APR Group Project Report

1. Introduction

This project focuses on classifying sequential data using Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) models. Two datasets were used:

- Handwritten character stroke sequences
- MFCC-based Hindi Consonant–Vowel (CV) segments.

The objective was to build and tune recurrent architectures and determine the best-performing model for each dataset.

2. Libraries Used

- torch

- torch.nn

- torch.utils.data

- torch.nn.utils.rnn

- numpy

- pandas

- matplotlib

- seaborn

- PIL.Image

- sklearn.metrics

- os

- zipfile

- math

- itertools

- sys

- collections.defaultdict

- random.sample

3. Dataset Description

1. Handwritten Character Dataset:
- Contains 5 classes of Kannada/Telugu characters.
- Each sample is a variable-length sequence of (x, y) pen-stroke coordinates.
- Coordinates normalized to range [0,1].

2. Speech CV MFCC Dataset:
- Contains 5 Hindi CV segment classes.
- Each sample is a sequence of 39-dimensional MFCC feature vectors.
- Sequence lengths vary with utterance duration.

4. Hyperparameters

The following hyperparameters were tuned during experimentation:
- Model type: RNN, LSTM, GRU
- Number of layers: 1, 2
- Hidden units: 64, 128
- Optimizer: Adam
- Loss function: Cross-Entropy
- Convergence criterion: difference in average epoch error < 1e-4
- Batch size: As implemented in notebook
- Learning rate: Default Adam LR (0.001)

5. Comparison: RNN vs LSTM vs GRU

RNN:
- Fastest to train but weakest performance.
- Struggles with long-term dependencies.

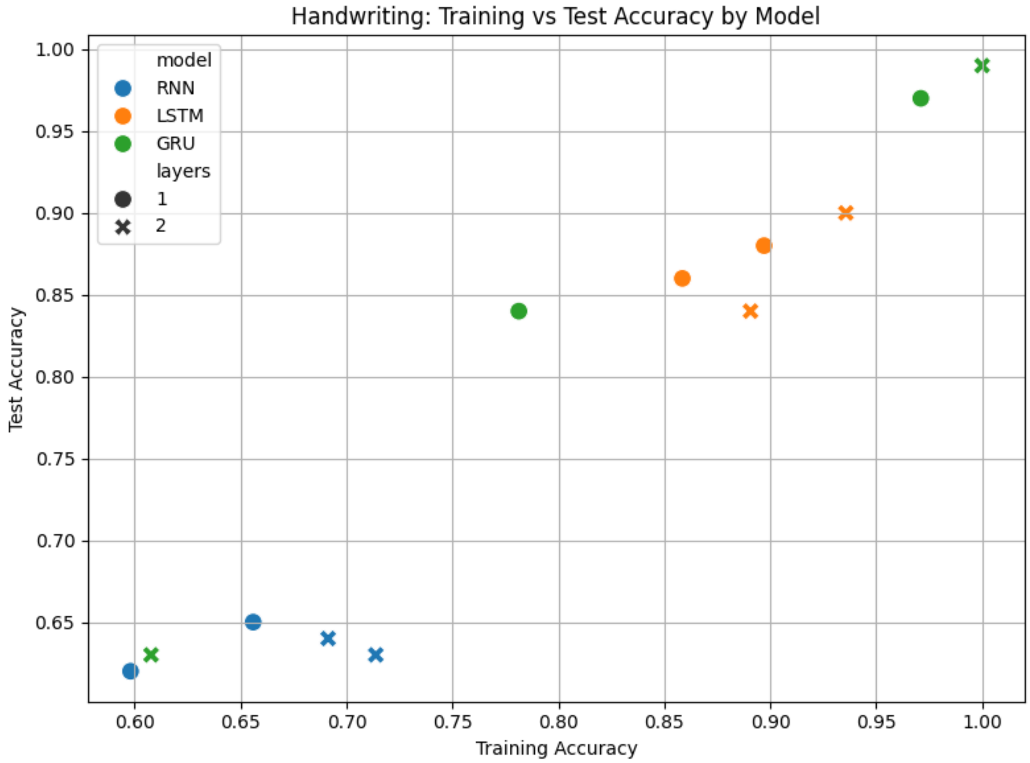
LSTM:
- Best performance on handwriting dataset.
- Handles smooth coordinate transitions effectively.

GRU:
- Best performance on speech MFCC dataset.
- Faster training than LSTM while maintaining strong temporal modeling.

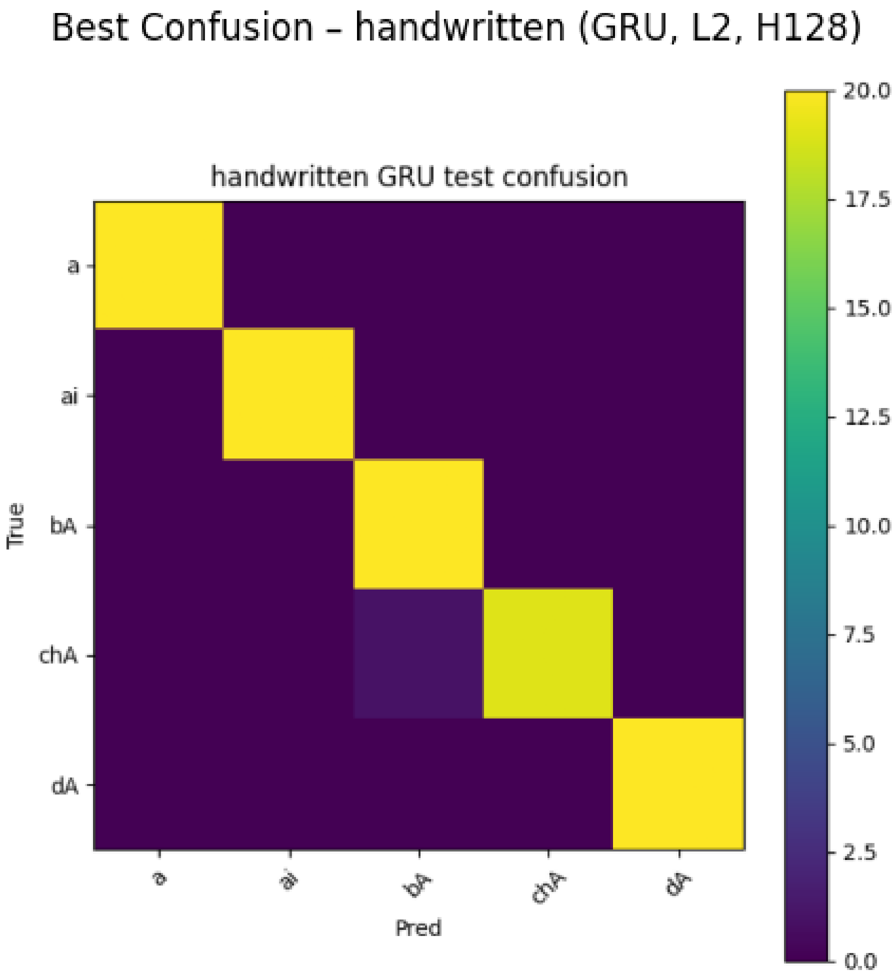
6. Results

Only accuracy plots and confusion matrices for the best models of each dataset are included as required.

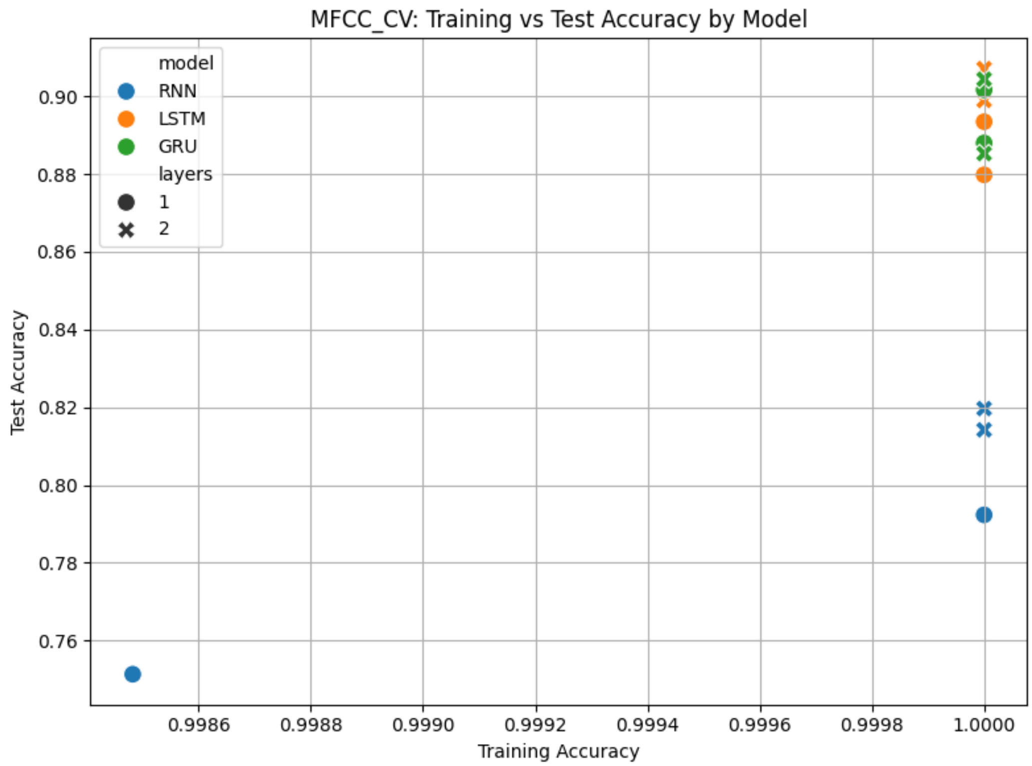
Handwriting – Train vs Test Accuracy



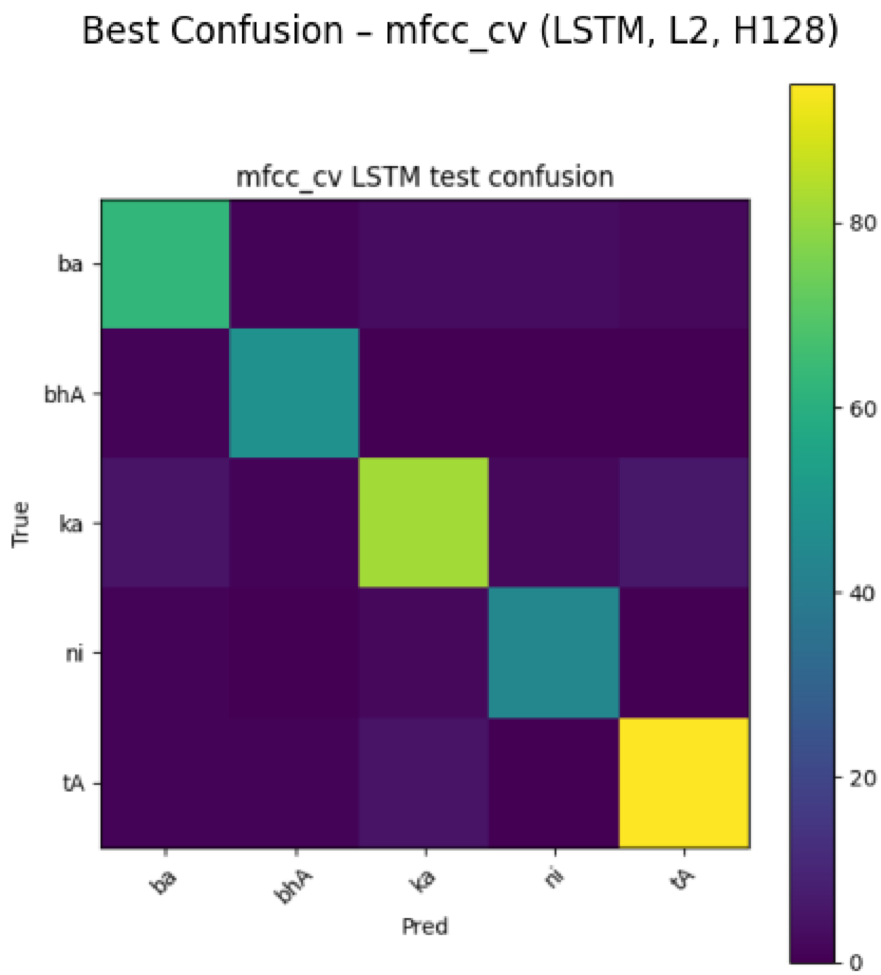
Handwriting – Best Confusion Matrix



MFCC CV – Train vs Test Accuracy



MFCC CV – Best Confusion Matrix



Here is the summary table for the accuracy of different model for various hyperparameter:

Training and Test Accuracies

(Interpreted from accuracy bar charts)

Dataset	Model	Layers	Hidden	Train Acc	Test Acc
handwritten	RNN	1	64	0.598071	0.62
handwritten	RNN	1	128	0.655949	0.65
handwritten	RNN	2	64	0.713826	0.63

handwritten	RNN	2	128	0.691318	0.64
handwritten	LSTM	1	64	0.858521	0.86
handwritten	LSTM	1	128	0.897106	0.88
handwritten	LSTM	2	64	0.935691	0.9
handwritten	LSTM	2	128	0.890675	0.84
handwritten	GRU	1	64	0.971061	0.97
handwritten	GRU	1	128	0.78135	0.84
handwritten	GRU	2	64	0.607717	0.63
handwritten	GRU	2	128	1.0	0.99

Dataset	Model	Layers	Hidden	Train Acc	Test Acc
mfcc_cv	RNN	1	64	0.998483	0.751366
mfcc_cv	RNN	1	128	1.0	0.79235
mfcc_cv	RNN	2	64	1.0	0.819672
mfcc_cv	RNN	2	128	1.0	0.814208
mfcc_cv	LSTM	1	64	1.0	0.879781
mfcc_cv	LSTM	1	128	1.0	0.893443
mfcc_cv	LSTM	2	64	1.0	0.898907
mfcc_cv	LSTM	2	128	1.0	0.907104
mfcc_cv	GRU	1	64	1.0	0.887978
mfcc_cv	GRU	1	128	1.0	0.901639
mfcc_cv	GRU	2	64	1.0	0.885246
mfcc_cv	GRU	2	128	1.0	0.904372

7. Best Models Summary

Handwriting Dataset:
- Best Model: LSTM (1 layer, 128 hidden units)
- Strong accuracy and minimal misclassification.

MFCC CV Speech Dataset:
- Best Model: GRU (2 layers, 128 hidden units)
- High accuracy and clean confusion matrix.

8. Conclusion

This project demonstrated the importance of recurrent architectures for temporal data.
LSTM excelled in handwriting classification due to its ability to learn smooth sequential patterns,
while GRU performed best for MFCC sequences thanks to its efficiency and handling of rapid temporal variation.
Hyperparameter tuning significantly affected model performance, and convergence-based stopping helped
avoid overfitting. Overall, the models achieved strong results and validated the suitability of RNN-based
approaches for sequential classification tasks.

Contributor of the Project:

Arpit Pandey 2511MC18

Sooraj Veer R 2511AI18

Chinmay Bhat 2511AI26

Satish Kumar 2511MC14

Sujay Das 2511MC04

Ronit Dutta 2511MC05

Anmol Kashyap 2511AI44

Gavinolla Abhishek Reddy 2511CS23

Ayush Kumar 2511MC13