Emotion Detection Via Tone of Voice

(USING DEEP-LEARNING)

PRESENTED BY:
AHAD ALMUTAIRI
AMJAD ALTHUNYAN

OVERVIEW

01 TOOLS

02 INTRO

03 WORK FLOW

04 EDA& PREPROCESSING

05 DATA AUGMENTATION

06 CLASSIFICATION MODELS

07 NN MODELS

08 TRANSFER LEARNING

09 PREDICTION

TOOLS

Keras

TensorFlow

librosa

scikit learn

pandas

pyAudio

SciPy

seaborn

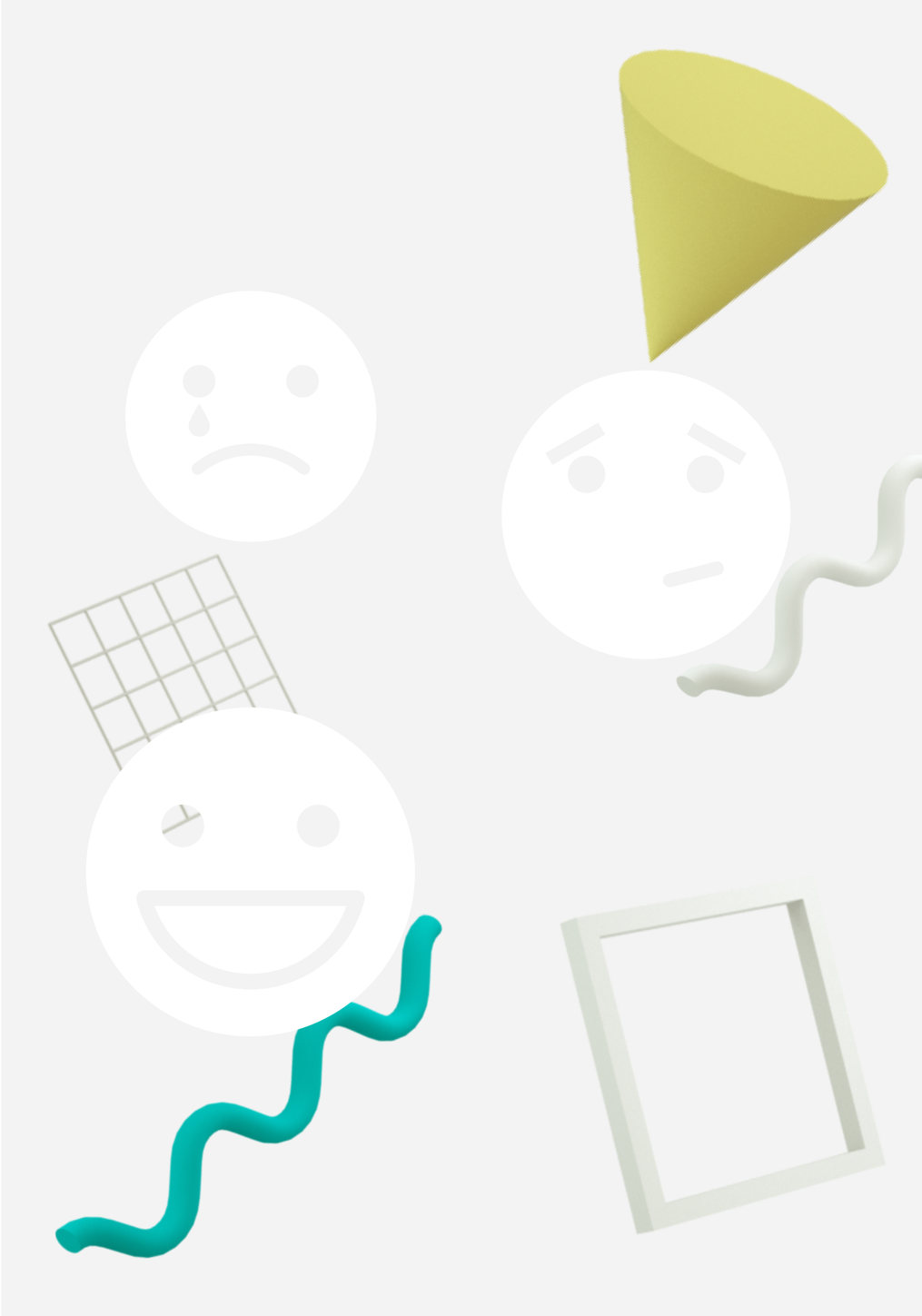
NumPy

TONE OF VOICE

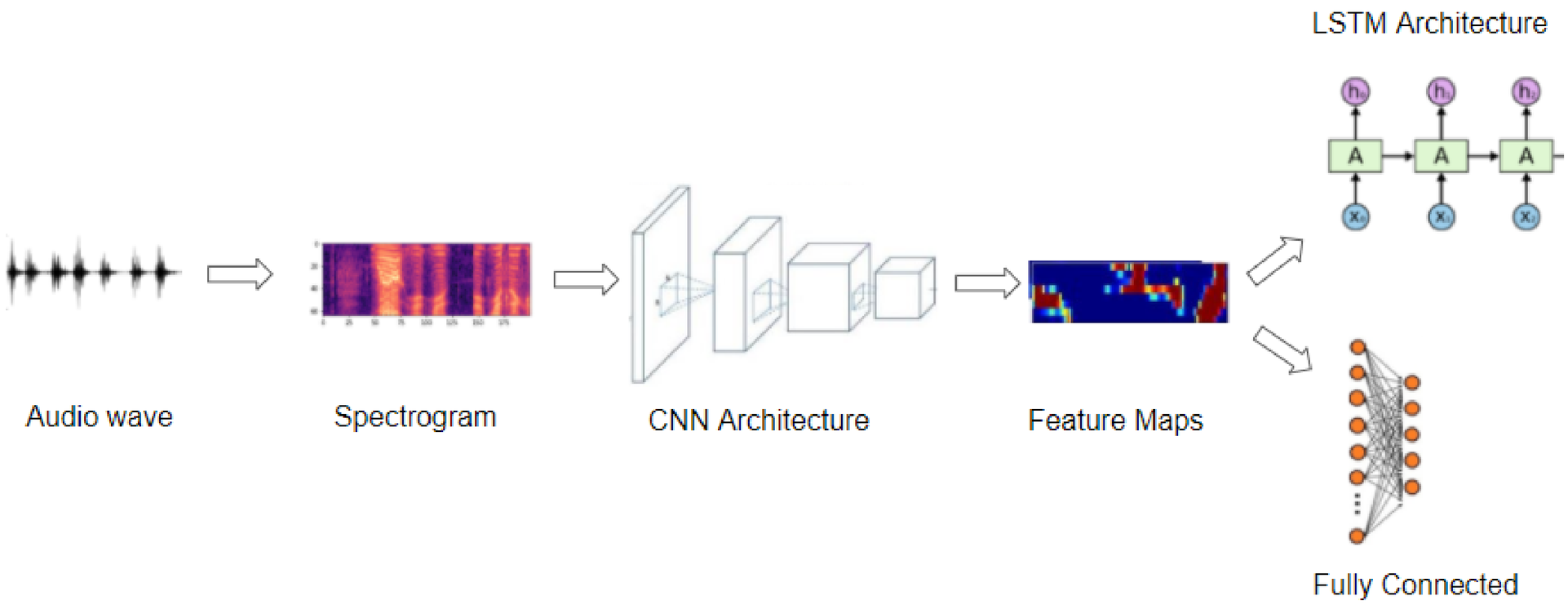
- Tone of someone's voice is **half** of the **message**
- **Humans** have the capability to distinguish emotions unlike **machines**
- Deep-Learning algorithms will help in **translation**

GOAL

- Detection of the **gender** and the **emotion** of the speaker through the application of **Deep-Learning**

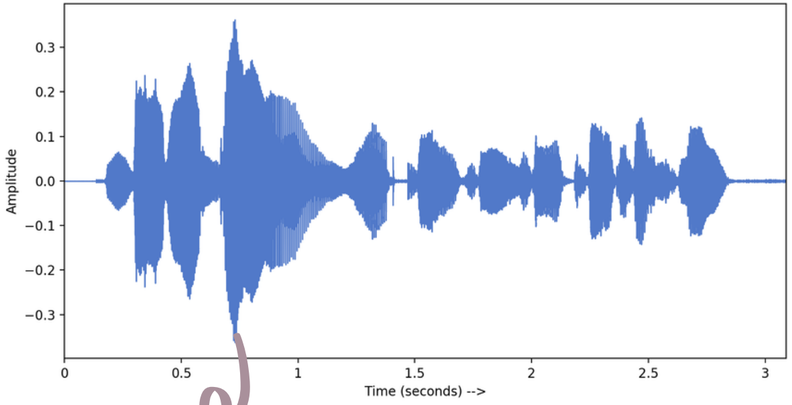


WORKFLOW


LSTM Architecture
Audio wave
Spectrogram
CNN Architecture
Feature Maps
Fully Connected

EDA

- Established sample & Sample Rate



amplitude

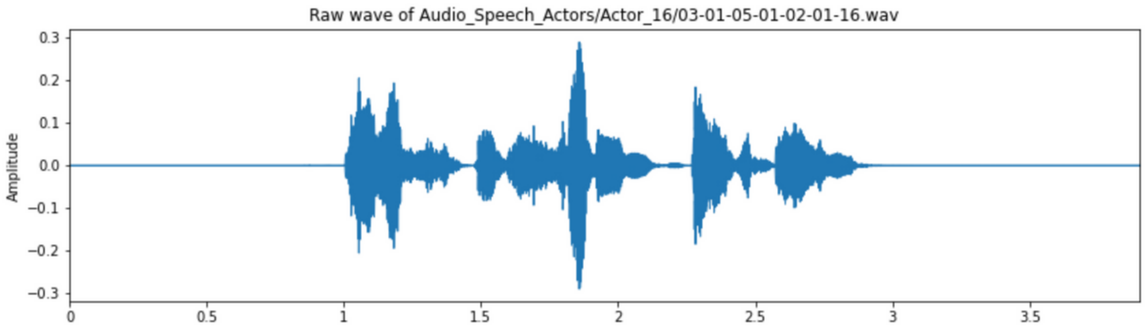
- Added new columns
  - Gender
  - Emotion
  - Intensity
  - Statement
  - Path
  - Duration

- Labeled recordings according to emotions

03-01-01-01-01-01-23.wav
03-01-01-01-01-02-23.wav
03-01-01-01-02-01-23.wav
03-01-01-01-02-02-23.wav
03-01-02-01-01-01-23.wav
03-01-02-01-01-02-23.wav
03-01-02-01-02-01-23.wav
03-01-02-01-02-02-23.wav
03-01-02-02-01-01-23.wav
03-01-02-02-01-02-23.wav
03-01-02-02-02-01-23.wav

audio
vocal channel
emotion
intensity
statement
repetition
actor

PREPROCESSING

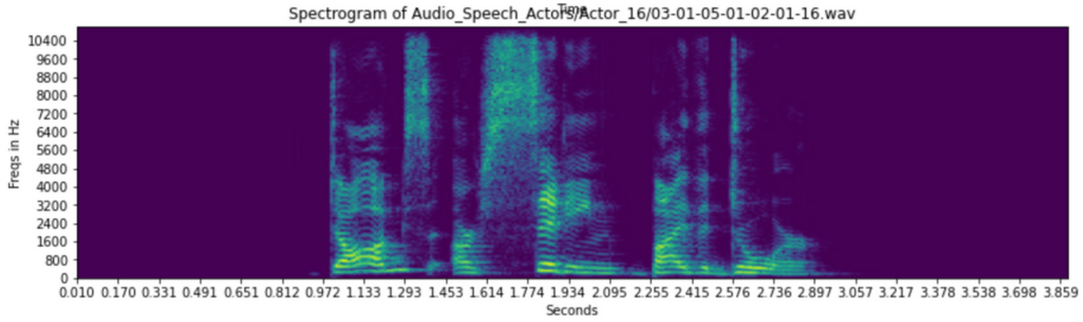


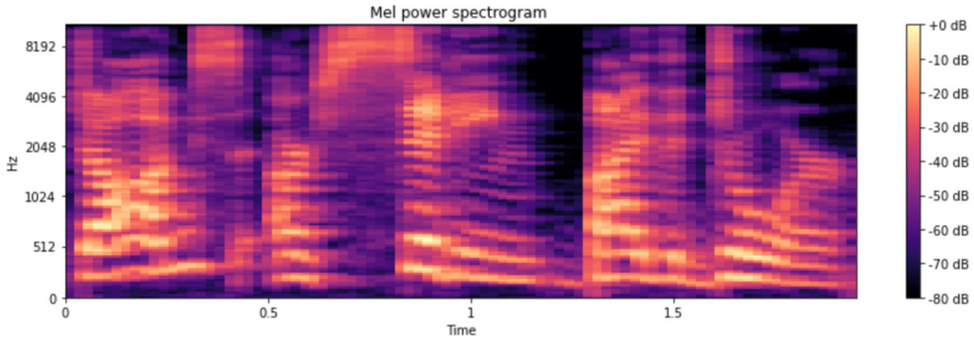
Raw wave indicates the amplitude over time

It gives us a sense of how loud or quiet a clip is at any point in time

A Spectrogram of a signal plots its Spectrum over time and is like a 'photograph' of the signal

Brighter color = High energy of signal - frequency

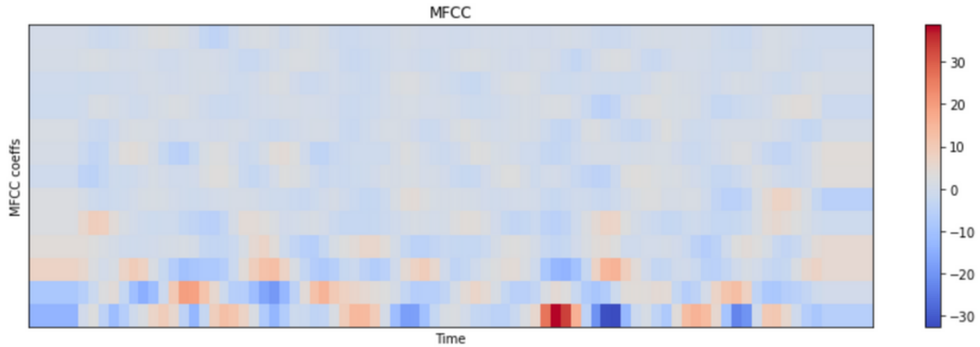




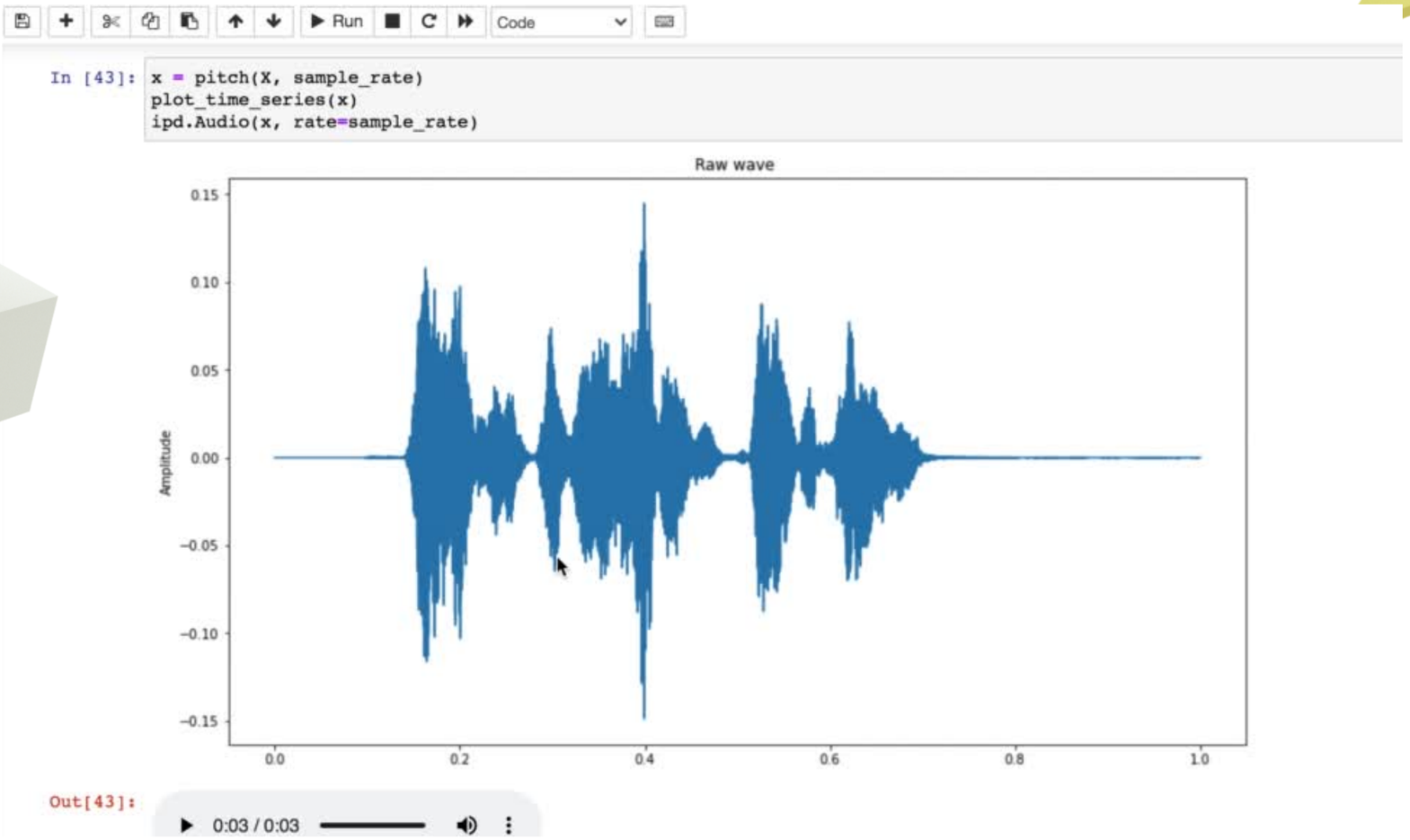
most sounds humans hear are concentrated in very small frequency and amplitude ranges

take Mel Spectrograms and apply a couple of further processing steps

The MFCC extracts a much smaller set of features that are most relevant in capturing the essential quality of the sound



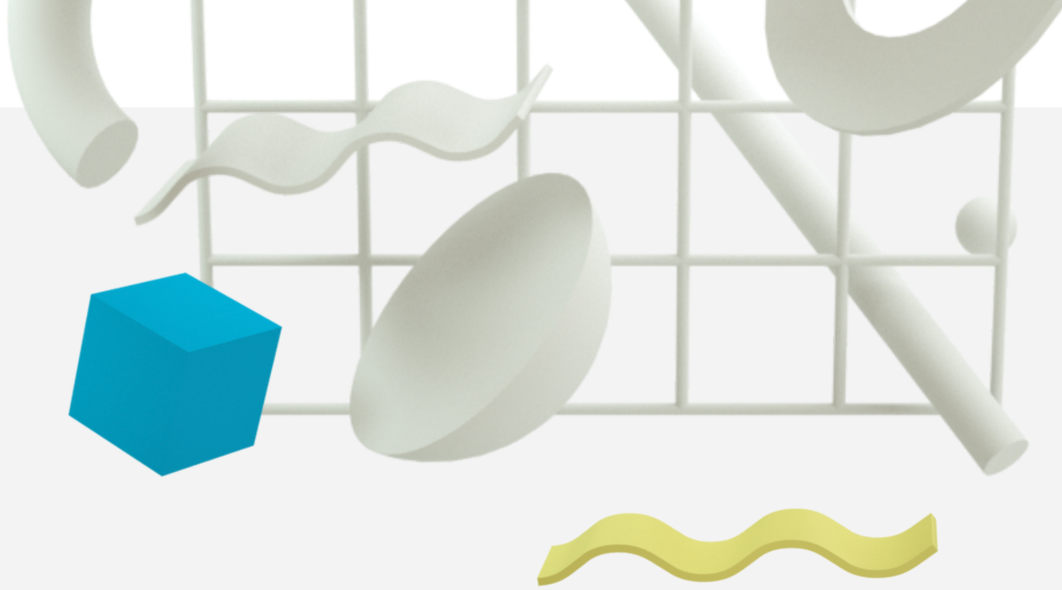
AUDIO AUGMENTATION



Implemented augmentation methods: Pitch - Noise - Stretch - Shift

- Data before augmentation 1440
- Data after augmentation 7200

CLASSIFICATION MODELS



- **Decision Tree** Classifier
  - Train : **1.0**
  - Validation : **0.3305**
  - Train + Validation : **0.8326**

- **Overfit**
- **Low scores**

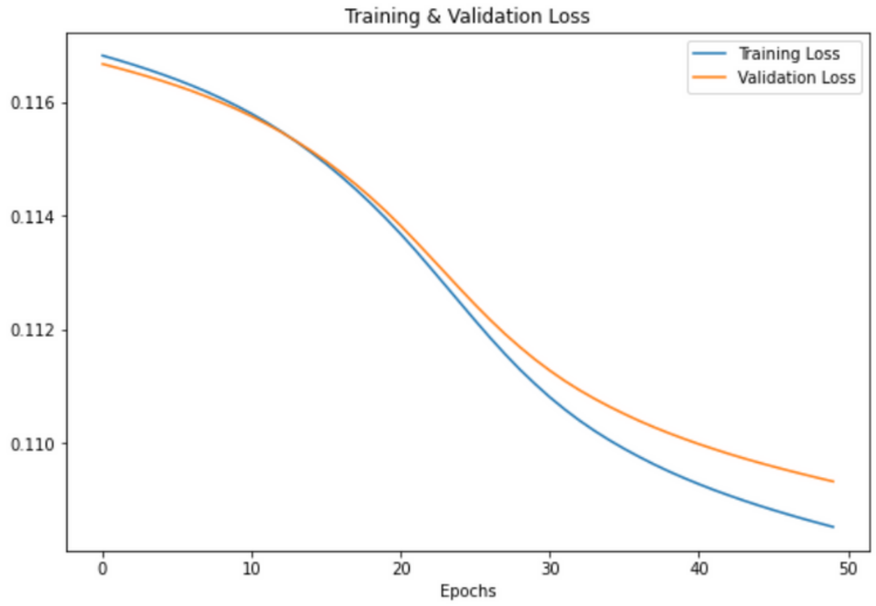
- **K Neighbors Classifier** Classifier
  - Train : **0.030**
  - Validation : **0.0256**
  - Train + Validation : **0.0289**

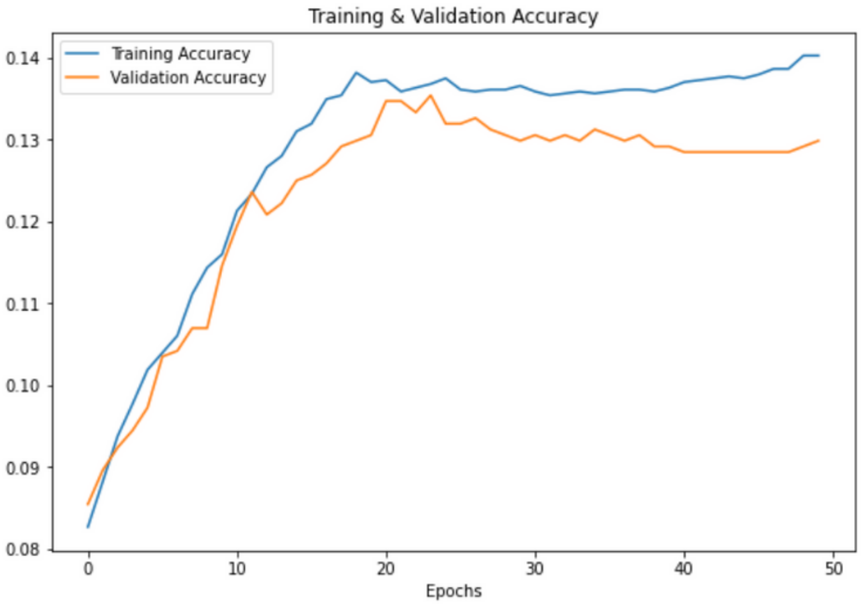
- **Incredibly low scores**

NEURAL NETWORKS MODELS

Feed-Forward Neural Networks (Baseline)

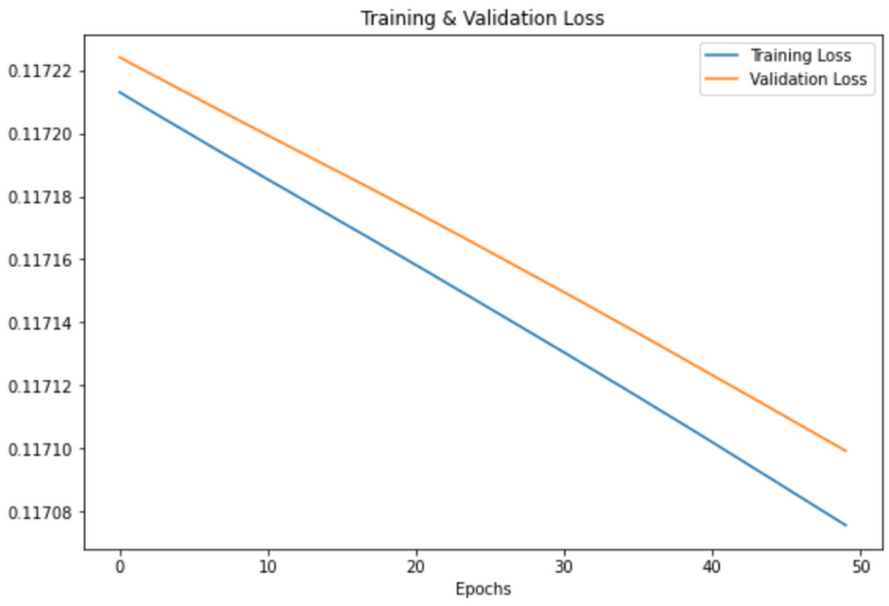
- 4 Layers
- Relu + Softmax
- Categorical accuracy
- 50 Epochs
- 64 Batch size
- 0.1403 Train
- 0.1299 Validation
- 16 Classes

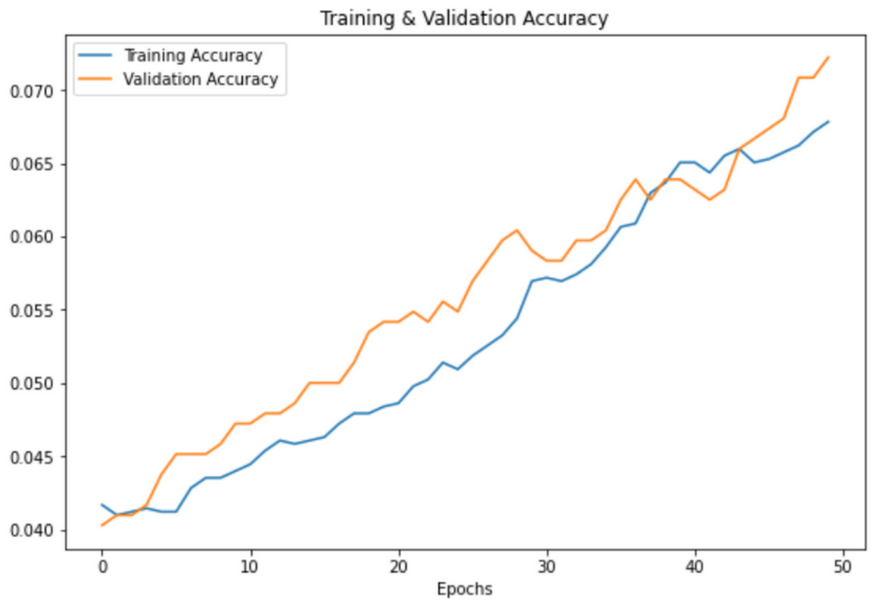




Feed-Forward Neural Networks (More Complex)

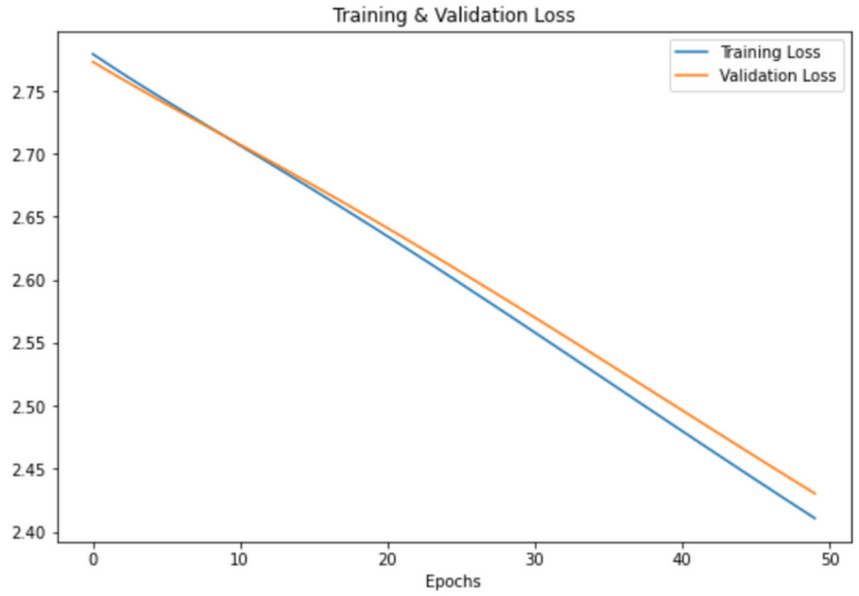
- 9 Layers
- Relu + Softmax
- Categorical accuracy
- 50 Epochs
- 64 Batch size
- 0.0678 Train
- 0.0722 Validation
- 16 Classes

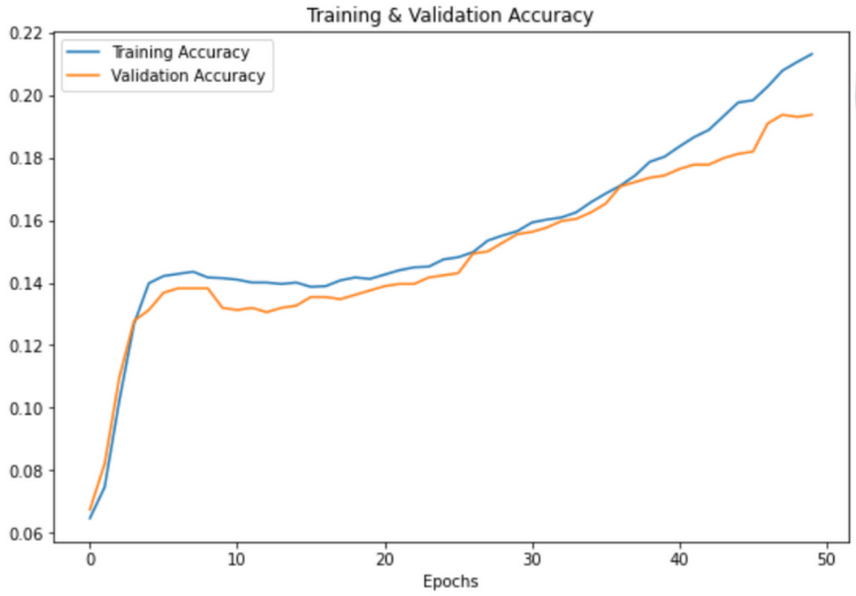




CNN (Baseline)

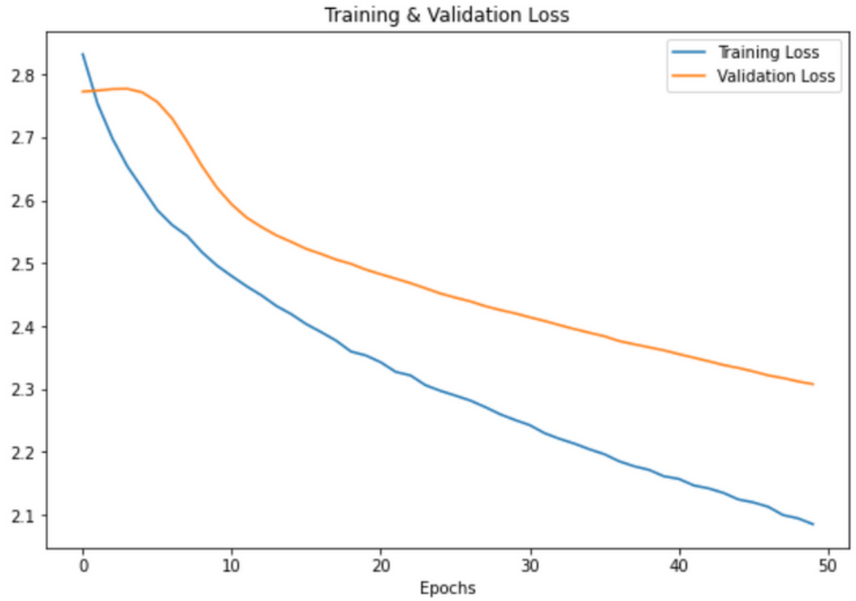
- 3 Layers
- Relu + Padding + Softmax
- Categorical accuracy
- 50 Epochs
- 64 Batch size
- 0.2132 Train
- 0.1937 Validation
- 16 Classes

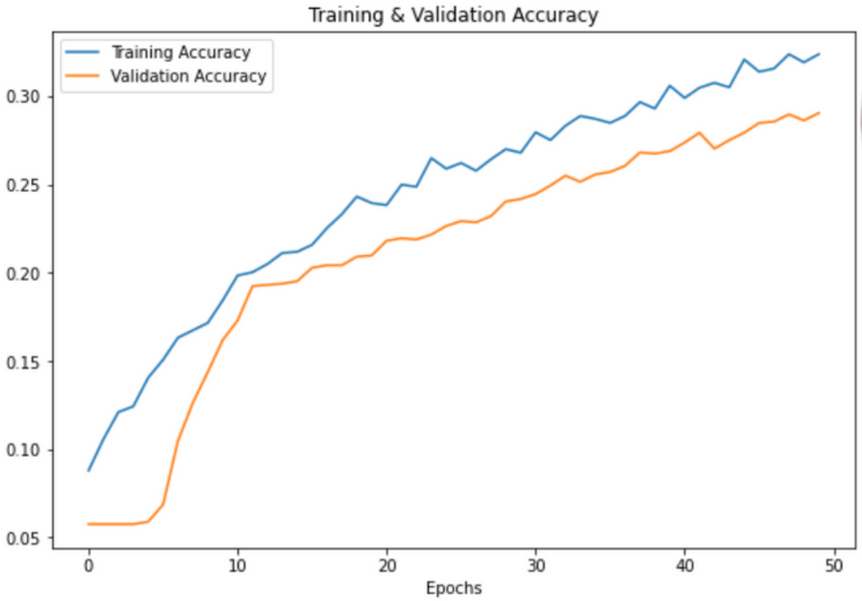




CNN (Complex)

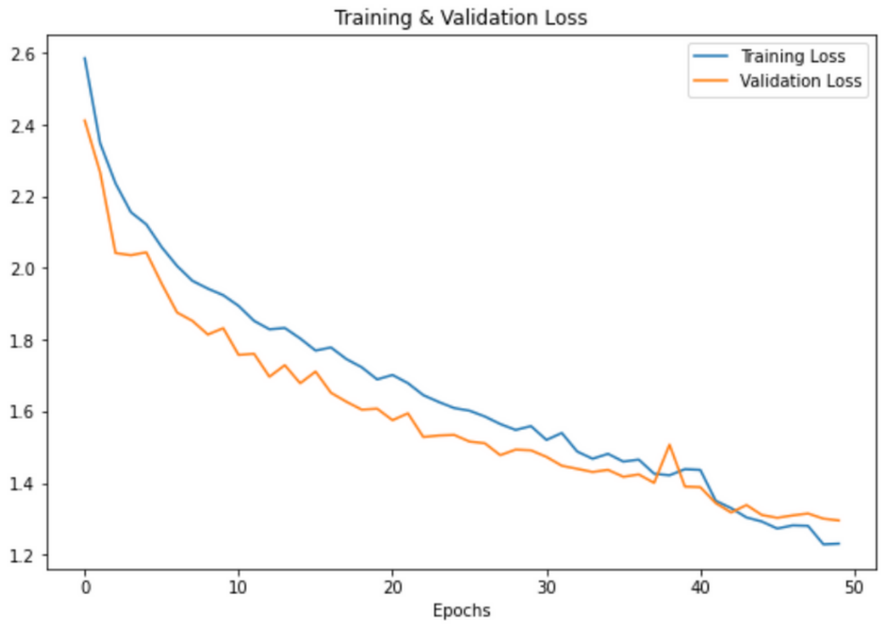
- 9 Layers
- Relu + Padding + Softmax + MaxPooling + Dropout(.25) + BatchNorm
- Categorical accuracy
- 50 Epochs
- 64 Batch size
- 0.3236 Train
- 0.2903 Validation
- 16 Classes

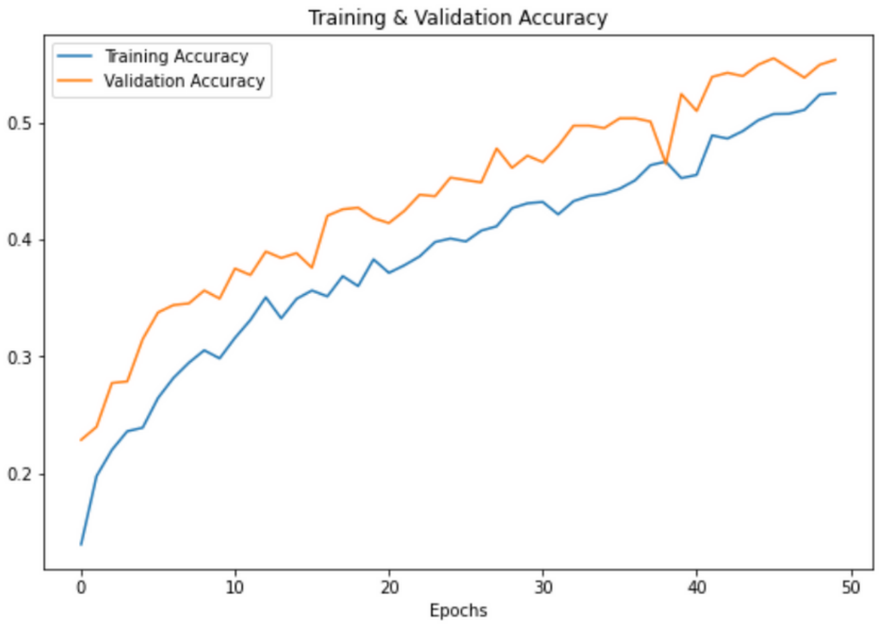




CNN (Complex 2)

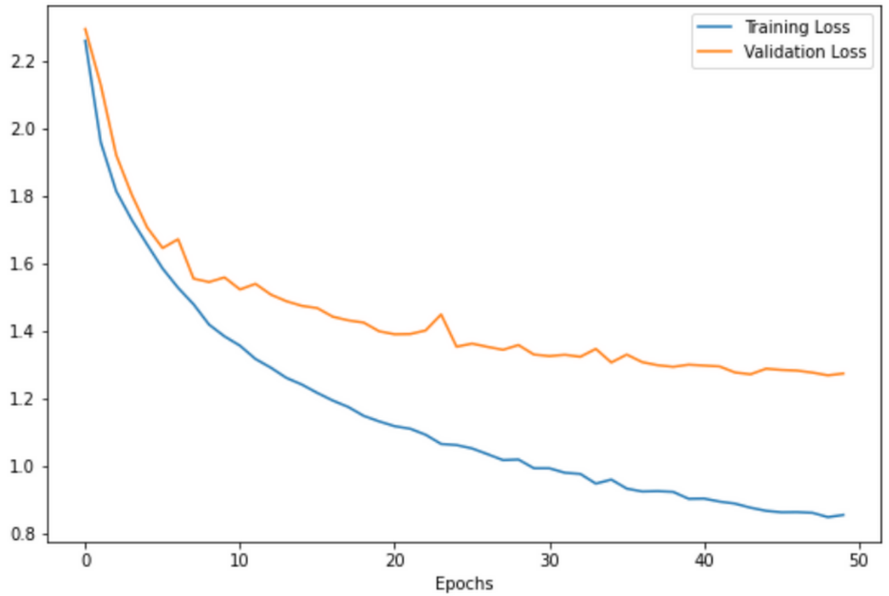
- 5 Layers
- Relu + Padding + Softmax + MaxPooling + Dropout(.25) + BatchNorm + Kernel size + stride
- Categorical accuracy
- 50 Epochs
- 30 Batch size
- 0.5250 Train
- 0.5535 Validation
- 16 Classes

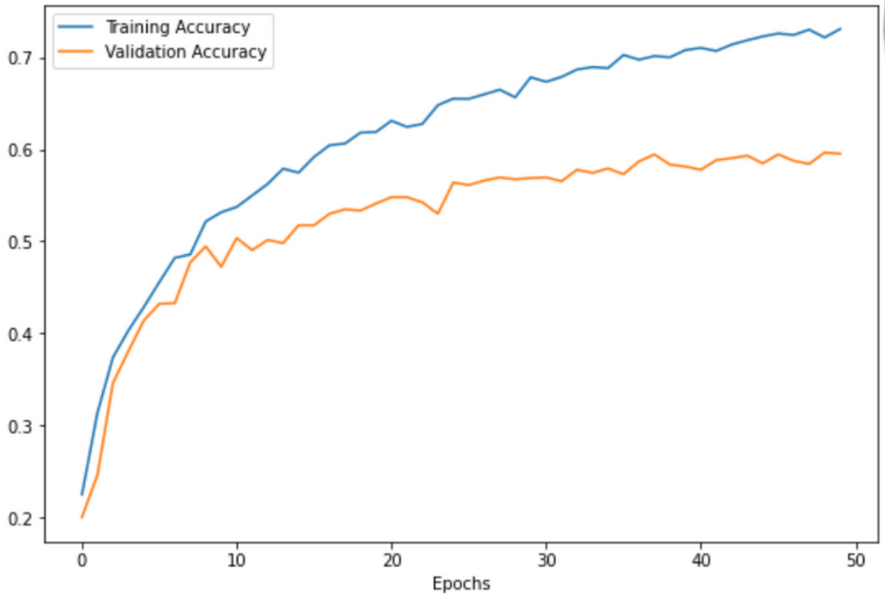




LSTM

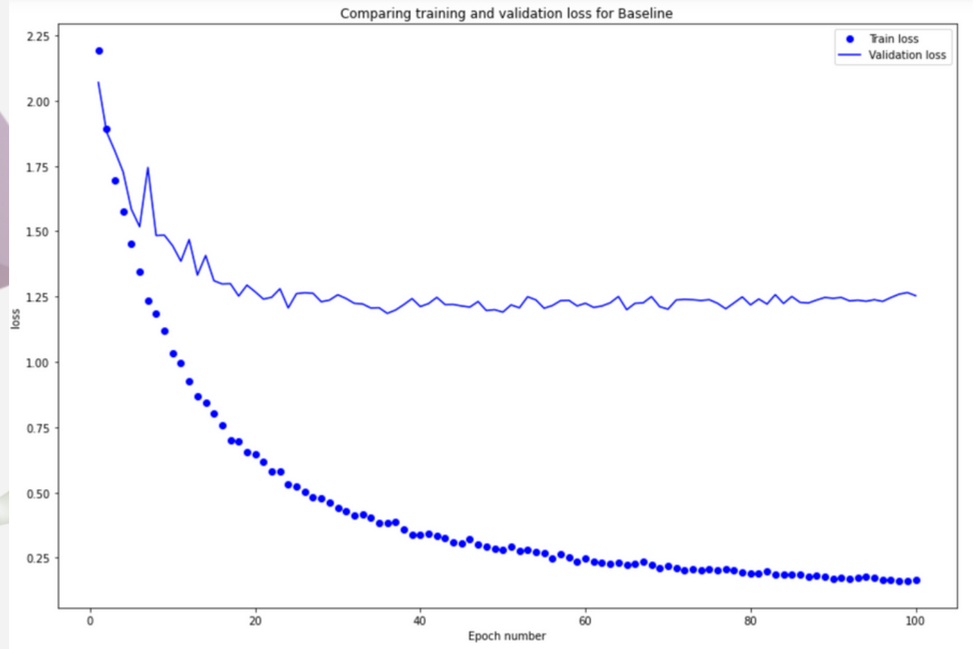
- 4 Layers
- Relu + Padding + Softmax + MaxPooling + Dropout(.25) + BatchNorm + Kernel size + stride
- Categorical accuracy
- 50 Epochs
- 30 Batch size
- 0.7308 Train
- 0.5951 Validation
- 16 Classes



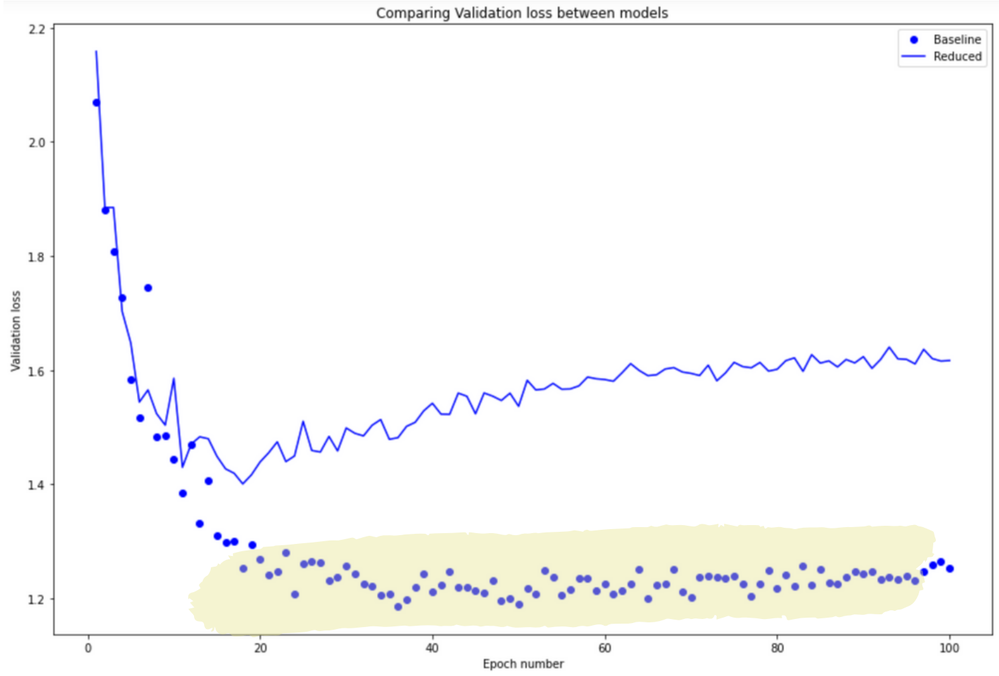


Further improvements on overfitting issue

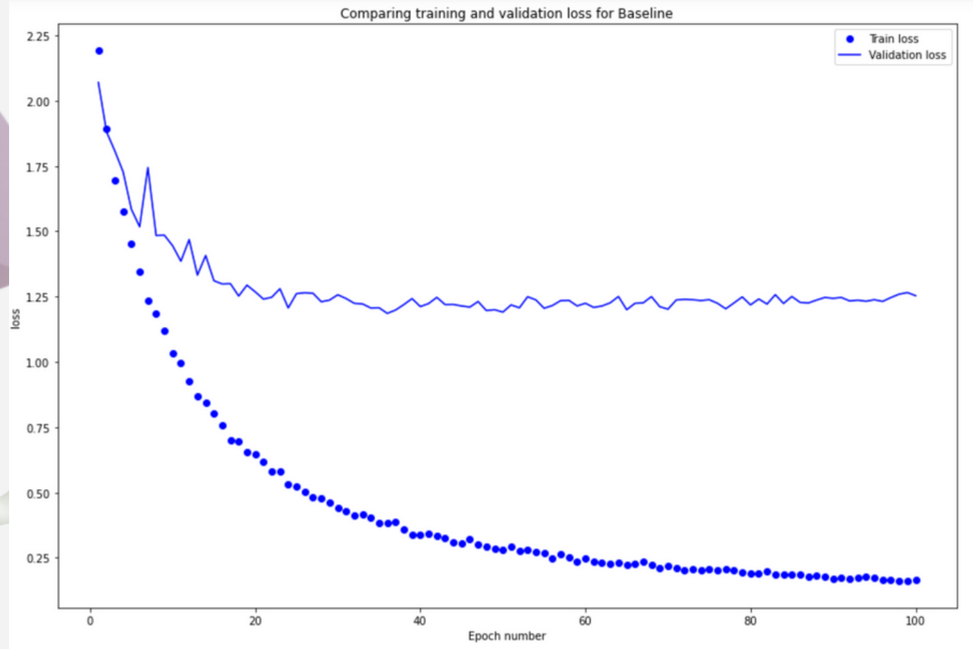
OVERFITTING SOLUTIONS



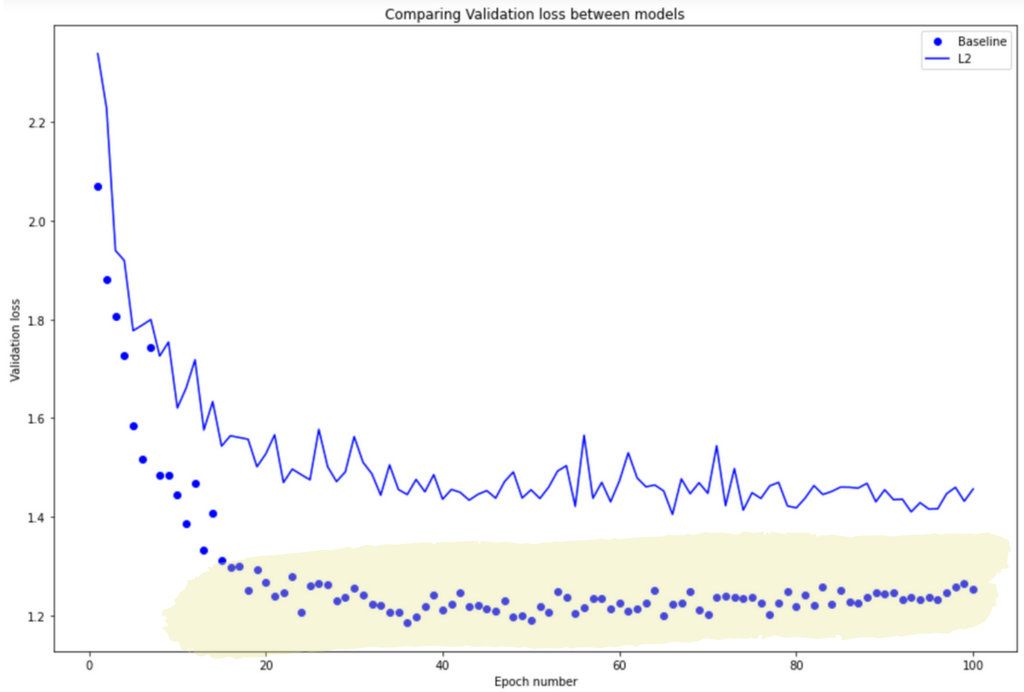
- Reduced # of layers from 4 to 3
- Reduced # of nodes in each layer



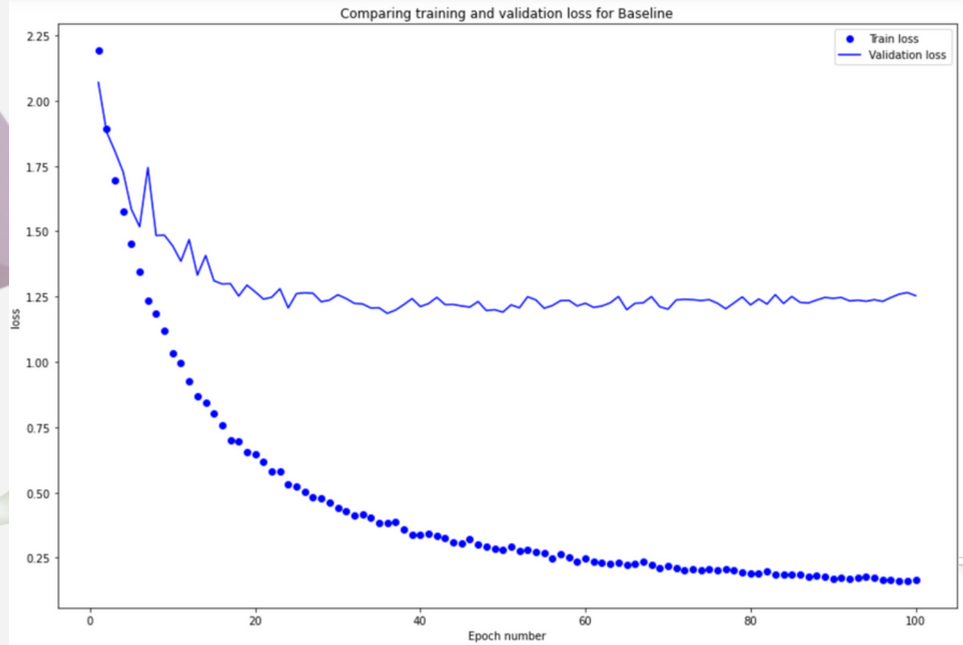
OVERFITTING SOLUTIONS



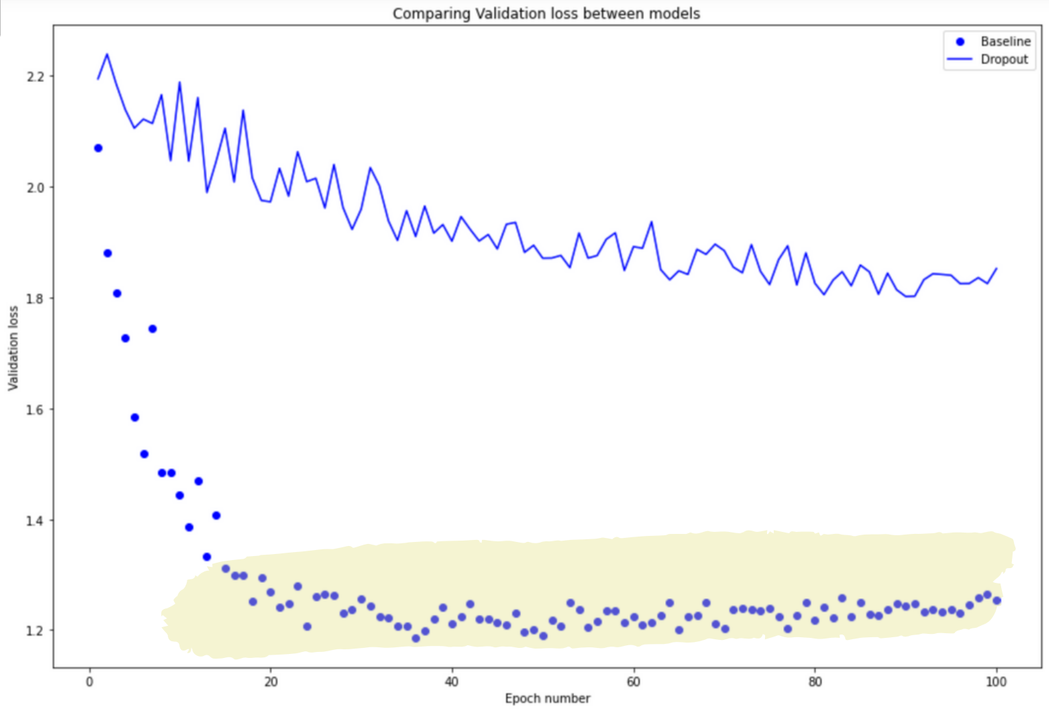
- Added L2 to 3 layers



OVERFITTING SOLUTIONS



- Increased the value of dropout to .6 / .8 / .9

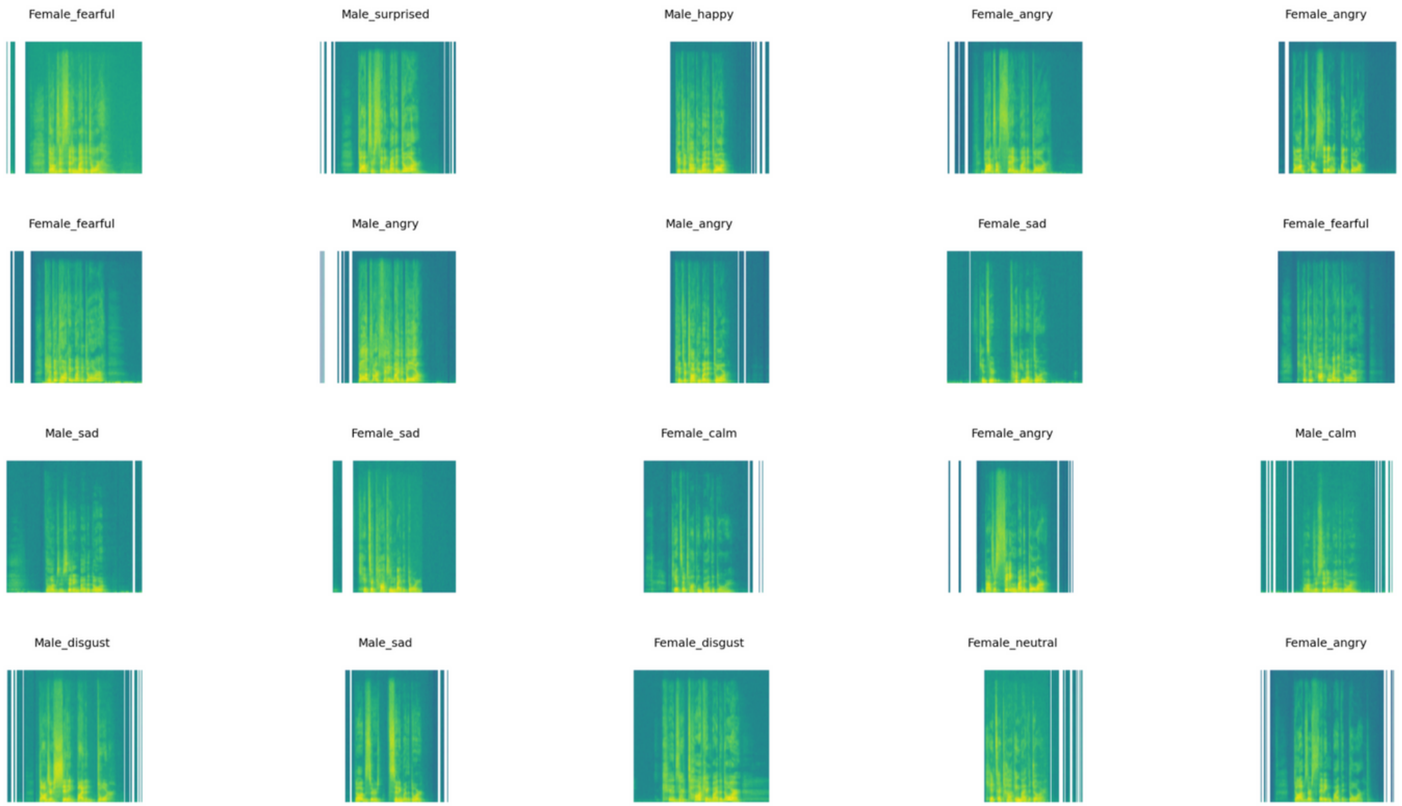


TRANSFER LEARNING

VGG16 MODEL

Spectrogram pics

Augmented by frequency blocking



Freezed first 19

Added 3 extra layers

Softmax + Dropout + relu

10 Epochs

16 Classes

Categorical accuracy

0.8728 Train

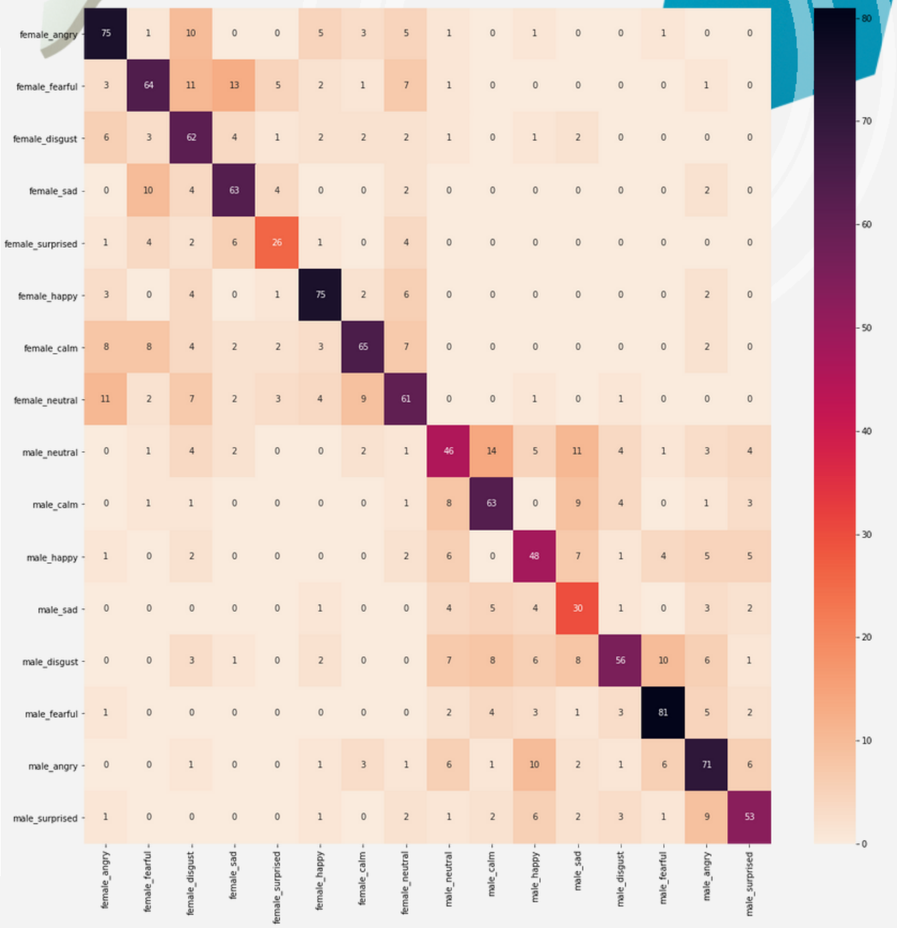
0.8996 Validation

30 Batch size

PREDICTION ON CHOSEN MODEL

- 70% on test
- Successfully detected gender
- Confusion in detection of emotions

	Actucal Emotions	Predicted Emotions
968	male_surprised	male_surprised
760	female_disgust	female_disgust
2849	male_calm	male_calm
590	male_sad	male_sad
357	male_angry	male_sad
2133	male_angry	male_fearful
400	male_surprised	male_surprised
854	female_sad	female_sad
2847	male_calm	male_calm
2641	female_angry	female_disgust



THANK YOU FOR LISTENING

ANY QUESTIONS?