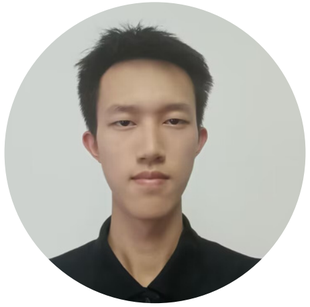
CS 172B

AutoAudio

Sound-Based Classification of Traffic

Team

Redford



Jiaye

Ceci

Lazar

Sebastian

Project Overview

Objective: Develop an algorithm to differentiate vehicle type using their sound emissions

Purpose: Advance autonomous acoustic sensing capabilities to accelerate deployment of smart cities and infrastructure

Data: 2 open-source datasets of vehicle sound data

1. New Delhi, India (**MVDA**)
2. Munich, Germany (**IDMT**)

Questions

Question 1

What model is best suited for general classification of ground vehicles?

Question 2

How do the models respond differently to the 2 datasets?

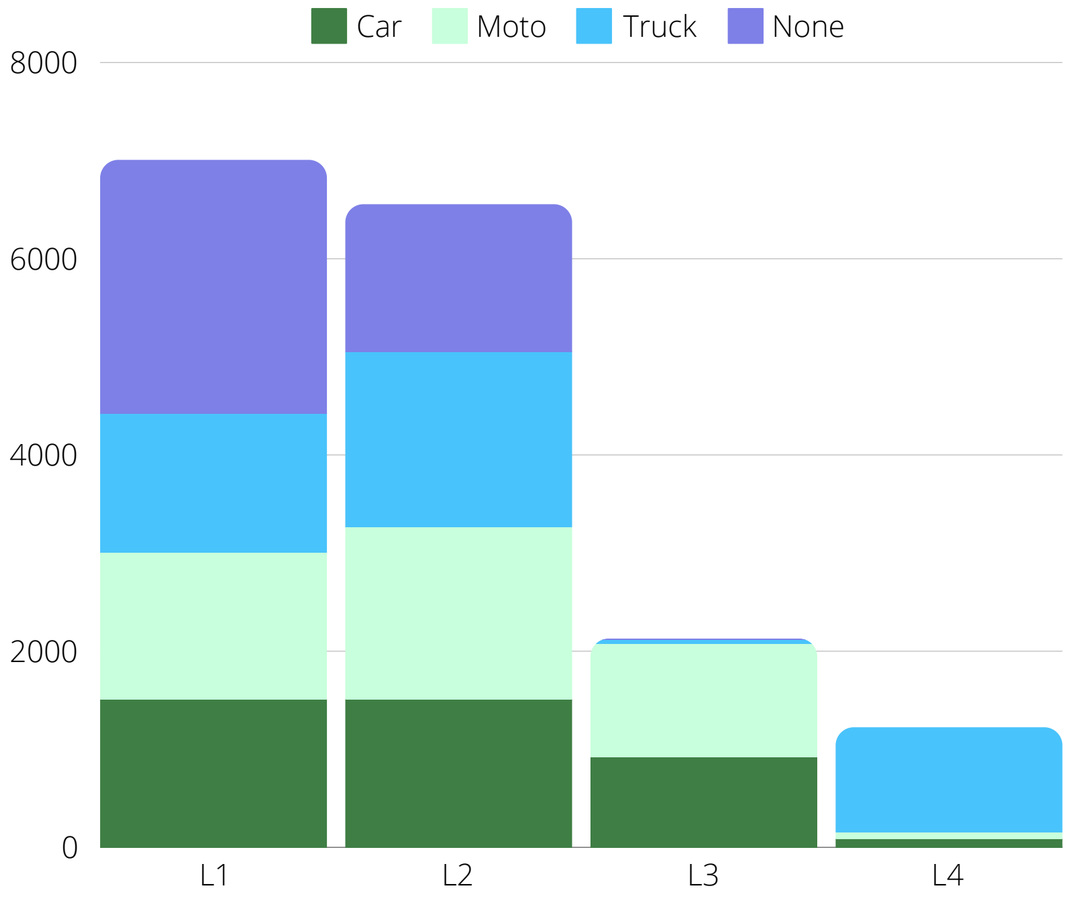
Question 3

What was the error rate relative to Motorcycles? (lane positioning, variable exhaust, etc.)

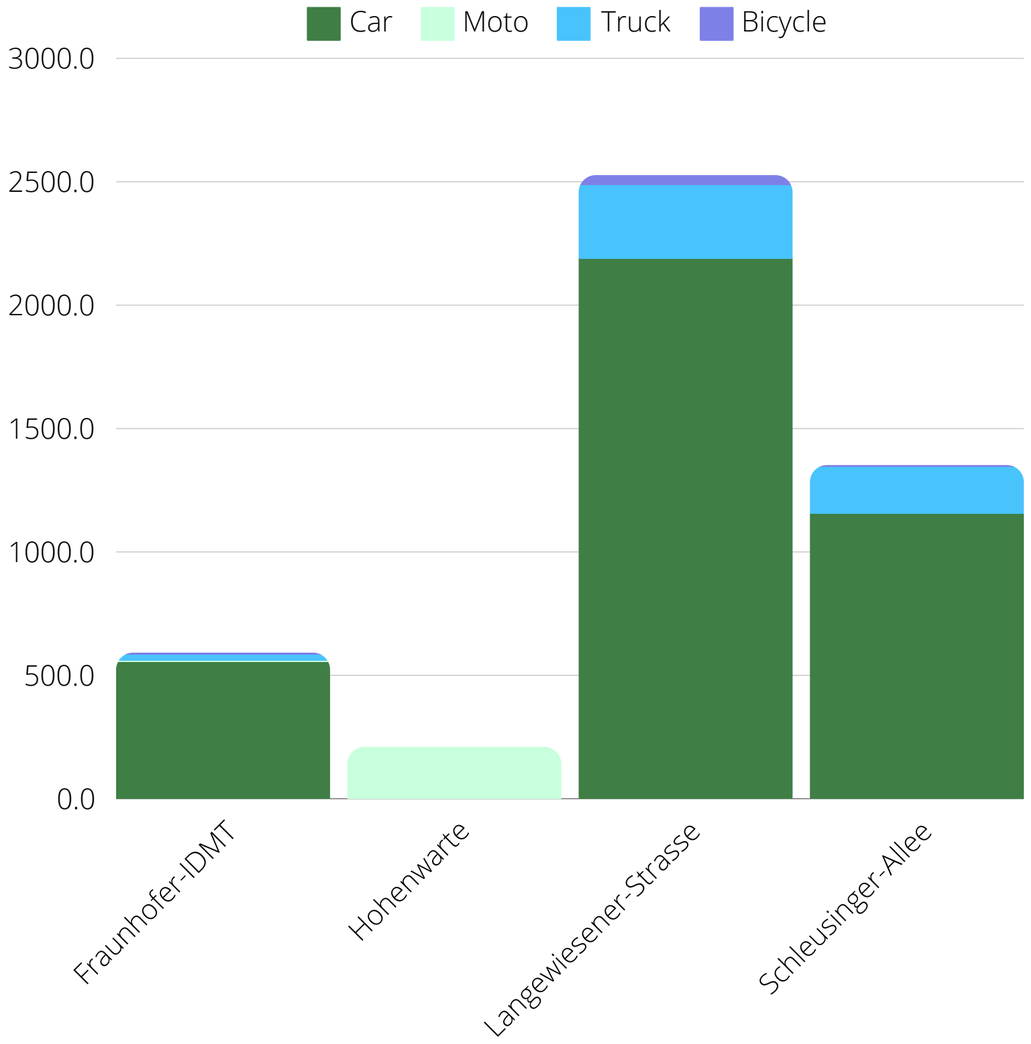
Datasets Comparison

	Germany (IDMT)	India (MVDA)
# of Instances	17,506	16,916
Length (seconds)	2	3
# of Recording Sites	4	4
Vehicle Classes	Car, Motorbike, Bus, Truck, Background	Car, Motorcycle, Truck, Pedestrian
Dry/Wet Included?	Yes	Yes
Recording Hardware	MEMS, sE8	MEMS
Sampling Rate	48	22

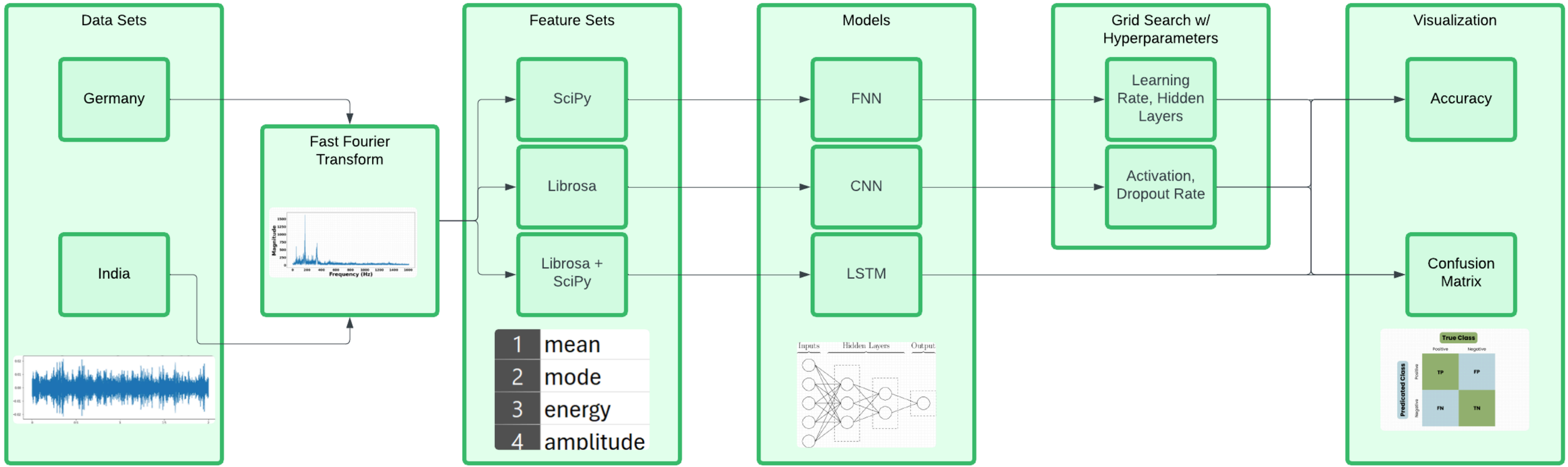
India (MVD)



Germany (IDMT)



Data Pipeline



Feature Sets

1. SciPy

- Mode
- IQR
- Variance
- Entropy
- 1st-4th Moments

2. Librosa (Audio Library)

- MFCCs (analyze frequency and rescale)
- Chroma-stft (how pitches change over time)
- Root Mean Square

3. Librosa + SciPy

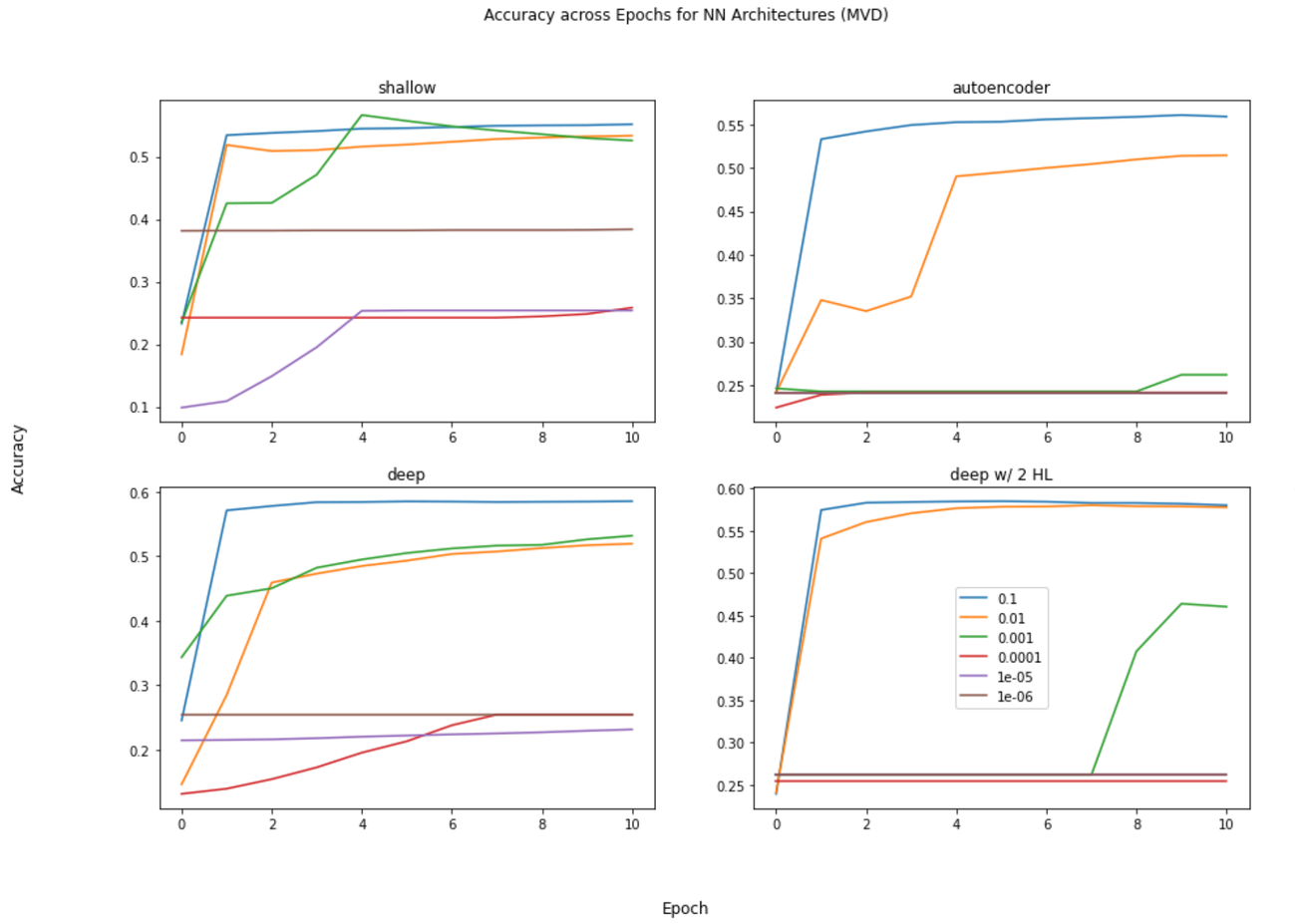
- Spectral Analysis
  - (*Centroid, Roll-off, Bandwidth, Flux*)
- ZCR, MFCCs, Energy, Pitch Freq

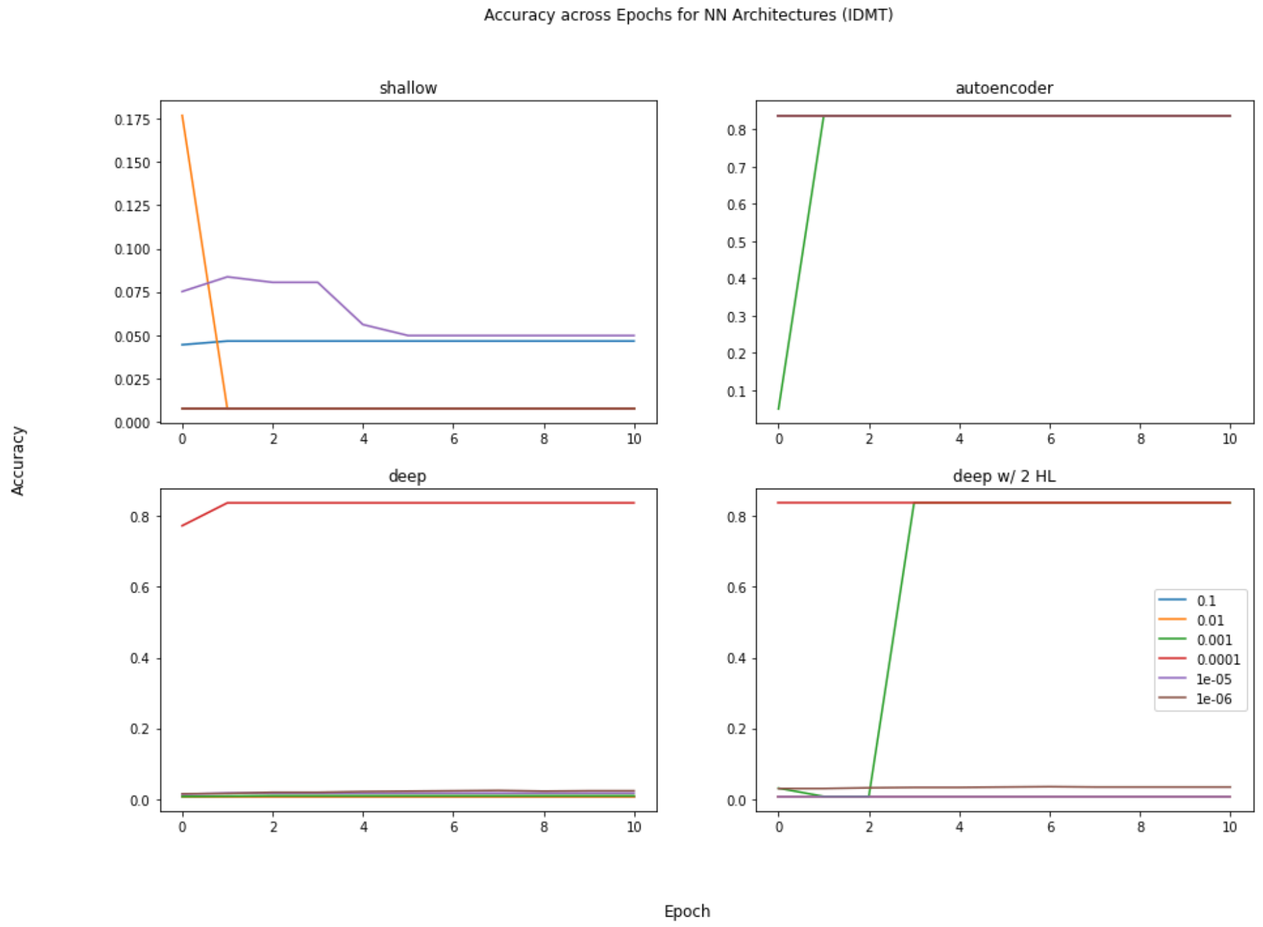
Final

Results

AutoAudio

Linear Network





- Stops improving after 2 epochs
- Upper accuracy bound is:
  - ~60% for India
  - ~80% for Germany

CNN (Only on India)

CNN + LSTM Model:

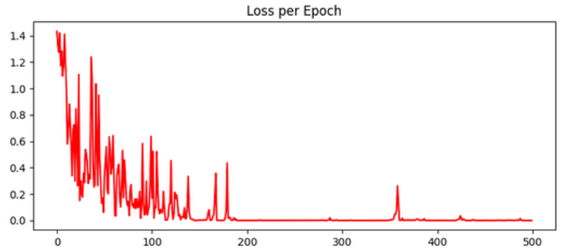
- 3 Convolution Layers
  (why Conv?)
- LSTM
  (time series)
- Dense layers

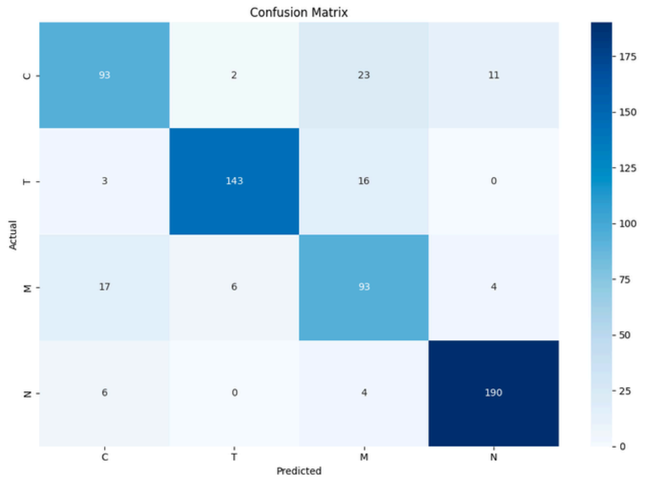
Loss Function:

- Cross Entropy

Optimizer:

- Adam(0.0001)





Overall Accuracy

Accuracy on the Testing Set

For C: 93/129 correctly, 72%.

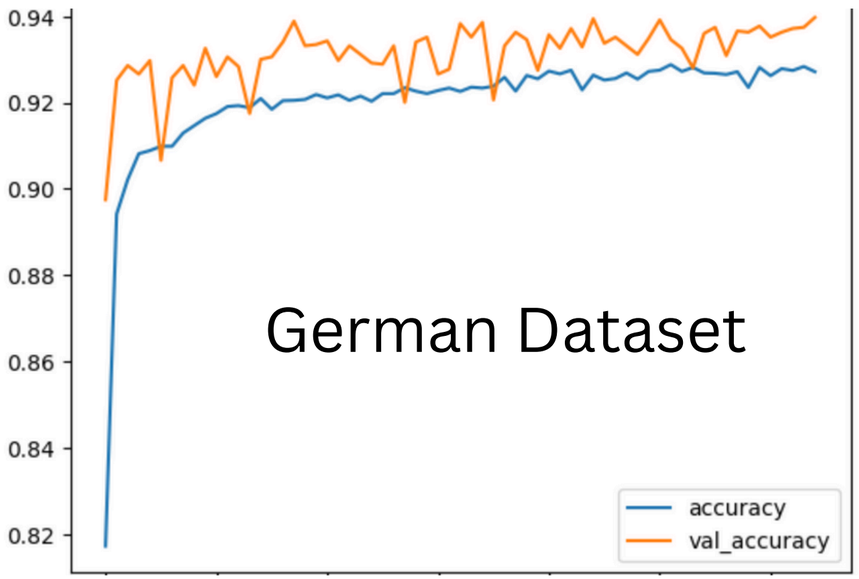
For T: 143/162 correctly, 88%.

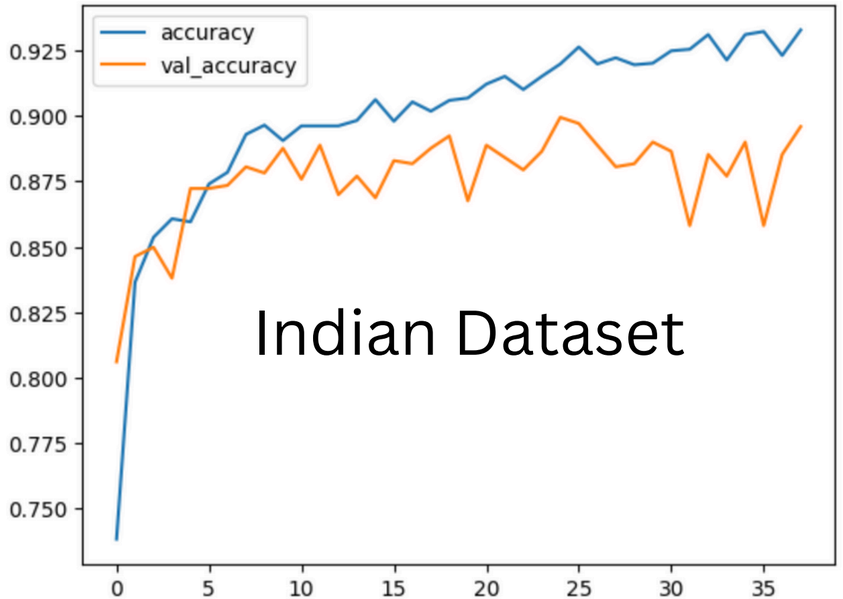
For N: 190/200 correctly, 95%.

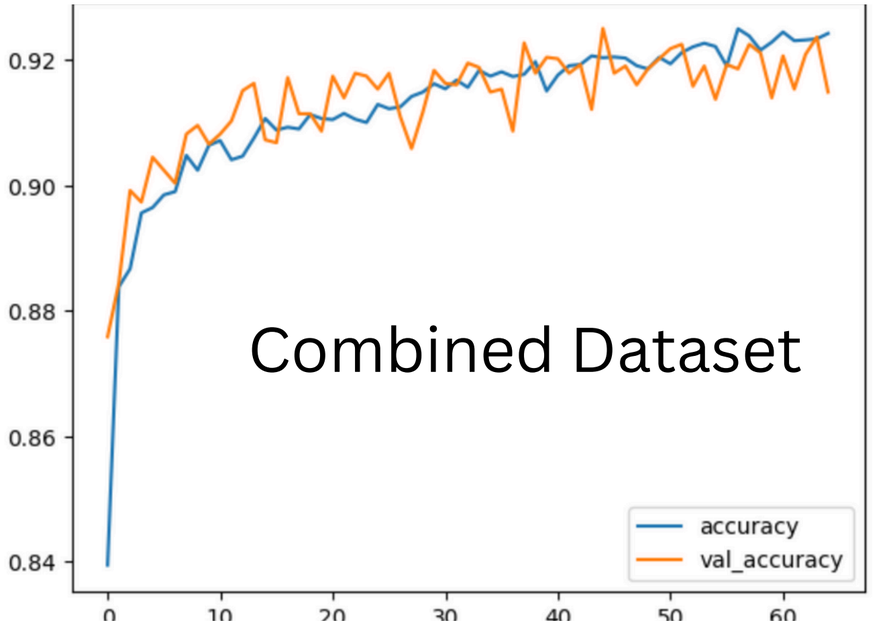
For M: 93/120 correctly, 78%.

Overall: 85%.

LSTM







Summary

- 100 epochs, Early Stopping
- optimal parameters for combined dataset
  - {units = 416, activation = 'relu',
  - number of lstm layers = 3, dropout_rate = 0.25}
- 92.5% accuracy

		LSTM Models		
		German	Indian	Joint
Data	German	93.6%	70.0%	94.3%
	Indian	48.3%	89.8%	90.7%
	Joint	84.1%	84.1%	92.5%

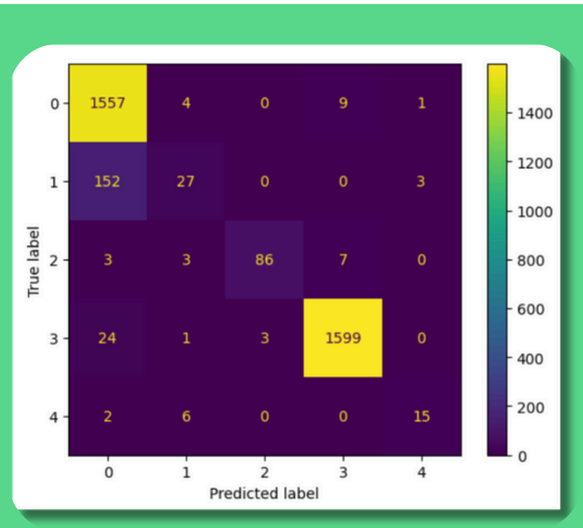
Evaluating on different models

We wanted to see the effect on accuracy rate if we evaluated the different sets on the three models that we had created.

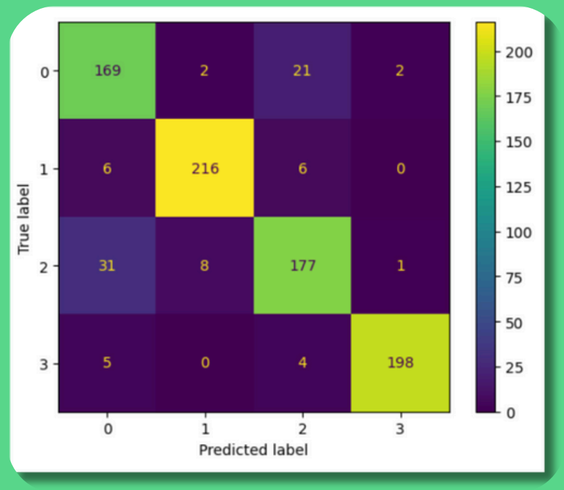
1. LSTM trained on German
    a. Grid-search hyperparams
2. LSTM trained Indian
    a. Keras_tune hyperparams
3. LSTM trained on joined datasets

Error by Vehicle Type (confusion matrix)

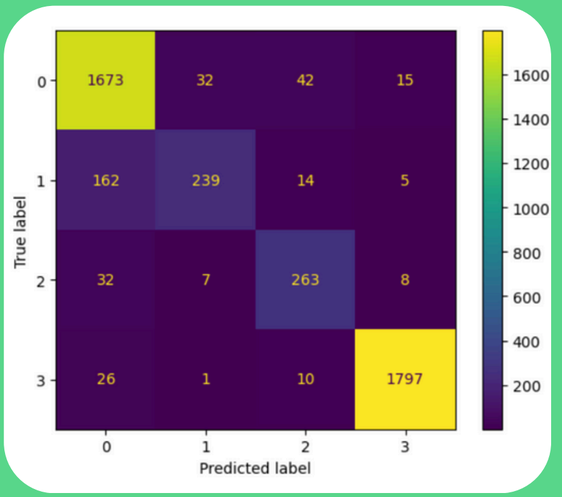
We decided to create a confusion matrix for each of the three models that we discussed in the previous slide. We had initially hypothesized that due to variable lane positioning of motorcycles, we may have a tough time actually accurately predicting it. (Car = 0, Truck = 1, Moto = 2, None/BG = 3, Bus = 4)



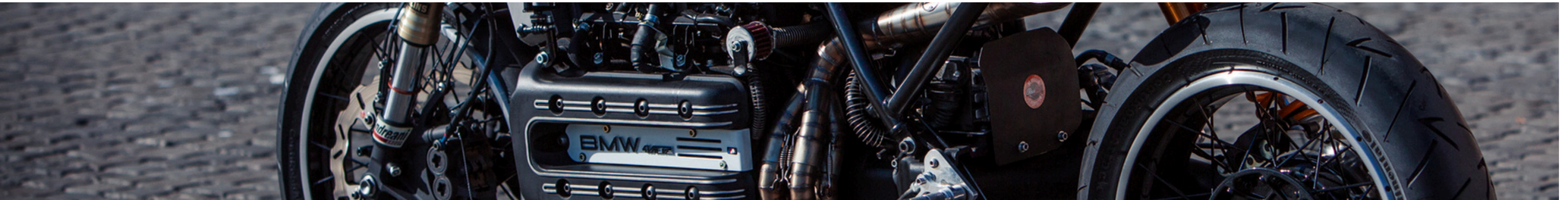
German



Indian



Joint



Future Directions

- Merge analysis
- Predict Speed
- Ensemble model for data fusion

Thanks!