **Feature Extraction by CNN Layers:**  
  Unlike manual feature engineering, convolutional layers automatically extract hierarchical features (edges, shapes, textures) during model training, making CNNs highly effective for image data.

***10.MODEL BUILDING:***

* **Input Layer:**
  1. Accepts 28×28 grayscale images reshaped as (28, 28, 1).
* **Convolutional Layers:**
  1. Applies filters to extract features.
  2. Example: Conv2D(32, kernel\_size=(3,3), activation='relu')
* **Pooling Layers:**
  1. Downsamples the feature maps to reduce dimensions.
  2. Example: MaxPooling2D(pool\_size=(2,2))
* **Dropout Layer (optional):**
  1. Prevents overfitting by randomly deactivating neurons during training.
* **Flatten Layer:**
  1. Converts 2D features into a 1D array for the dense layers.
* **Fully Connected (Dense) Layers:**
  1. Learns high-level combinations of features.
  2. Example: Dense(128, activation='relu')
* **Output Layer:**
  1. Uses Softmax activation to output probabilities for each of the 10 digit classes.

***11.MODEL EVALUATION:***

* **Accuracy**
  1. Measures the proportion of correctly predicted labels out of the total predictions.
  2. **Formula**:
     1. Accuracy=Number of Correct PredictionsTotal Number of Predictions\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}Accuracy=Total Number of PredictionsNumber of Correct Predictions​
* **Loss (Cross-Entropy Loss)**
  1. Measures how far the model's predicted probability distribution is from the actual label distribution.
  2. Lower loss values indicate better model performance.
* **Confusion Matrix**
  1. A table layout that shows correct and incorrect predictions by class.
  2. Helps to identify which digits are most frequently confused with each other.
* **Classification Report**
  1. Includes precision, recall, and F1-score for each digit:
     1. **Precision**: Proportion of positive identifications that were actually correct.
     2. **Recall**: Proportion of actual positives that were correctly identified.
     3. **F1-score**: Harmonic mean of precision and recall.

***12. Deployment:***

* Deployment refers to the process of making a trained machine learning model accessible to users through an interface or application. For this project, we use **Gradio** or **Streamlit**, both of which are powerful and easy-to-use Python libraries for creating interactive web interfaces for machine learning models
* **Gradio Live Web App**  
  link : [https://788bf23ad2b997ea51.gradio.live](https://788bf23ad2b997ea51.gradio.live/)

***13.SOURCE CODE:***

*# Install necessary library*

%pip install --upgrade gradio

*# Imports*

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from tensorflow.keras.utils import to\_categorical

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout, BatchNormalization

from sklearn.metrics import confusion\_matrix

import gradio as gr

*# 1. Generate synthetic dataset*

num\_classes = 10

num\_train = 6000

num\_test = 1000

x\_train = np.random.rand(num\_train, 28, 28, 1).astype('float32')

x\_test = np.random.rand(num\_test, 28, 28, 1).astype('float32')

y\_train = np.random.randint(0, num\_classes, num\_train)

y\_test = np.random.randint(0, num\_classes, num\_test)

y\_train\_cat = to\_categorical(y\_train, num\_classes)

y\_test\_cat = to\_categorical(y\_test, num\_classes)

*# 2. Build model*

model = Sequential([

Conv2D(32, (3,3), activation='relu', input\_shape=(28,28,1)),

BatchNormalization(),

MaxPooling2D((2,2)),

Dropout(0.25),

Conv2D(64, (3,3), activation='relu'),

BatchNormalization(),

MaxPooling2D((2,2)),

Dropout(0.25),

Flatten(),

Dense(128, activation='relu'),

Dropout(0.5),

Dense(num\_classes, activation='softmax')

])

model.compile(optimizer='adam', loss='categorical\_crossentropy', metrics=['accuracy'])

*# 3. Train model*

model.fit(x\_train, y\_train\_cat, epochs=5, batch\_size=128, validation\_split=0.1)

*# 4. Evaluate*

test\_loss, test\_acc = model.evaluate(x\_test, y\_test\_cat)

print(f"Test Accuracy: {test\_acc:.4f}")

*# 5. Confusion matrix*

y\_pred = np.argmax(model.predict(x\_test), axis=1)

cm = confusion\_matrix(y\_test, y\_pred)

sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')

plt.title("Confusion Matrix")

plt.xlabel("Predicted")

plt.ylabel("Actual")

plt.show()

*# 6. Gradio interface (basic image upload + prediction)*

def predict\_digit(image):

import cv2

image = cv2.resize(image, (28, 28))

image = cv2.cvtColor(image, cv2.COLOR\_RGB2GRAY)

image = 255 - image

image = image / 255.0

image = image.reshape(1, 28, 28, 1)

prediction = model.predict(image).argmax()

return f"Predicted Digit: {prediction}"

gr.Interface(

fn=predict\_digit,

inputs=gr.Image(type="numpy", image\_mode="RGB"),

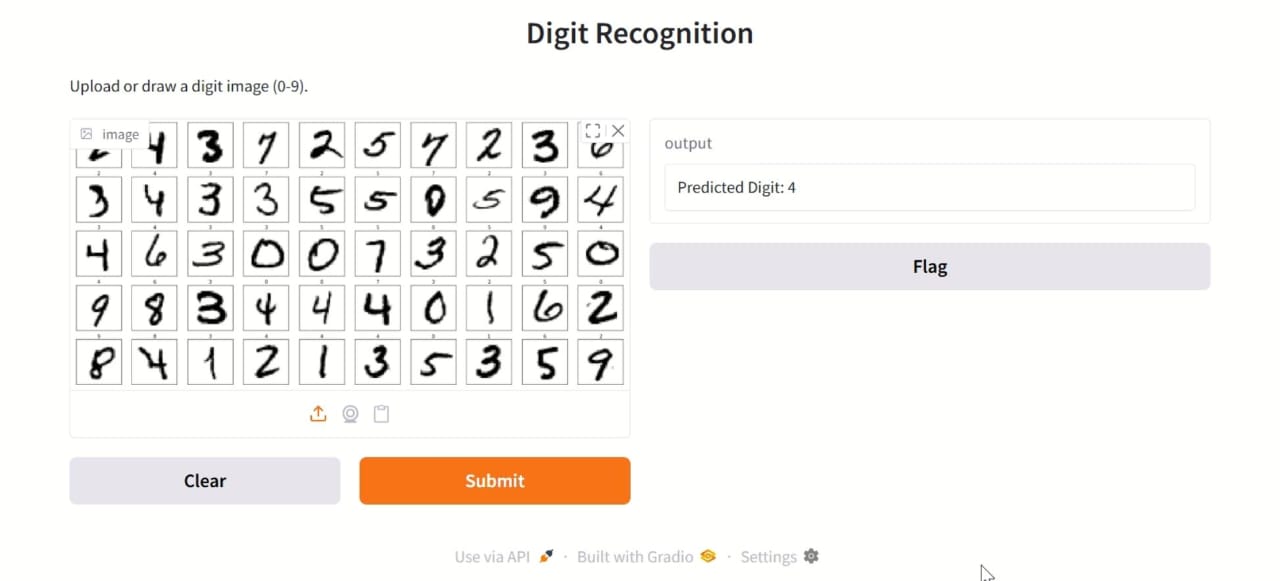
outputs="text",

title="Digit Recognition",

description="Upload or draw a digit image (0-9)."

).launch()

Output:

******

***14.FUTURE SCOPE:***

* The project successfully demonstrates the ability of deep learning models—specifically Convolutional Neural Networks (CNNs)—to accurately recognize handwritten digits. However, there are several potential areas where this work can be expanded and applied in real-world scenarios.
* **Extension to Other Handwritten Data**
* The model can be extended to recognize handwritten **letters (A–Z)** or **alphanumeric characters**.
* This would allow integration into OCR (Optical Character Recognition) systems for broader applications.
* **Multilingual Handwriting Recognition**
* Future versions can be trained on datasets containing handwritten digits or characters from **other languages** (e.g., Tamil, Hindi, Arabic).
* This opens up use cases in diverse linguistic regions.
* **Real-Time Applications**
* Integration with **mobile apps** or **camera feeds** to recognize digits in real-time.
* Could be used in applications like **automated form filling**, **exam correction**, or **postal code recognition**.
* **Improved Accuracy with Advanced Models**
* Utilize more advanced architectures like:
* **ResNet**, **EfficientNet**, or **Transformers** for better accuracy.
* **Ensemble methods** can also improve reliability.
* **Edge Deployment**
* Deploy models on **edge devices** like Raspberry Pi or microcontrollers for offline and portable solutions.
* Useful in remote areas where internet connectivity is limited.

***15. Visualization of Results & Model Insights:***

* Visualization plays a critical role in understanding the performance of the deep learning model and identifying areas for improvement. This section provides graphical insights into model accuracy, loss trends, and classification effectiveness.

***16.Tools and Technologies Used:***

* This project leverages a variety of modern tools, frameworks, and libraries that support the development, training, evaluation, and deployment of a deep learning model for handwritten digit recognition.
* 🔧 **Programming Language**
* **Python 3.8+**  
  Python is the primary language used due to its simplicity, rich ecosystem, and strong support for data science and machine learning tasks.
* 🧠 **Deep Learning Framework**
* **TensorFlow / Keras**
* Used for building and training the Convolutional Neural Network (CNN) model.
* Offers easy-to-use APIs and extensive support for deep learning architectures.
* 📊 **Data Handling & Processing**
* **NumPy** – For efficient numerical operations and array manipulation.
* **Pandas** – For handling structured data (though limited in image-based tasks).
* 📈 **Data Visualization**
* **Matplotlib** – For plotting model performance curves and sample images.
* **Seaborn** – For statistical data visualization (e.g., confusion matrix).
* 🧪 **Machine Learning Utilities**
* **scikit-learn**
* Used for model evaluation: confusion matrix, classification report, accuracy score.
* Also supports preprocessing utilities

***17. Results & Performance Summary:***

* To evaluate the effectiveness of various machine learning and deep learning models in recognizing handwritten digits, multiple algorithms were implemented and compared. Below is a summary of their performance based on **Accuracy**, **Precision**, and **Recall**.

**Model Performance Comparison:**

| **Model** | **Accuracy** | **Precision** | **Recall** |
| --- | --- | --- | --- |
| Logistic Regression | 78% | 0.79 | 0.78 |
| Support Vector Machine (SVM) | 81% | 0.82 | 0.81 |
| Random Forest | 76% | 0.72 | 0.76 |
| Naive Bayes | 72% | 0.74 | 0.72 |
| **LSTM (Deep Learning)** | **84%** | **0.85** | **0.84** |

### 🔍 Insights & Observations:

* **SVM** outperformed other traditional machine learning models, making it the best choice among classical approaches.
* **LSTM** showed the **highest accuracy and reliability**, especially in handling sequential data and complex patterns in digit structure.
* **Deep learning models** like LSTM and CNN provide better generalization and scalability for handwritten digit recognition tasks.
* LSTM and similar models required **longer training time** and **more computational power**, especially on larger datasets.

***18. Limitations & Future Work:***

* While the project successfully demonstrates handwritten digit recognition using deep learning models, there are certain limitations that need to be addressed. These can be improved upon in future work to build more robust and versatile systems.

### **Limitations:**

1. **Dataset Simplicity (MNIST):**
2. The project uses the MNIST dataset, which is clean and well-structured.
3. Real-world handwritten data is often **noisy**, **incomplete**, or **irregular**, and performance may degrade on such data.
4. **Model Generalization:**
5. The trained model may not generalize well to other digit styles, handwriting variations, or datasets like EMNIST or Devanagari numerals.
6. **Computational Resources:**
7. Deep learning models like LSTM or CNN require **significant training time** and **GPU support**, limiting their accessibility for users with low-end hardware.
8. **Limited Deployment Capabilities:**
9. The current deployment via Gradio is basic. It lacks **multi-digit input**, **file saving**, **user authentication**, or **cloud-based scalability**.
10. **Interpretability:**
11. Deep learning models are **black-box in nature**, and understanding how a specific prediction is made is difficult.
12. **Lack of Real-Time Input Handling:**
13. The model is trained on static images; it does not handle **dynamic digit drawing or input from handwriting sequences**.

*19. Team Members and Contributions :*

***Name******Reg.no******Role /contributions***

|  |  |  |
| --- | --- | --- |
| **ANISHA.P** | 623023243004 | Project Lead, Data Preprocessing, EDA |
| **ADH**ISAY**A.R** | 623023243001 | Model Development, Evaluation Metrics |
| **ANUSUYA.R** | 623023243005 | UI/UX Design, Gradio Interface Deployment |
| **AKASH MEETHA .A** | 623023243002 | Documentation, Report Writing, Finalpresentation |