Submitted by:
Amir Hamza Bashir Ahmed Shaikh

Title of Project:

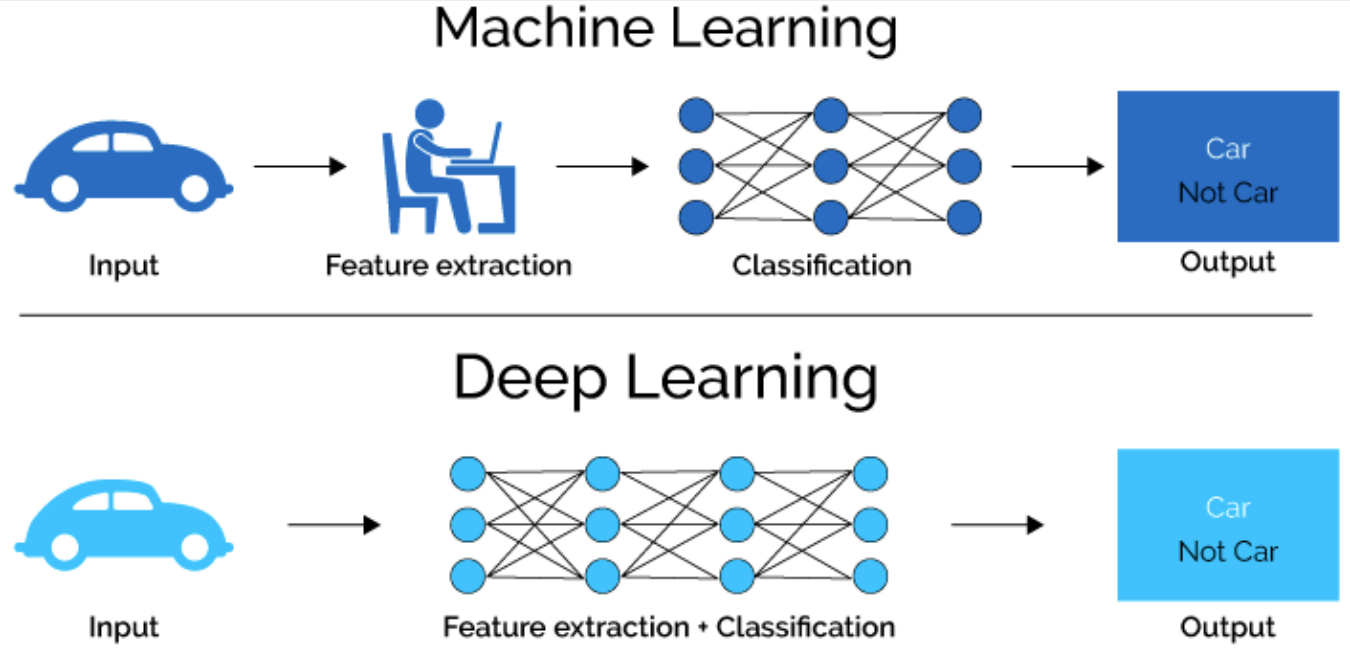
Binary Sound classification of Emergency and Non Emergency vehicles Using Deep Learning Models Through Time Domain and Spectrogram

Abstract

Recent advances in **Deep Learning models** for tackling difficult classification tasks have sparked increased interest in sound classification strategies. Since all aspects of human life depend on sound. Sound is a key component in the development of AI systems for these fields, from critical monitoring to personal safety.

The aim is to perform audio classification that can classify ambulance and non-emergency vehicles using two of the important deep learning models. Based on the nature of the audio and complexities associated with it, here deep learning algorithms such **as Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM)** are used to build the models. The preliminary results shows that LSTM performs much better than CNN **in time domain**. Where as in **the CNN perform better than LSTM in spectrogram domain**. Both the models were validated with standard datasets and results very are encouraging.

INTRODUCTION



Research Gaps

All papers used Spectrogram as input? Why

CNN is commonly used? WHY?

Some authors have made Deep learning model(s) **VS** NON-Deep learning model(s)

Is it possible that DL performs better without spectrogram?

Is fine tuning of model is proportional to validation accuracy?

Target dish : Paneer Pulao
Two main Ingredients: Paneer + Rice
Each of 4 pots have paneer of different brands: Amul, Mother Dairy , Aashirvad ,Nandini
Each of 4 pots have Rice of different brands: DAAWAT,India Gate,Delhi Rice,Kohinoor

Does the pulao taste the same in all three pots?
Does it take the same amount of time to cook regardless of the shape of the pot?
Does the name of the dish change as different brands of paneer and rice are used?




Main type of classification: Sound Classification
Two main Ingredients: ML model1+ ML model2
Both models have trained and tested in various applications: Microphone classification, Environmental classification, event classification, recitation(maqam) classification

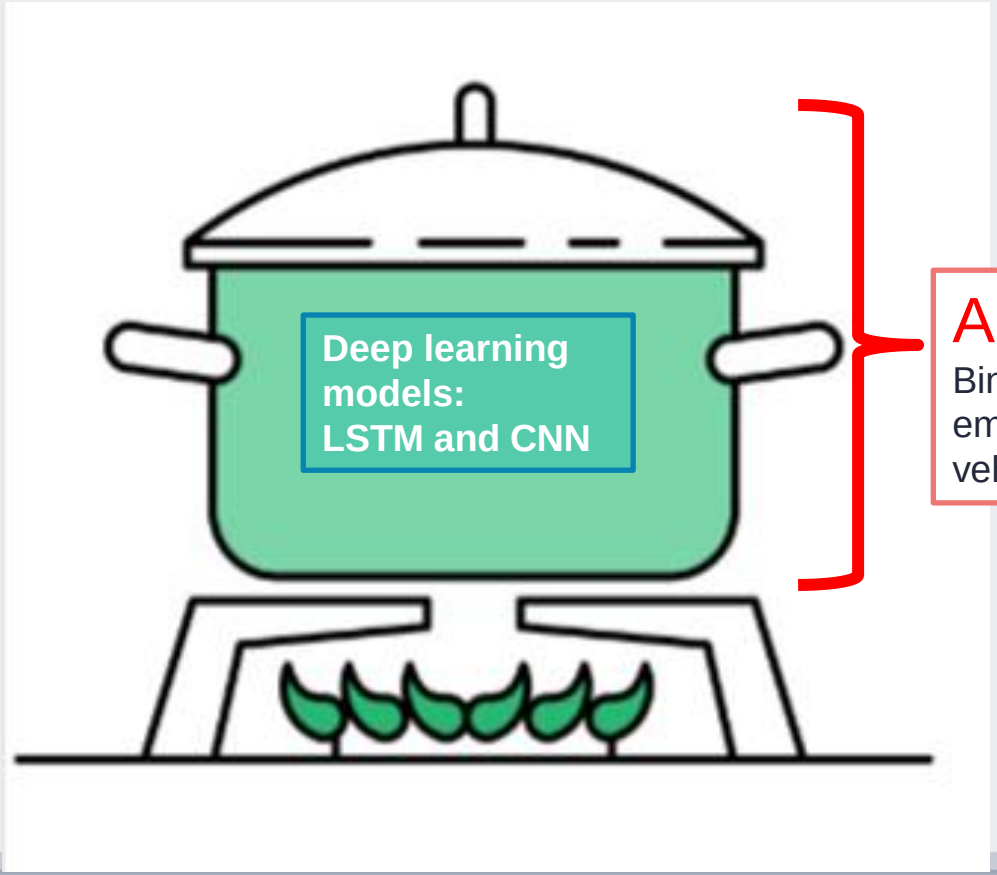
Does the outcome Sound classification remains same ?
Does it take the same amount of time to train and test the ML models?
It is possible to change name ie. Sound classification to Image classification irrespective of ML models and applications.

Application:
Microphone classification

Application:
Environmental classification

Application:
Event classification

Application:
Recitation classification



Application:

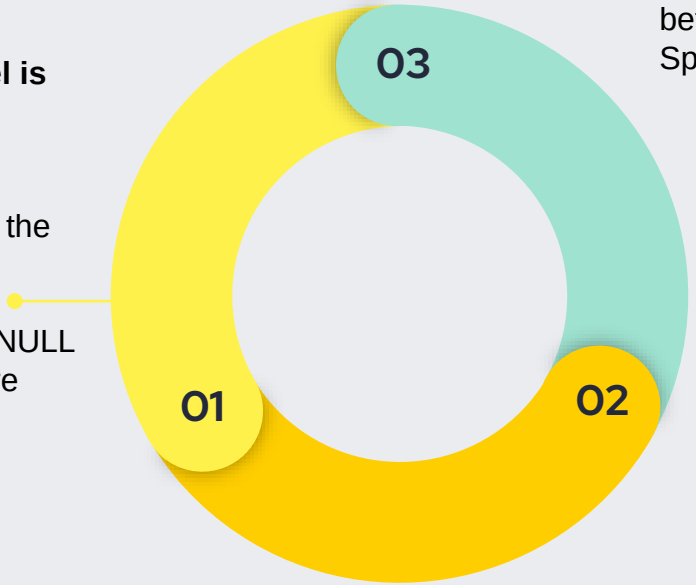
Binary Sound Classification of emergency and Non-emergency vehicles.

(OBJECTIVES) Motivations

NULL Hypothesis: Whether the deep learning model work better in taking the audio source in time domain.

ALTERNATE Hypothesis: If NULL Hypothesis fails, Then model works better in in taking the audio source in Spectrogram.

Which deep learning model is better when audio input is given in Time Domain .
NULL Hypothesis: In Time domain, IS LSTM dominates the CNN with respect to model accuracy.
ALTERNATE Hypothesis: If NULL Hypothesis fails, CNN is more superior to LSTM model.



Which deep learning model is better when audio input as Spectrogram .
NULL Hypothesis: In Spectrogram, IS LSTM dominates the CNN with respect to model accuracy
ALTERNATE Hypothesis: If NULL Hypothesis fails, CNN is more superior to LSTM model.

PROBLEM STATEMENT

THE HINDU

MENU

HOME NEWS OPINION BUSINESS SPORT ENTERTAINMENT CROSSWORD+ SCIENCE

CITIES BENGALURU CHENNAI COIMBATORE DELHI HYDERABAD KOCHI KOLK

NEWS > CITIES > MUMBAI

MUMBAI

Helping patients get to hospital on time

SHARE ARTICLE PRINT A A A

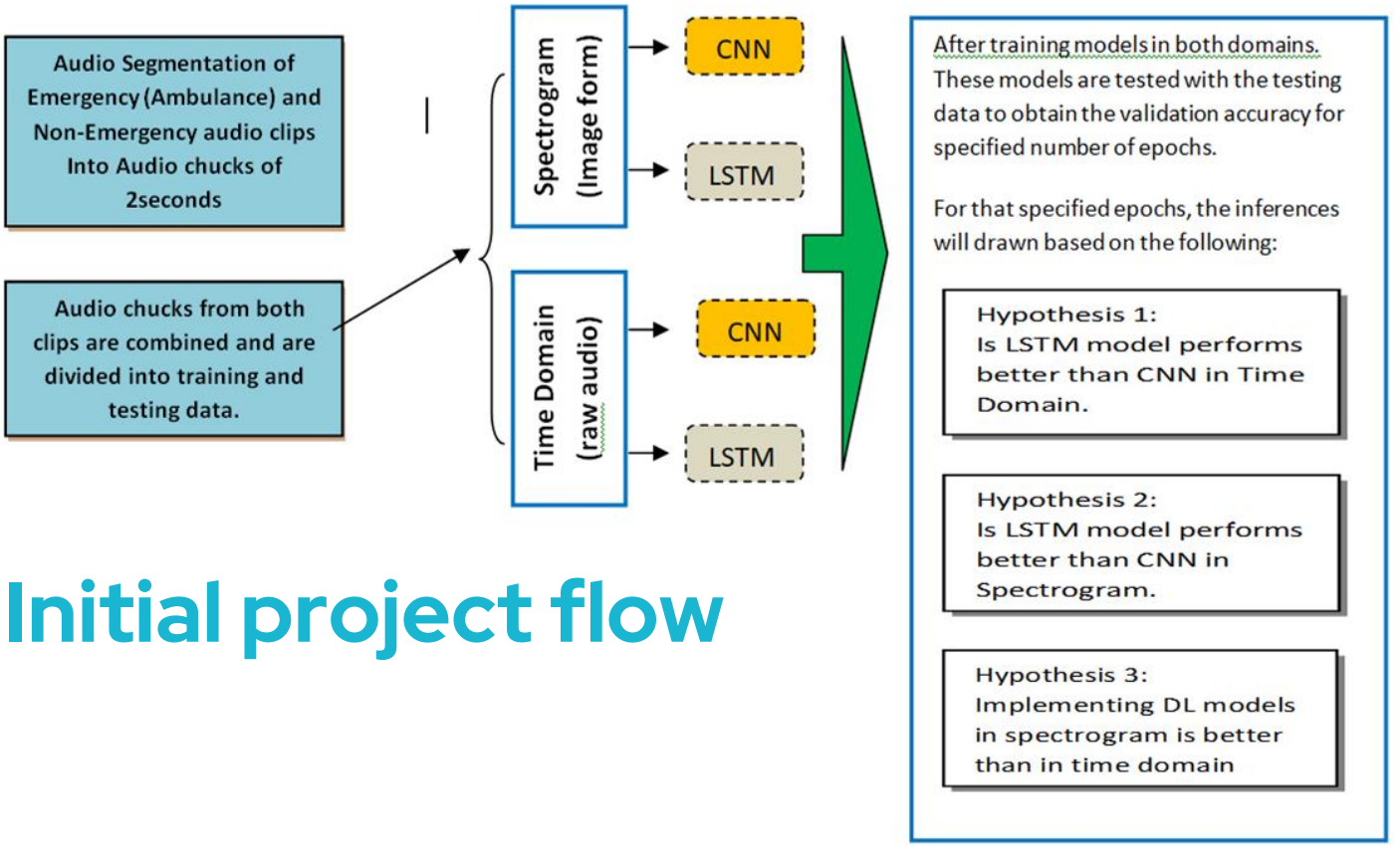
A foundation is creating awareness among motorists to give ambulances the right of way on the road

According to the National Crime Records Bureau, nearly 24,012 people die each day due to a delay in getting medical assistance. These patients have suffered heart attacks, brain haemorrhage, suicide attempts, accidents and strokes. The Foundation found — from national crime, birth and death records in January 2014 — that heart attacks, at 19% of the total, are the leading cause of death, while brain haemorrhage is fourth on the list. Early stabilisation of a brain haemorrhage patient is critical in saving his life.

Accidents are 10th on the list — nearly 4,40,042 cases are reported across the country each year, of which 1,39,091 people lose their life. The first hour after the incident, or the Golden Hour, is critical. Many accident victims wait for help at the site, and a delay costs them their life. "The reasons could range from ambulances stuck in traffic to the fire brigade not being able to reach the site on time due to traffic jams," says Dr. Savla, who has been distributing pamphlets

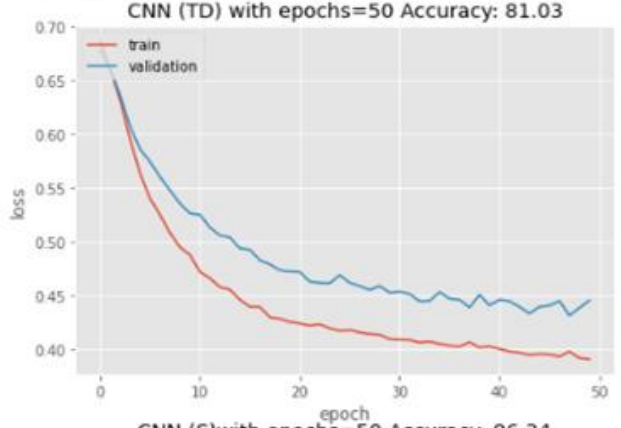
According to the National Crime Records Bureau, nearly 24 thousand people die each day due to a delay in getting medical assistance. Many accident victims wait for help at the site, and a delay costs them their lives. The reasons could range from ambulances stuck in traffic to the fire brigade not being able to reach the site on time due to traffic jams.

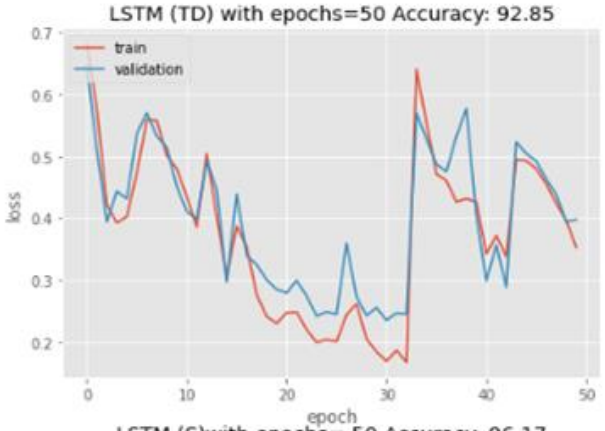
The solution to the above problem is to create a system that can automatically detect the emergency vehicles before they reach the traffic signals and accordingly the traffic signals changes.

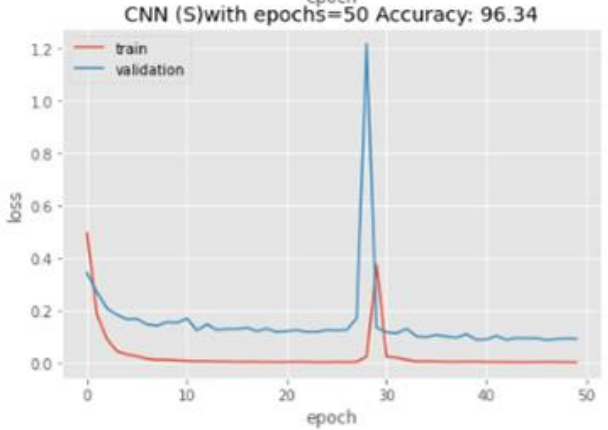


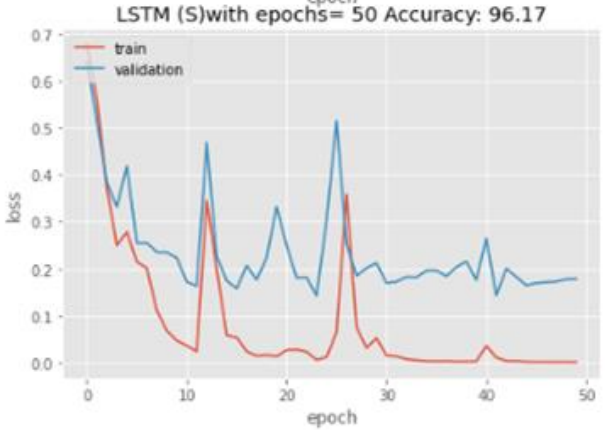
Initial project flow

Text(0.5, 1.0, 'LSTM (S)with epochs= 50 Accuracy: 96.17')



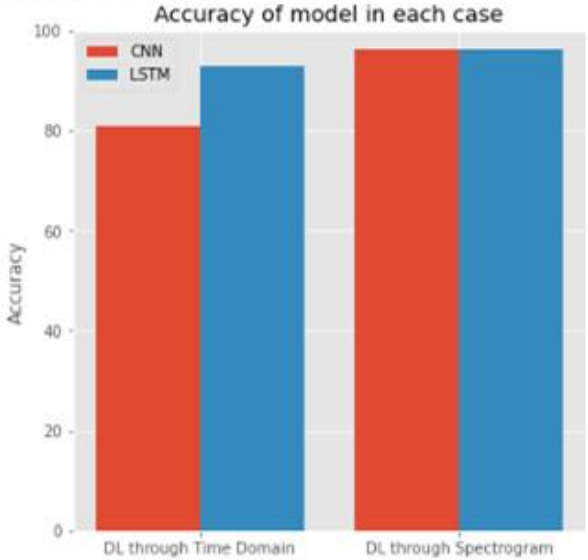




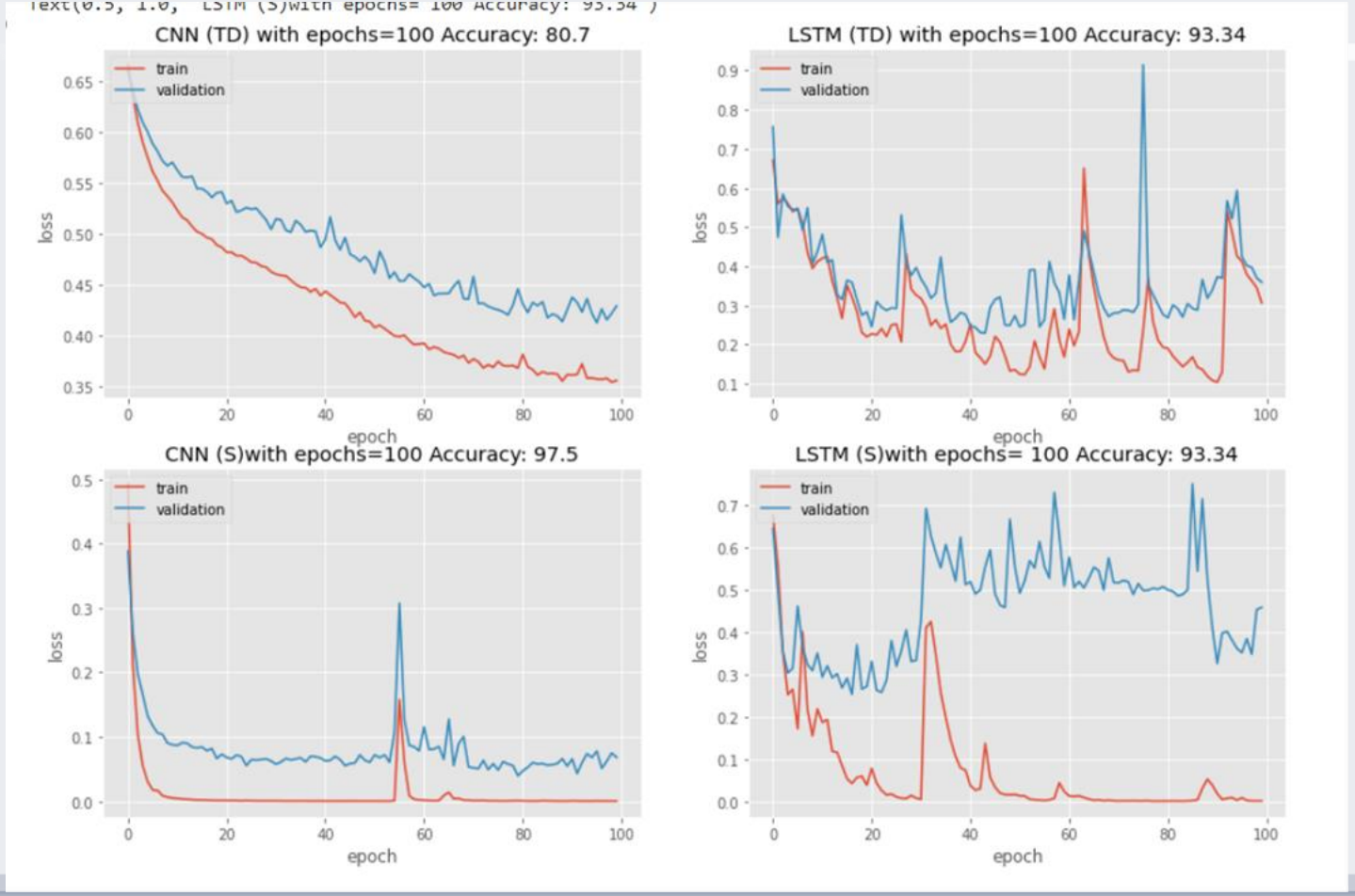


```
st1,st2,st3=hypothesis_testing(CNN_TD_acc,LSTM_TD_acc,CNN_S_acc,LSTM_S_acc)
# st1,st2,st3=hypothesis_testing(50.22,50,34.33,80.45)
```

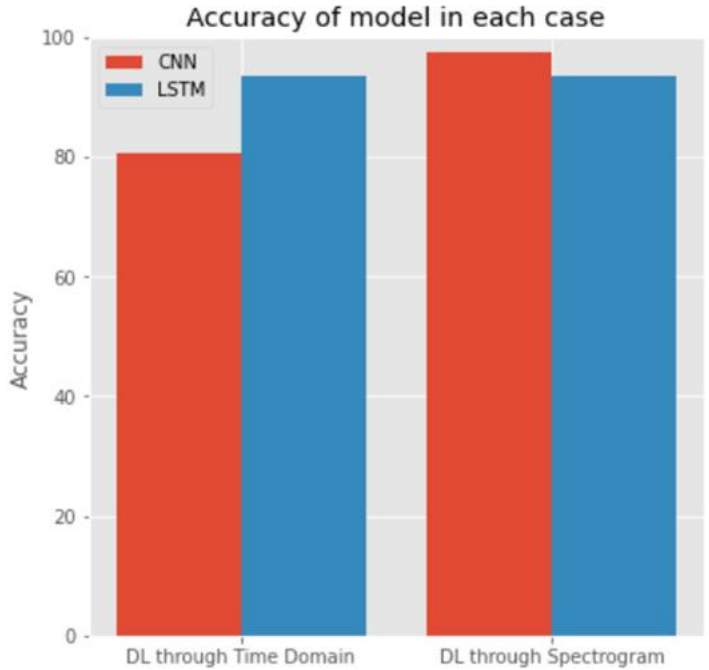
```
DATA:
{'Time Domain': (['CNN', 81.03], ['LSTM', 92.85]), 'Spectrogram': (['CNN', 96.34], ['LSTM', 96.17])}
THE Accuracy Graph
```




```
INFERENCES FROM THE ABOVE GRAPH
The model LSTM performs well in Time Domain as compared to CNN
Both models CNN & LSTM have equal performance in Spectrogram
 OVERAll:Both models  (LSTM and CNN) performs better in Spectrogram
```

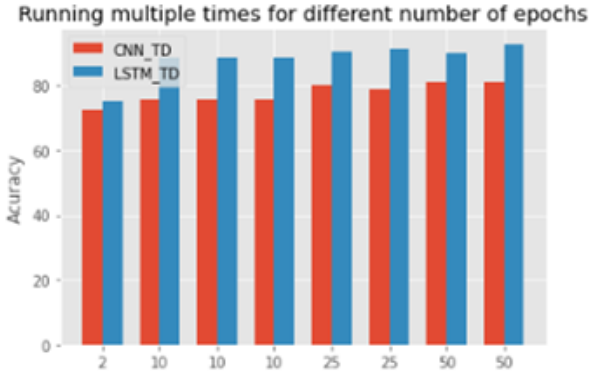

CNN (TD) with epochs=100 Accuracy: 80.7
LSTM (TD) with epochs=100 Accuracy: 93.34
CNN (S)with epochs=100 Accuracy: 97.5
LSTM (S)with epochs= 100 Accuracy: 93.34
train
validation
loss
epoch

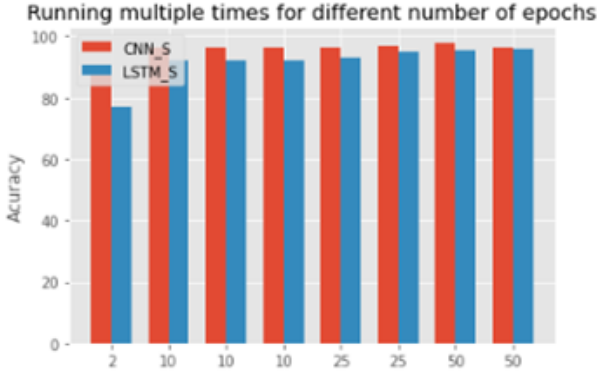
DATA:
{'Time Domain': (['CNN', 80.7], ['LSTM', 93.34]), 'Spectrogram': (['CNN', 97.5], ['LSTM', 93.34])}
THE Accuracy Graph

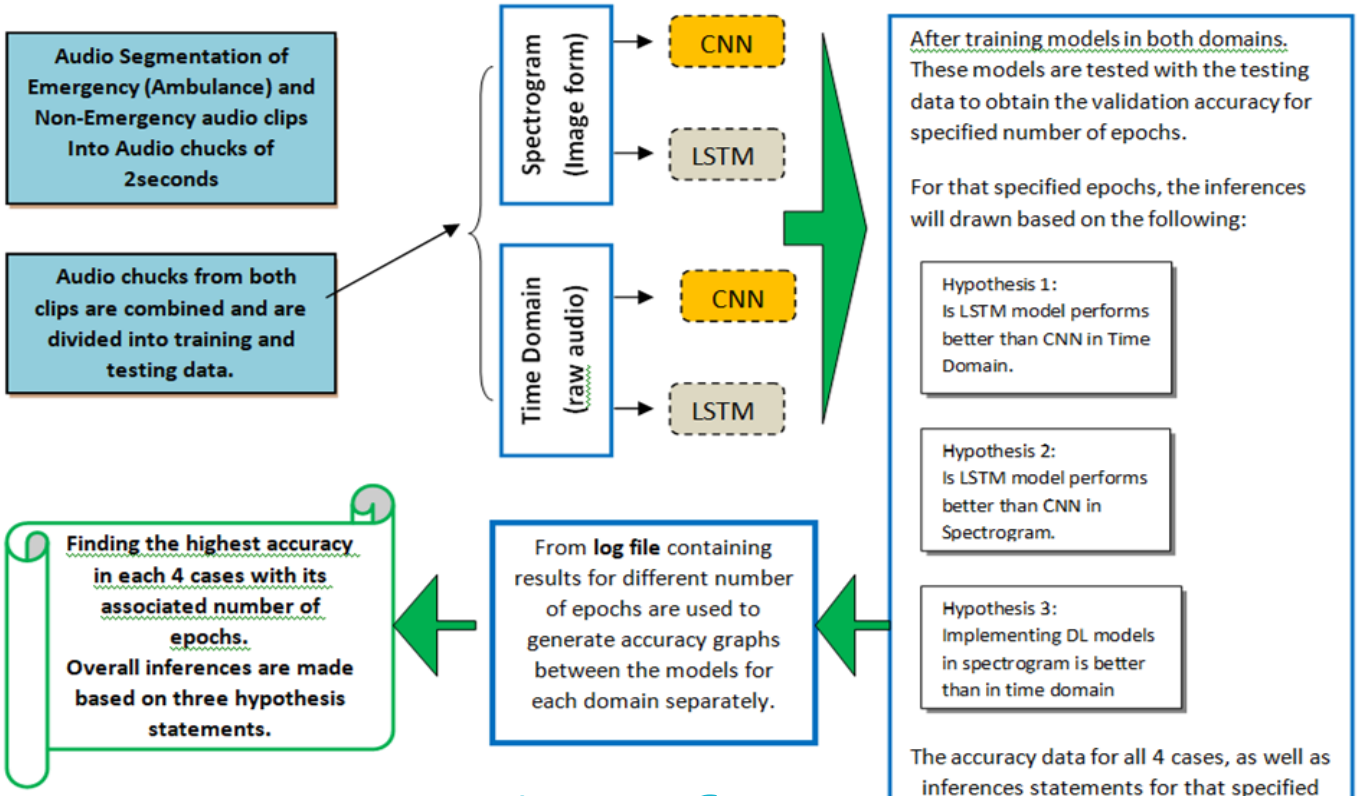


INFERENCES FROM THE ABOVE GRAPH
The model LSTM performs well in Time Domain as compared to CNN
The model CNN performs well in Spectrogram as compared to LSTM
OVERAll: one model performs better in time domain and other in spectrogram

done

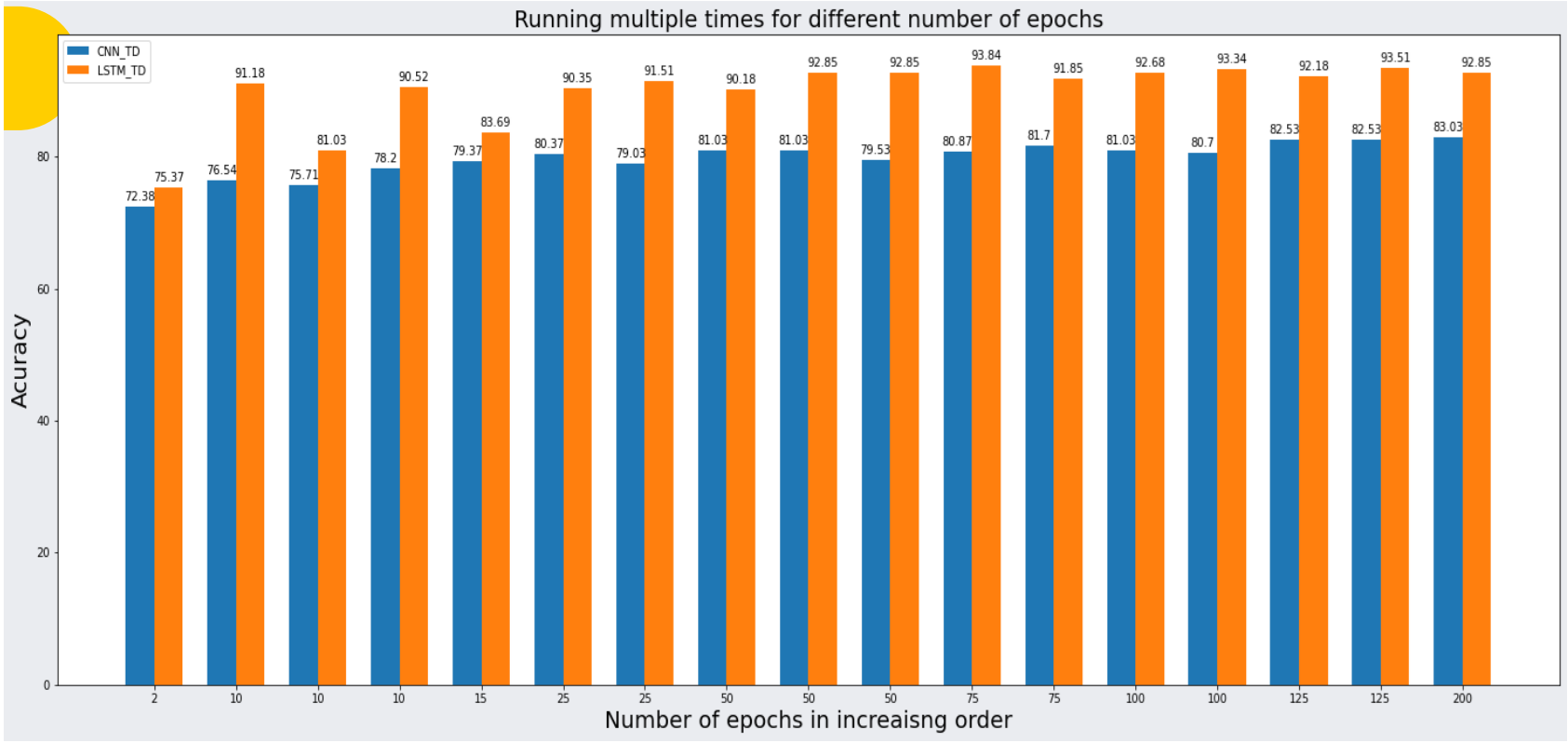

Running multiple times for different number of epochs
CNN_TD
LSTM_TD
Acuracy
0
20
40
60
80
2
10
10
10
25
25
50
50

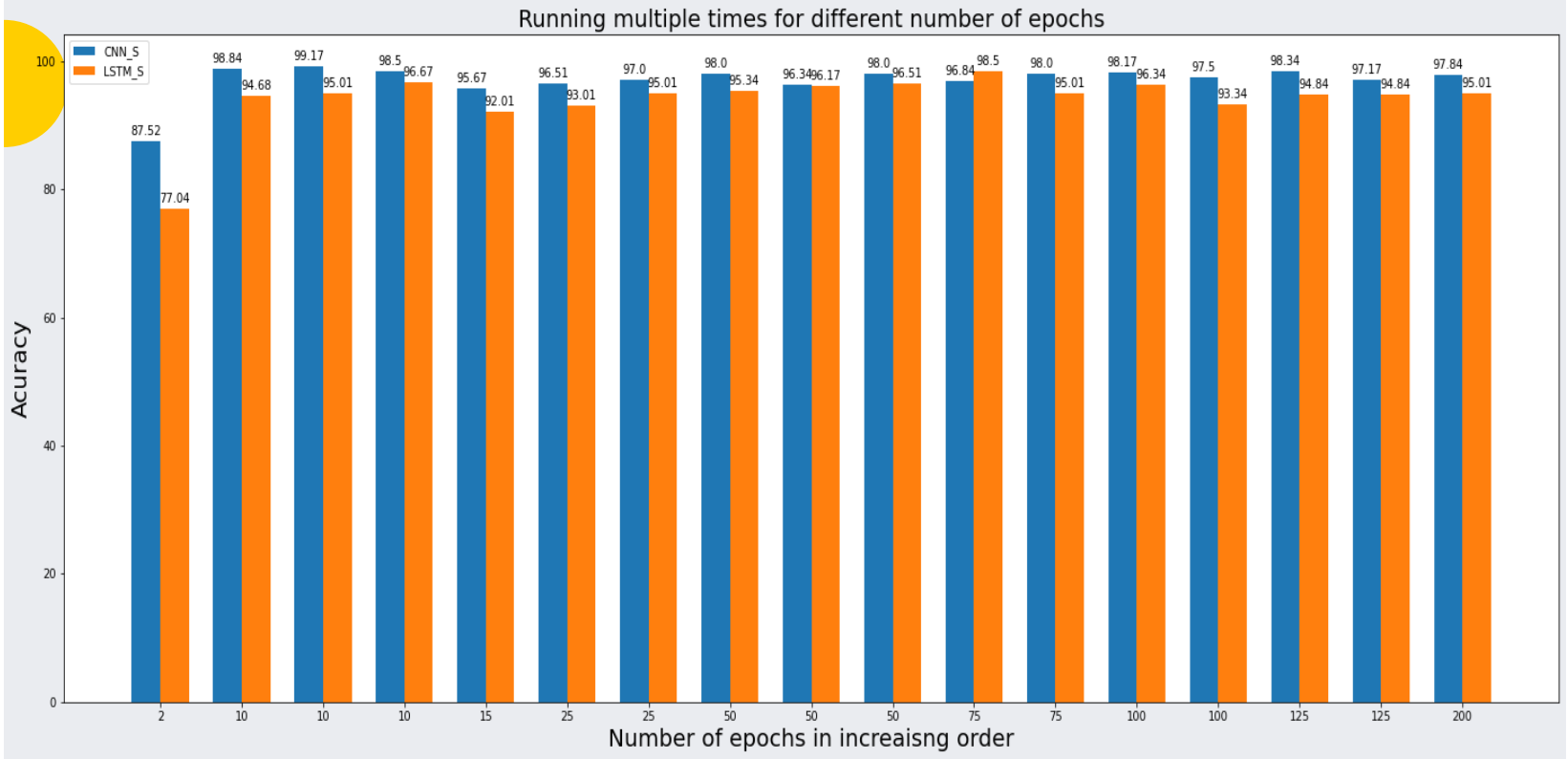

Running multiple times for different number of epochs
CNN_S
LSTM_S
Acuracy
0
20
40
60
80
100
2
10
10
10
25
25
50
50



Updated project flow

2|72.38|75.37|87.52|77.04|Both models CNN & LSTM have equal performance in Time Domain|The model LSTM performs well in Spectrogram as compared to CNN|OVERAll: one model performs better in time domain and other in spectrogram
10|75.54|88.69|96.51|92.18|Both models CNN & LSTM have equal performance in Time Domain|The model LSTM performs well in Spectrogram as compared to CNN|OVERAll: one model performs better in time domain and other in spectrogram
10|75.54|88.69|96.51|92.18|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
10|75.54|88.69|96.51|92.18|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
25|80.37|90.35|96.51|93.01|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
25|79.03|91.51|97.0|95.01|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
50|81.03|90.18|98.0|95.34|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
50|81.03|92.85|96.34|96.17|The model LSTM performs well in Time Domain as compared to CNN|Both models CNN & LSTM have equal performance in Spectrogram| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
100|81.03|92.68|98.17|96.34|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
75|80.87|93.84|96.84|98.5|The model LSTM performs well in Time Domain as compared to CNN|The model LSTM performs well in Spectrogram as compared to CNN| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
125|82.53|92.18|98.34|94.84|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram
125|82.53|93.51|97.17|94.84|The model LSTM performs well in Time Domain as compared to CNN|The model CNN performs well in Spectrogram as compared to LSTM| OVERAll:Both models (LSTM and CNN) performs better in Spectrogram


Running multiple times for different number of epochs
CNN_TD
LSTM_TD
Acuracy
Number of epochs in increaisng order
0
20
40
60
80
2
10
10
10
15
25
25
50
50
50
75
75
100
100
125
125
200
72.38
75.37
76.54
91.18
75.71
81.03
78.2
90.52
79.37
83.69
80.37
90.35
79.03
91.51
81.03
90.18
81.03
92.85
79.53
92.85
80.87
93.84
81.7
91.85
81.03
92.68
80.7
93.34
82.53
92.18
82.53
93.51
83.03
92.85


Running multiple times for different number of epochs
CNN_S
LSTM_S
Acuracy
0
20
40
60
80
100
87.52
77.04
98.84
94.68
99.17
95.01
98.5
96.67
95.67
92.01
96.51
93.01
97.0
95.01
98.0
95.34
96.34
96.17
98.0
96.51
96.84
98.5
98.0
95.01
98.17
96.34
97.5
93.34
98.34
94.84
97.17
94.84
97.84
95.01
2
10
10
10
15
25
25
50
50
50
75
75
100
100
125
125
200
Number of epochs in increaisng order

Results

Hypothesis 1: Deep Learning Models performs well in Spectrogram domain than time domain after observing the accuracy graphs .

Hypothesis 2: LSTM performs well as compared to CNN model in time Domain as LSTM in almost all of the epochs

Hypothesis 3: It had very tough competition between CNN and LSTM model in Spectrogram. CNN leads in majority of epochs. One important point to note that CNN performed well within short number epochs that is, when epochs=10

Future Scope

As this work focuses on number of epochs for deep learning models. The further work can also be done by varying the layer configurations, dropout values, sampling rate, number of samples, batch size, and even sample size in order to have a more detailed comparison between the two models in their respective domains.

Thank You !!