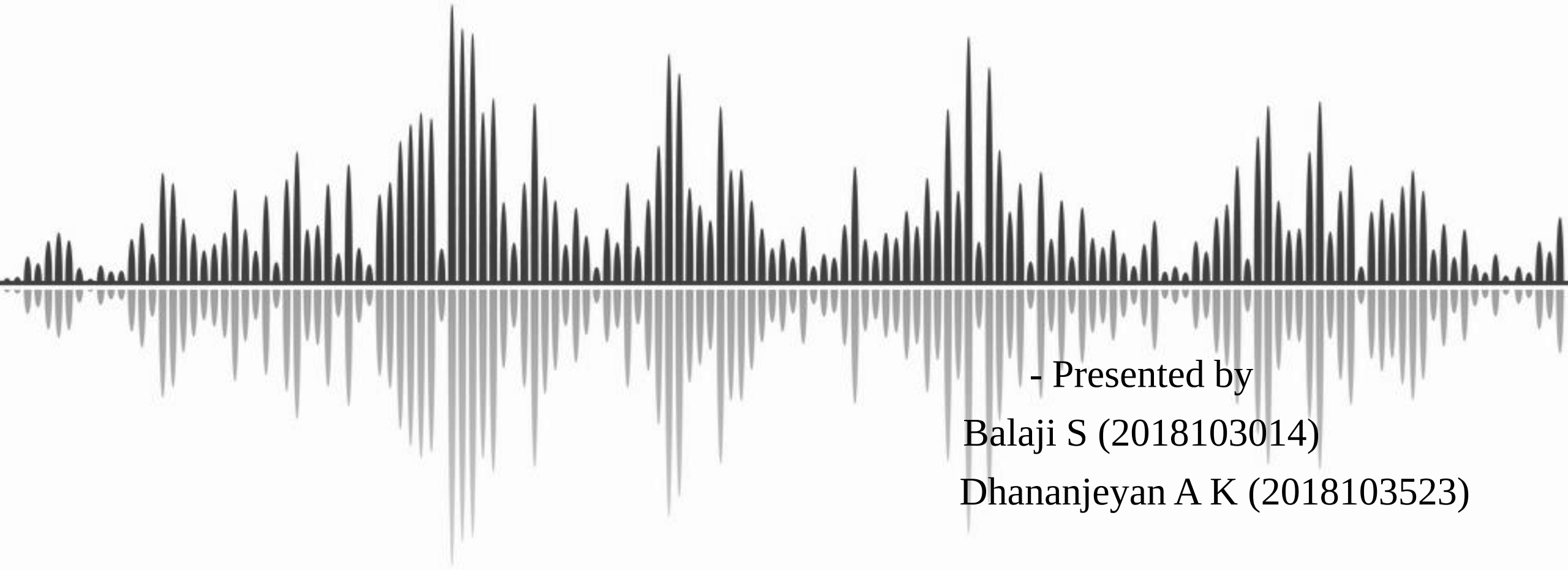
# CS6030 - NATURAL LANGUAGE PROCESSING

# SPEECH EMOTION RECOGNITION



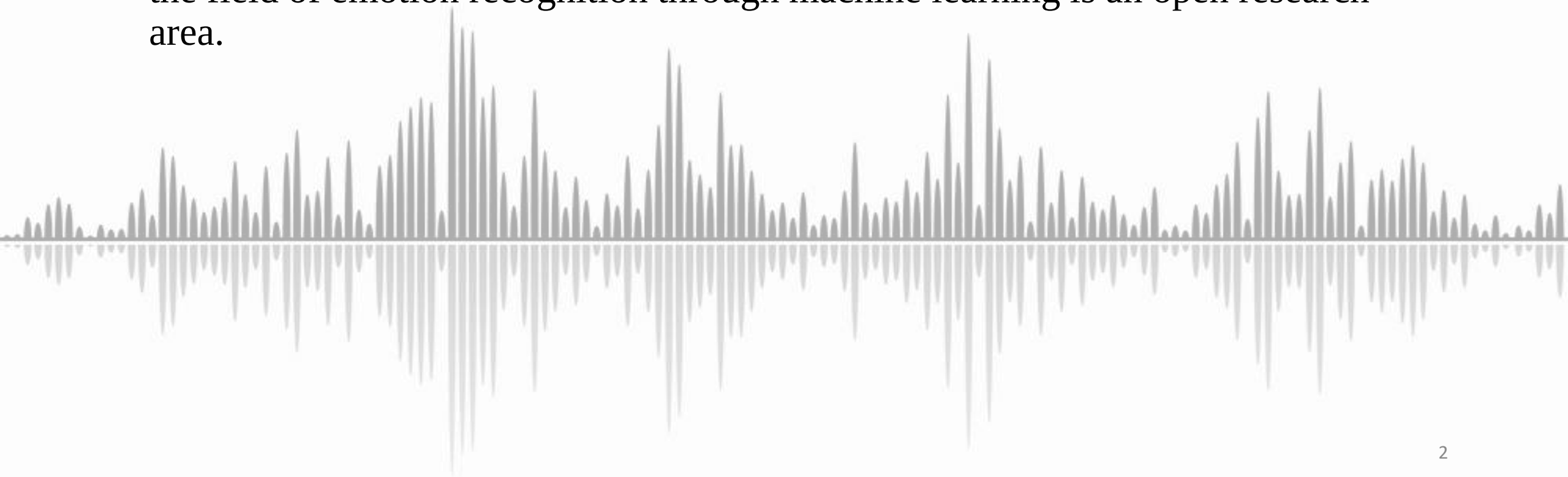
- Presented by

Balaji S (2018103014)

Dhananjeyan A K (2018103523)

# INTRODUCTION:

- The human voice is very versatile and carries a multitude of emotions. Emotion in speech carries extra insight about human actions. Through further analysis, we can better understand the motives of people, whether they are unhappy customers or cheering fans. Humans are easily able to determine the emotion of a speaker, but the field of emotion recognition through machine learning is an open research area.



## OBJECTIVES:

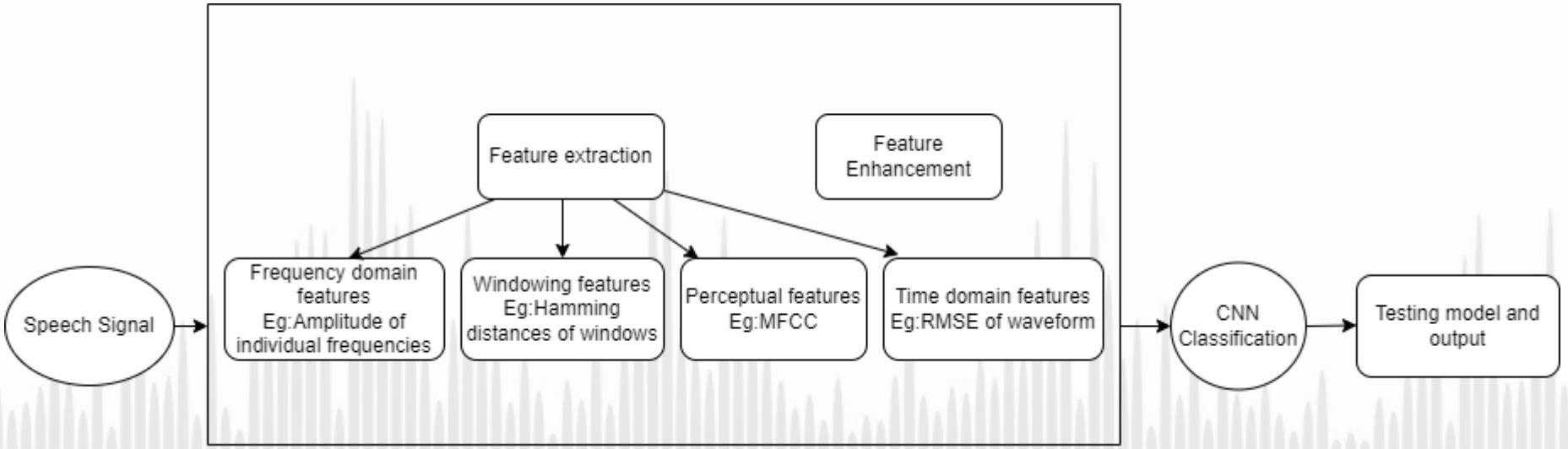
- The main objective is to enable computers to understand the emotional states expressed by the human subjects, so that personalized responses can be delivered accordingly in the cases of voice assistance and recommendation systems.

## DATASET:

- Input signals are from Ryerson Audio Visual Database of Emotional speech and song (RAVDESS) against classifier models. (24 actors).

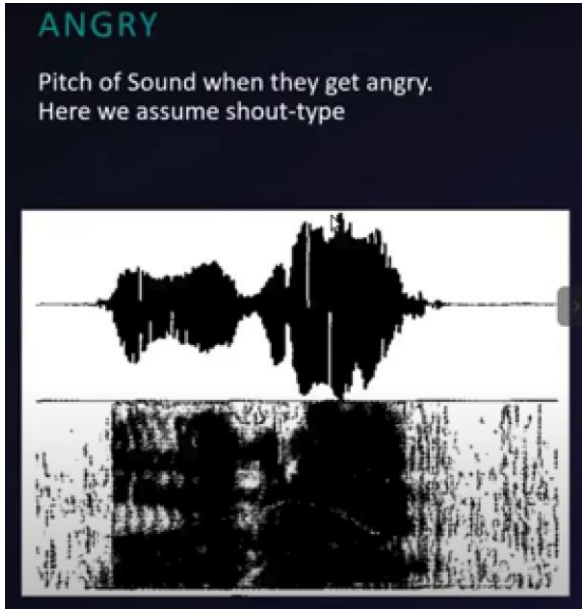
- Surrey Audio Visual Expressed Emotion (SAVEE) dataset

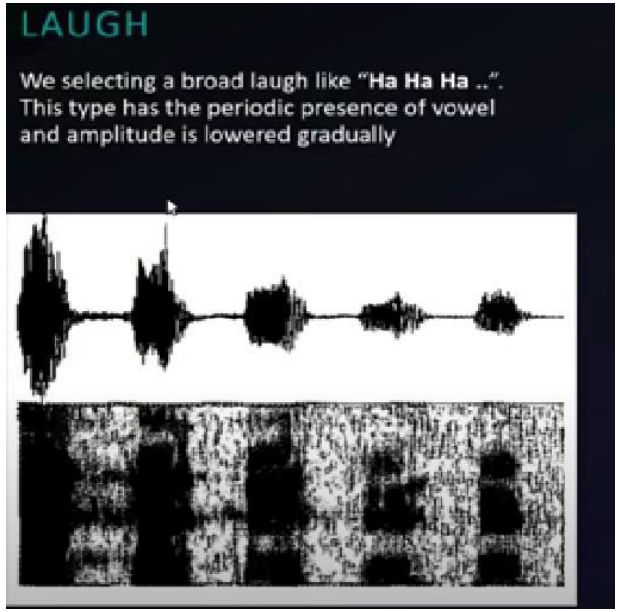
# METHODOLOGY DIAGRAM:




- Converting audio files into Machine understandable format.(Data preparation) using Sampling and frequency domain plotting.
- Data preprocessing – Extracting features from audio using Librosa and preparing for CNN –Features extracted using algorithms of MFCC , MEDC
- Features are then extracted using windowing techniques.

# METHODOLOGY CONTD.





- SER Model Creation (CNN)
- Model Training and Evaluation
- Test set prediction using loaded model
- Testing with audio data

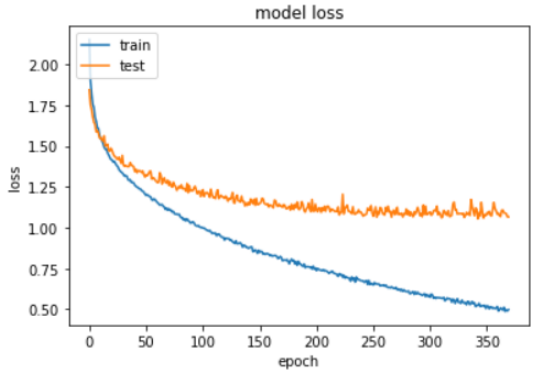
# PERFORMANCE:

- Val Loss and accuracy
- Accuracy – **61.7%**

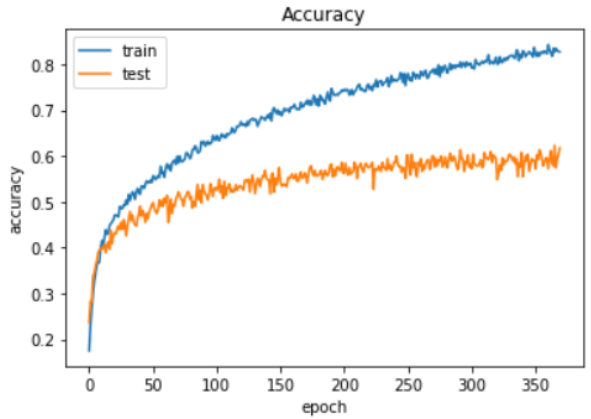
```
Epoch 367/370
128/128 [==============================] - 6s 47ms/step - loss: 0.4691 - accuracy: 0.8458 - val_loss: 1.0879 - val_accuracy:
0.5742
Epoch 368/370
128/128 [==============================] - 6s 47ms/step - loss: 0.5058 - accuracy: 0.8251 - val_loss: 1.0851 - val_accuracy:
0.5938
Epoch 369/370
128/128 [==============================] - 6s 48ms/step - loss: 0.4935 - accuracy: 0.8179 - val_loss: 1.0662 - val_accuracy:
0.6035
Epoch 370/370
128/128 [==============================] - 6s 47ms/step - loss: 0.5056 - accuracy: 0.8244 - val_loss: 1.0660 - val_accuracy:
0.6172
```

**Loss Vs Iterations**

```
plt.plot(cnnhistory.history['loss'])
plt.plot(cnnhistory.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()
```




```
plt.plot(cnnhistory.history['accuracy'])
plt.plot(cnnhistory.history['val_accuracy'])
plt.title('Accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()
```




| | actualvalues | predictedvalues |
|---|---|---|
| 130 | 1 | 1 |
| 131 | 5 | 0 |
| 132 | 1 | 1 |
| 133 | 3 | 3 |
| 134 | 4 | 4 |
| 135 | 0 | 0 |
| 136 | 3 | 3 |
| 137 | 0 | 0 |
| 138 | 1 | 1 |
| 139 | 6 | 6 |