**INFOSYS SPRINGBOARD INTERNSHIP**

Application Logo

An Internship Report

On

Handwritten Digit Recognition Model

**Problem Statement:**

This project entitled *“HANDWRITTEN DIGIT RECOGNITION*” is a practical project based on some trends of computer science. Every day the world is searching new techniques in the field of computer science to upgrade the human limitations into machines to get more and more accurate and meaningful data. The way of machine learning and artificial intelligence has no negative slop it has only the slop having positive direction. This project is a very basic idea of those concepts. This project deals with the very popular learning process called Neural Network. There are various ways by which one can achieve the goal to a desired output, but in machine learning Neural network gives a way that machine learns the way to reach the output.

**Project Mentor:** Narendra Kumar

**Team Members:**

|  |  |
| --- | --- |
| Sai Keerthan Suthraye | saikeerthan\_s@srmap.edu.in |
| Chaithra M | chaithravani11@gmail.com |
| Digambar Dhanagar | diguddit@gmail.com |
| Jyothika Sujith | jyothikasujith7@gmail.com |

**TABLE OF CONTENTS**

1. **Introduction1**
2. **Design2**
3. **About the Dataset3**
4. **Data Preprocessing3**
5. **Methodology4**
   1. LeNet-54
   2. Basic MLP4
   3. Basic CNN4
   4. ResNet-504
6. **Comparative Analysis6**
7. **Result8**

**Conclusion9**

**User Interface10**

**Future Work12**

**Abstract**

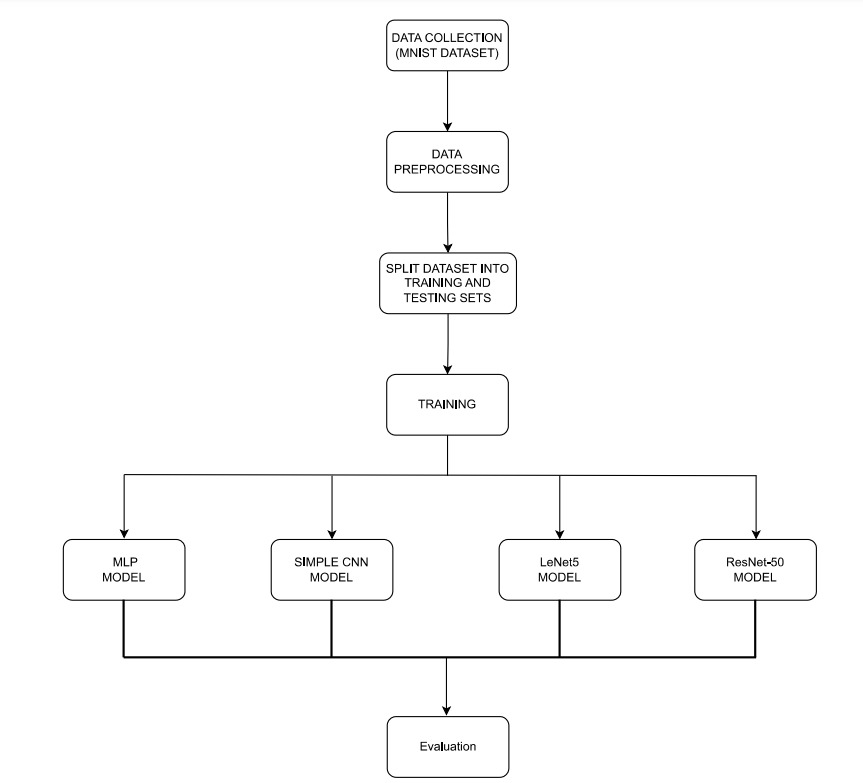
This report compares four network models. LeNet 5, a simple Multi-Layer Perceptron (MLP) a basic Convolutional Neural Network (CNN), with data augmentation and ResNet 50. Using the MNIST dataset. The analysis assesses these models based on their structure training methods and performance results. The findings suggest that the basic CNN with data augmentation performs better than the models in terms of accuracy emphasizing the importance of data augmentation in improving model performance. The LeNet 5 model strikes a balance between accuracy and computational efficiency while the basic MLP despite its accuracy rate is less efficient due to its larger number of parameters. ResNet 50, designed for datasets exhibits lower performance on the MNIST dataset indicating that model complexity should align with the specific task, at hand.

**Introduction**

The field of machine learning has made progress with one of its challenges being the recognition of handwritten digits. This issue is not fundamental but also acts as a gateway, to comprehending intricate image classification tasks. The focus of this report lies on the "Handwritten Digit Recognition" project, which utilizes the MNIST dataset—a component in image recognition.

Within the MNIST dataset are images of digits each tagged with its respective digit. This dataset has been widely utilized for training and testing machine learning models. Our project investigates the performance of four network architectures; LeNet 5, a simple Multi-Layer Perceptron (MLP) a basic Convolutional Neural Network (CNN), with data augmentation and ResNet-50. Through comparing these models, we seek to grasp their variances training approaches and resulting accuracies.

Our research not spotlights the aspects of these models but also underscores practical implementation and performance assessment. We delve into data preprocessing methods, model training protocols and evaluation criteria. Moreover, an intuitive interface was created to enable users to interact with the trained models by uploading images of digits and obtaining predictions.

**Design**

**About the Dataset**

The MNIST dataset comprises 70,000 grayscale images of handwritten digits, with 60,000 images in the training set and 10,000 images in the test set. Each image is 28x28 pixels, normalized and centred. The dataset includes 10 classes, representing the digits 0 through 9. Its simplicity and balanced class distribution make it ideal for benchmarking machine learning algorithms.

* **Total Samples:** 70,000 images
* **Training Set:** 60,000 images
* **Test Set:** 10,000 images
* **Image Size:** 28x28 pixels
* **Image Type:** Grayscale
* **Number of Classes:** 10 (digits 0 through 9)

**Data Preprocessing**

* **Normalization:** Pixel values scaled to [0, 1].
* **Reshaping:** For MLP, images flattened into 784-dimensional vectors; for CNNs, images reshaped to 28x28x1 arrays.
* **One-Hot Encoding:** Labels converted into one-hot encoded vectors.

**Methodology**

**MODEL ARCHITECTURE**

* 1. **LeNet-5**
* **Description:** LeNet 5 is a network containing two convolutional layers, followed by average pooling layers, fully connected layers and a softmax output layer.
* **Training Approach:** Utilized Adam optimizer and sparse entropy loss, for training over 10 epochs with a batch size of 128.
* **Total Parameters:** 61,706
  1. **Basic MLP**
* **Description:** A multi-layer perceptron comprising two dense layers each accompanied by a dropout layer and finalized with a SoftMax output layer.
* **Training Methodology:** Employed Adam optimizer and sparse entropy loss during training for 10 epochs using a batch size of 128.
* **Total Parameters:** 669,706
  1. **Basic CNN with Data Augmentation**
* **Description:** This model features a network with two convolutional layers incorporating max pooling and dropout layers. It also includes layers. Concludes with a SoftMax output layer. Data augmentation techniques like rotations, shifts, shear and zoom were implemented.
* **Training Procedure:** Applied Adam optimizer along with sparse entropy loss for training, over 12 epochs using a batch size of 128.
* **Total Parameters:** 62,006
  1. **ResNet-50**
* ResNet50 is a neural network (CNN) architecture introduced by Microsoft Research in the year 2015. It is based on the ResNet concept that emphasizes "Residual Network."
* The number "50”, in the title indicates that the network consists of 50 layers each layer being deep. The accuracy during training is 97.51% while the accuracy, during testing is 46.15%.

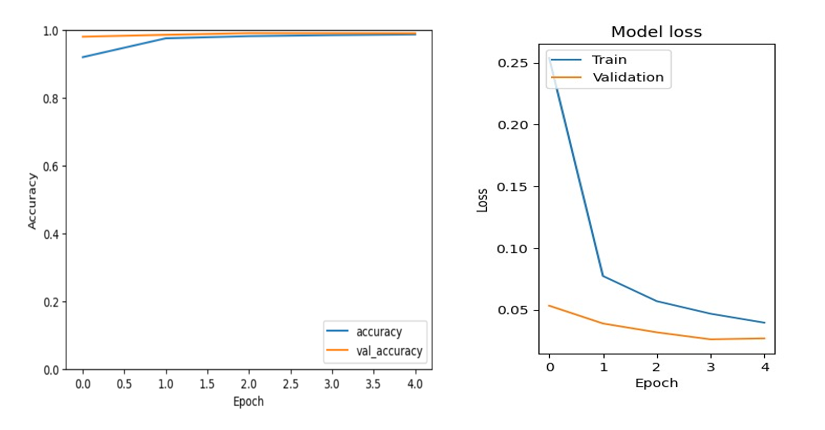
**The poor performance of ResNet-50**

The poor performance of ResNet-50 on the MNIST dataset is primarily due to the model's excessive complexity relative to the simplicity of the MNIST images. ResNet-50, designed for high-resolution, three-channel images like those in the ImageNet dataset, is prone to overfitting when trained on the smaller, low-resolution, single-channel MNIST digits. Additionally, the preprocessing steps—resizing and duplicating channels—may introduce artifacts that further degrade performance. The model's large number of parameters leads to memorizing training data rather than learning generalizable features, resulting in poor test accuracy. Fine-tuning fewer layers and using a less complex model could improve results.

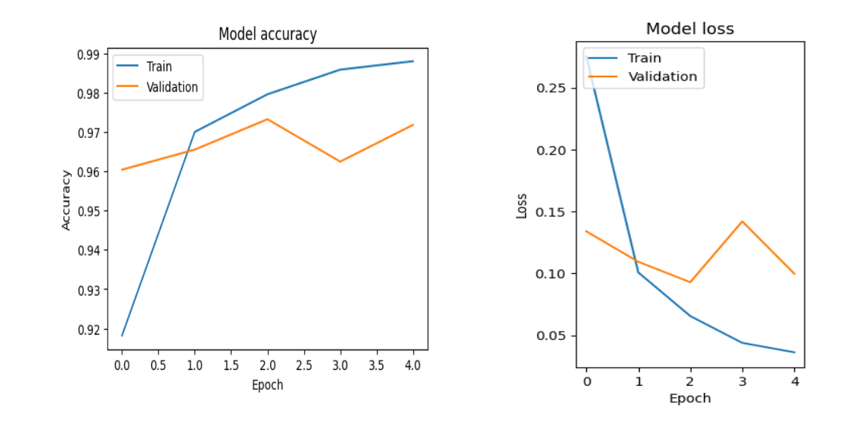
**6. Comparative Analysis:**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Training Accuracy | Validation Accuracy | Parameters |
| LeNet-5 | ~99.2% | ~98.9% | 61,706 |
| Basic MLP | ~98.5% | ~98.0% | 669,706 |
| Basic CNN + Aug | ~99.4% | ~99.4% | 62,006 |
| ResNet-50 | ~97.51% | ~46.15% | - |

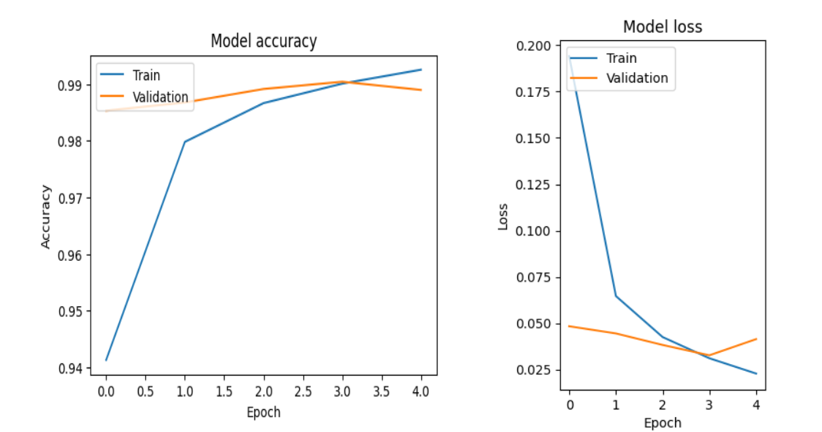
* **LeNet-5:** Achieves high accuracy with relatively few parameters, but lacks data augmentation which could improve generalization.
* **Basic MLP:** Achieves good accuracy but requires significantly more parameters, making it less efficient.
* **Basic CNN with Data Augmentation:** Achieves the highest accuracy with efficient parameter usage, demonstrating the benefits of data augmentation.
* **ResNet-**50: Poor performance when compared to other models.

****

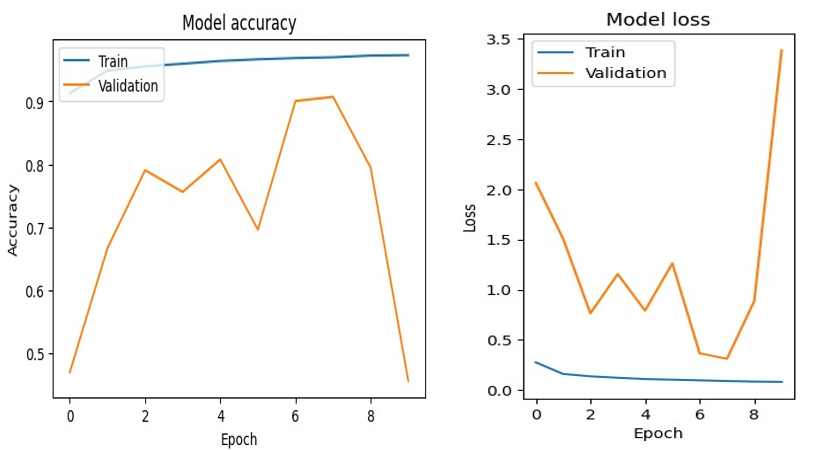
**LeNet-5 Model**



**Basic MLP Model**

****

**Basic CNN Model**



**ResNet-50 Model**

**Results**

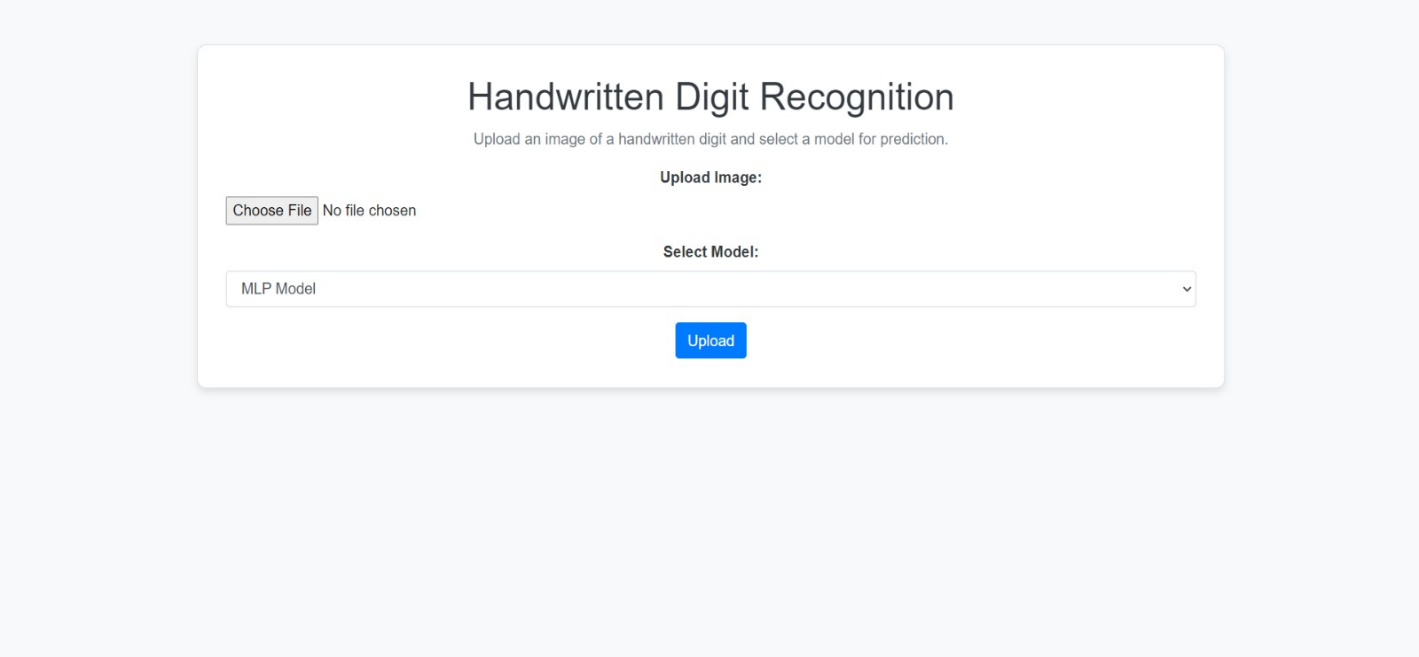
* **LeNet-5:** Training accuracy ~99.2%, validation accuracy ~98.9%.
* **Basic MLP:** Training accuracy ~98.5%, validation accuracy ~98.0%.
* **Basic CNN with Data Augmentation:** Training accuracy ~99.6%, validation accuracy ~99.4%.
* **ResNet-50:** training accuracy~97.51%, Test accuracy ~46.15%.

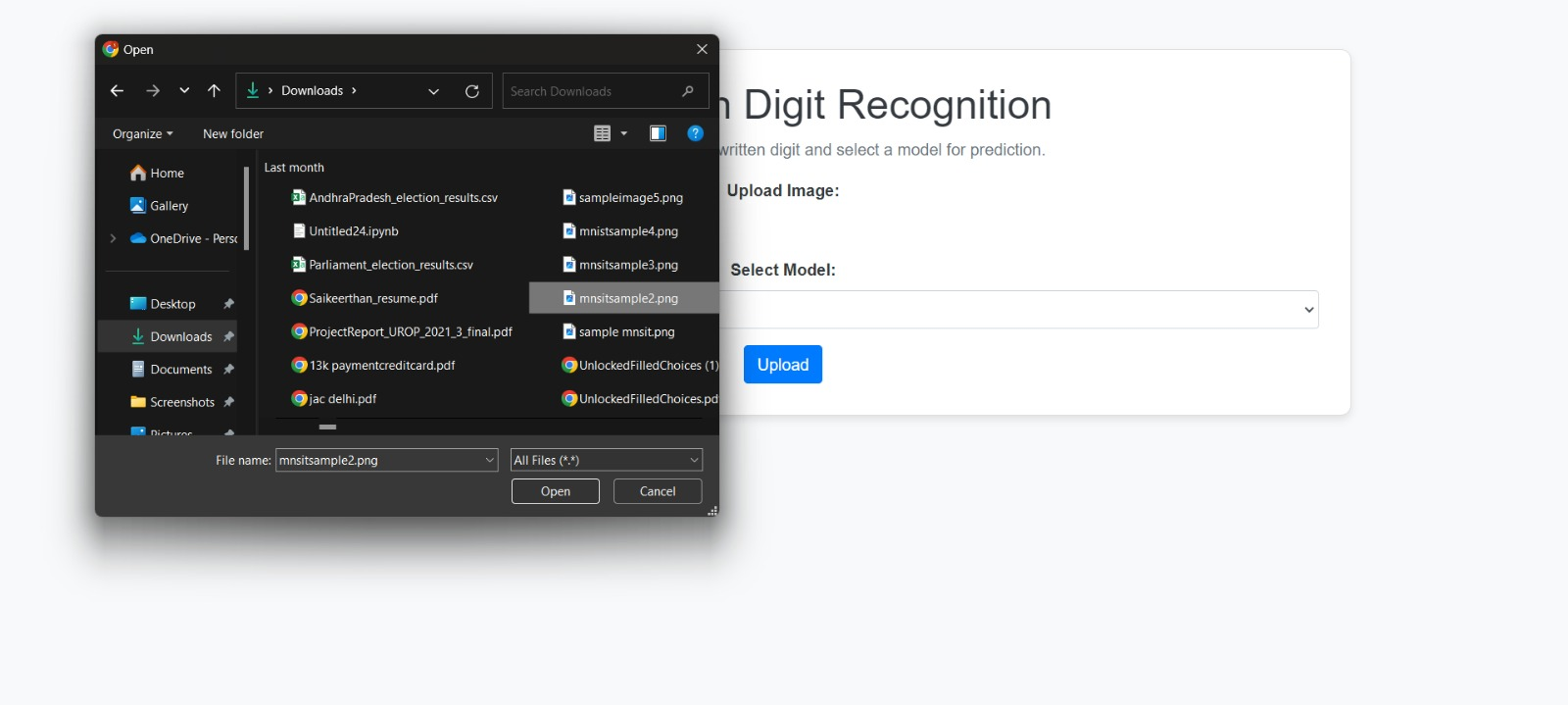
**Conclusion**

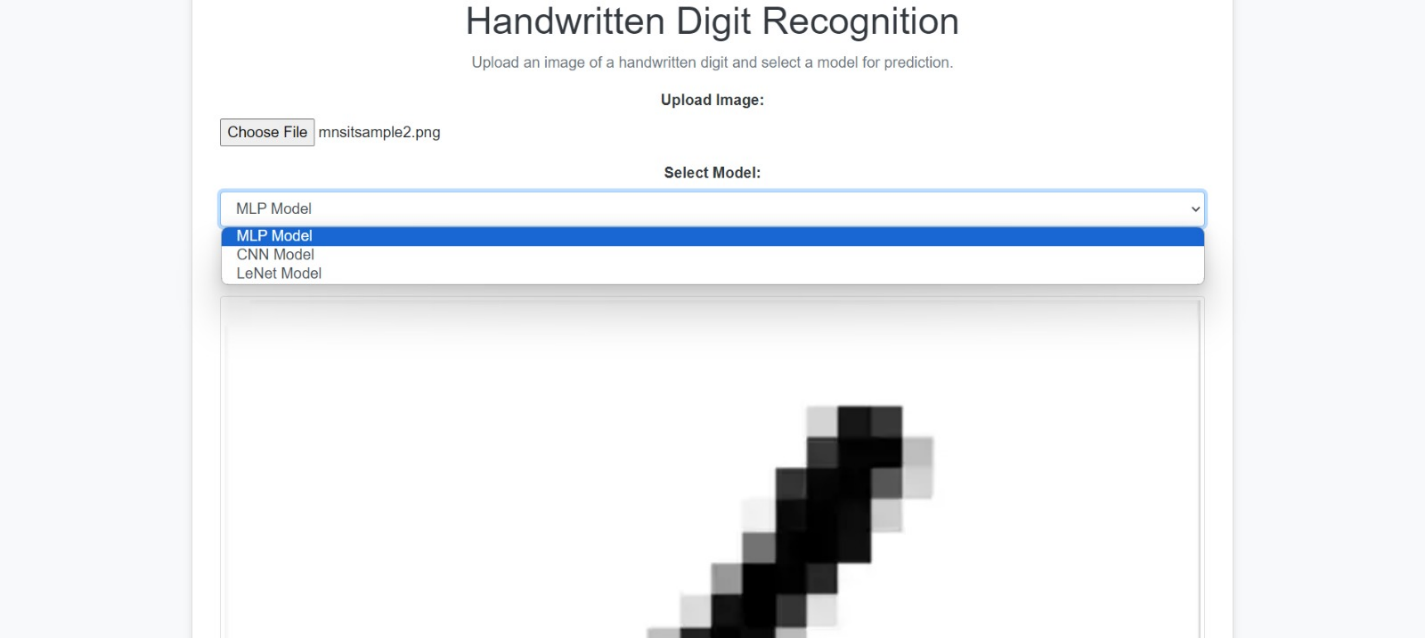
The research shows that using the CNN, with data enhancement performs better than both LeNet 5 and the basic MLP in terms of training and validation accuracy. Data enhancement greatly improves the model’s capacity to adapt to data leading to performance. The LeNet 5 model strikes a balance between accuracy and computational efficiency whereas the basic MLP although achieving accuracy is less efficient because of its larger parameter count. ResNet-50 gives poor performance when compared to other models.

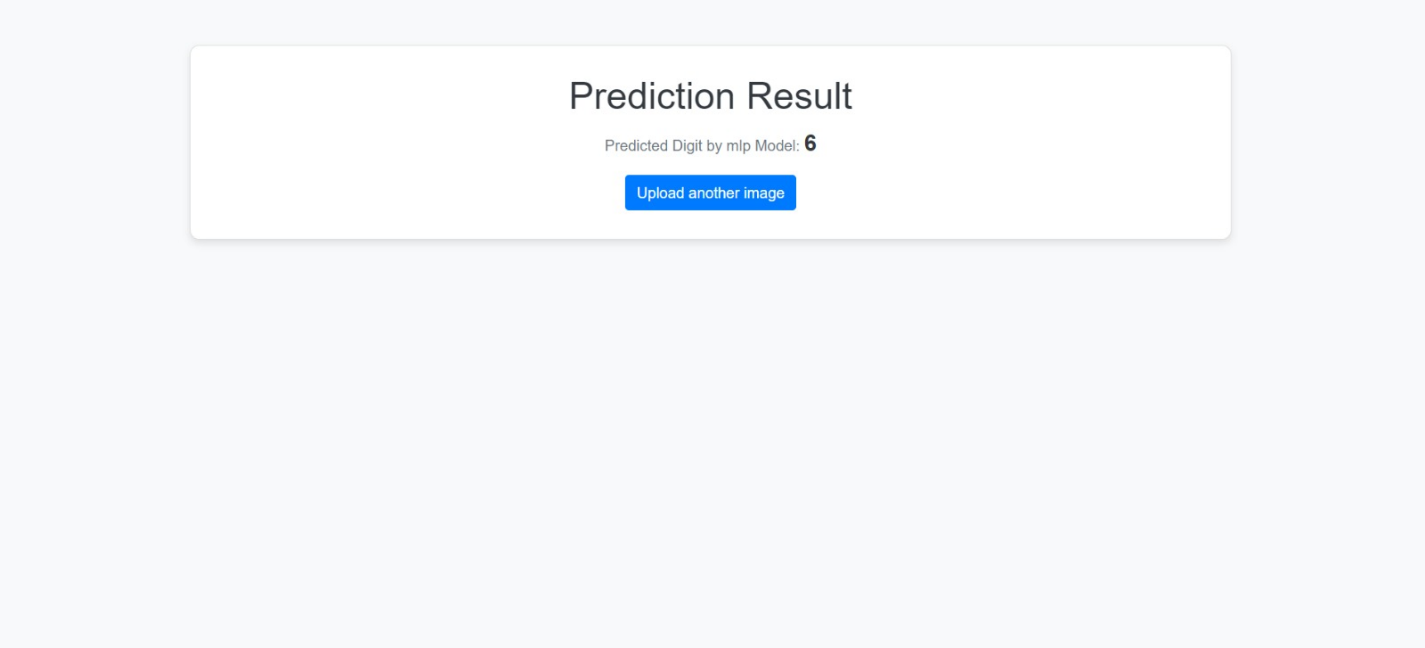
**User Interface**

A Flask-based web application was developed to provide a user-friendly interface for uploading images and receiving predictions from the trained models.



****

****

****

**Future Work**

* **Data Augmentation:** Incorporate data augmentation techniques to improve model robustness and generalization.
* **Efficiency Considerations:** For resource-limited applications, the LeNet-5 model offers a good balance between performance and efficiency.
* **Experimentation:** Further experimentation with advanced architectures and hyperparameters could yield even better performance.