HANDWRITTEN DIGIT RECOGNITION WITH LENET5 MODEL IN PYTORCH

PROJECT REPORT BY ROHAN PODDAR (GROUP 2)

INFOSYS SPRINGBOARD INTERNSHIP

OVERVIEW

•Handwritten digit recognition is a fundamental problem in pattern recognition and machine learning.

•Significant progress has been made with advancements in deep learning, particularly Convolutional Neural Networks (CNNs).

•LeNet-5, proposed by Yann, is one of the pioneering CNN architectures for handwritten digit recognition.

PROJECT FOCUS

THIS REPORT WILL DETAIL THE METHODOLOGY USED TO IMPLEMENT AND EVALUATE THE LENET-5, MLP, AND

CNN MODELS. IT WILL COVER ASPECTS INCLUDING DATA PREPROCESSING, MODEL ARCHITECTURES, TRAINING

PROCEDURES WITH OPTIMIZED HYPERPARAMETERS, AND RIGOROUS EVALUATION USING ESTABLISHED METRICS.

THE REPORT WILL ALSO INCLUDE COMPARATIVE ANALYSES OF THE MODELS' PERFORMANCE, DISCUSSING THEIR

COMPUTATIONAL EFFICIENCY AND PROVIDING RECOMMENDATIONS FOR FURTHER IMPROVEMENTS. THROUGH THIS

COMPREHENSIVE EVALUATION, THE REPORT AIMS TO CONTRIBUTE INSIGHTS INTO THE EFFICACY OF DIFFERENT

NEURAL NETWORK ARCHITECTURES FOR HANDWRITTEN DIGIT RECOGNITION.

OBJECTIVE

The primary objective of this project is to evaluate and compare the effectiveness of LeNet-5, MLP, and CNN models in accurately classifying handwritten digits from the MNIST dataset. By implementing these models, the project aims to analyze their performance metrics such as accuracy, precision, recall, and F1-score. Additionally, the project aims to provide insights into the strengths and weaknesses of each model, identify the best-performing architecture for this specific task, and explore potential applications in digit recognition tasks.

INSIGHTS

- Analyze strengths and weaknesses of each model.
- Identify the best-performing architecture.
- Explore potential applications in digit recognition tasks.

DATASET DESCRIPTION

MNIST Handwritten Digit Dataset:

- Contains 10,000 grayscale images of digits (0-9).
- Each image is 28x28 pixels.
- Divided into 7,000 training images and 3,000 testing images.

DATA PREPROCESSING

One Hot Encoding:
•Convert labels to one hot encoding if necessary.

Data Reshaping:
•Reshape data to fit input dimensions required by the LeNet-5 model.

Data Augmentation:
•Apply techniques to increase the diversity of the training data (e.g., rotations, shifts, flips).

Data Sampling:
•Sample the dataset to balance classes or reduce size for quicker iterations.

MODELS DESCRIPTION

LENET-5

● **Architecture**: Classic CNN architecture consisting of two convolutional layers with average

pooling, followed by three fully connected layers.

● **Training**: Trained for 20 epochs with SGD optimizer and cross-entropy loss.

MULTI-LAYER PERCEPTRON (MLP)

● **Architecture**: It consists of a single hidden layer with ReLU activation followed by a SoftMax

output layer.

● **Training**: Trained for 20 epochs with Adam optimizer and cross-entropy loss

CONVOLUTIONAL NEURAL NETWORK (CNN)

● **Architecture**: Custom CNN with three convolutional layers, batch normalization, maxpooling,

and dropout for regularization.

● **Training**: Trained for 20 epochs with Adam optimizer and categorical cross-entropy loss.

FLASK APPLICATION

User Uploads Image: User uploads an image through the web interface.

Image Received by Flask Server: The Flask server receives the uploaded image.

Image Preprocessing: The server preprocesses the image if necessary (e.g., resizing, normalization).

Model Predictions: The pre-processed image is passed to three different models for prediction.

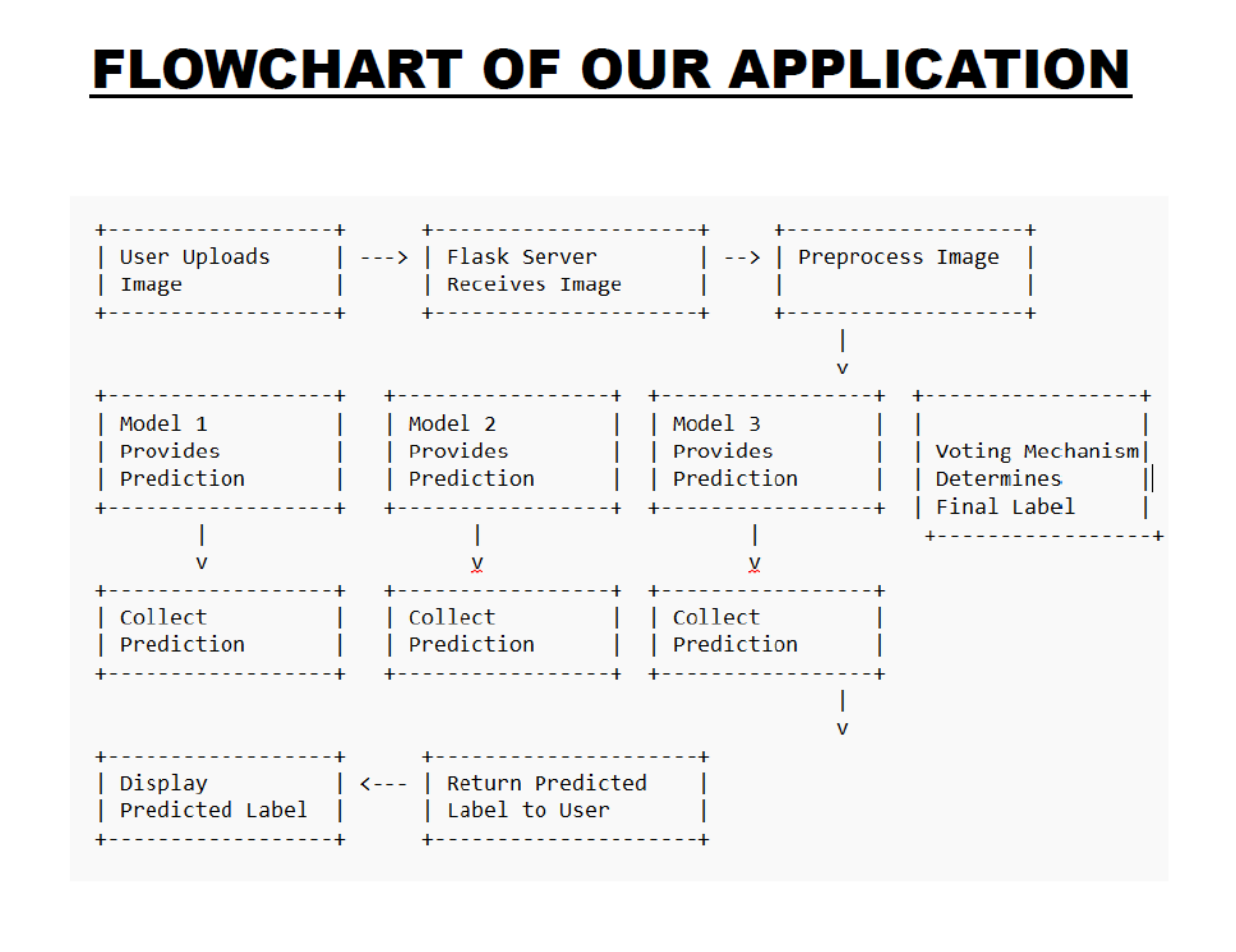
Model 1: Provides its prediction.

Model 2: Provides its prediction.

Model 3: Provides its prediction.

Voting Mechanism: The predictions from all three models are collected and a voting mechanism is used to determine the final class label.

Result Displayed: The predicted label is displayed to the user.



Comparative Analysis

Feature	LeNet 5	MLP	Custom CNN
Architecture	Convolutional Layers + Fully Connected Layers	Fully Connected Layers	Convolutional Layers + Fully Connected Layers
Objective	Digit Classification	Digit Classification	Digit Classification
Parameters	61,706	397,510	114,538
Training Time (per epoch)	2 minutes	22 seconds	2 minutes
Train Accuracy	99.33%	99.94%	98.90%
Test Accuracy	98.70%	97.95%	97.50%
Activation	ReLU	ReLU	ReLU
Input Size	1*28*28	784 (flattened 28*28)	1*28*28
Hidden Layers	2 Conv, 2 Pool, 1 FC	1 FC	3 Conv, 3 Pool, 1 FC, 1 Dropout
Output Size	10 (Classes)	10 (Classes)	10 (Classes)
Epochs	10	20	20
Optimizer	Adam	Adam	Adam
Loss Function	Cross -Entropy	Cross- Entropy	Categorical Cross-Entropy
Learning Rate	0.001	0.001	0.001
Complexity	Low	Low	High
Performance	Moderate	Moderate	High
Strengths	Simple, Efficient	Simple, Efficient	High Performance, Regularization
Weakness	Limited Capacity	Lacks Special Feature Extraction	Computational Intensive

PERFORMANCE SUMMARY

- LeNet-5 and Custom CNN outperform MLP in validation and test accuracies.

- CNN shows the best overall performance with highest test accuracy.

KEY TAKEAWAYS

- MLP provides baseline performance but lacks ability to capture spatial features.
- LeNet-5 and CNN effectively leverage convolutional layers for higher accuracy.
- CNN demonstrates superior performance in extracting and learning complex features.

THANK YOU

INFOSYS SPRINGBOARD INTERNSHIP