DATATHON
Powered by

CUSTOMER SENTIMENT ANALYSIS

AI-Based Speech Emotion Recognition

CODANIANS

“Companies worldwide spend

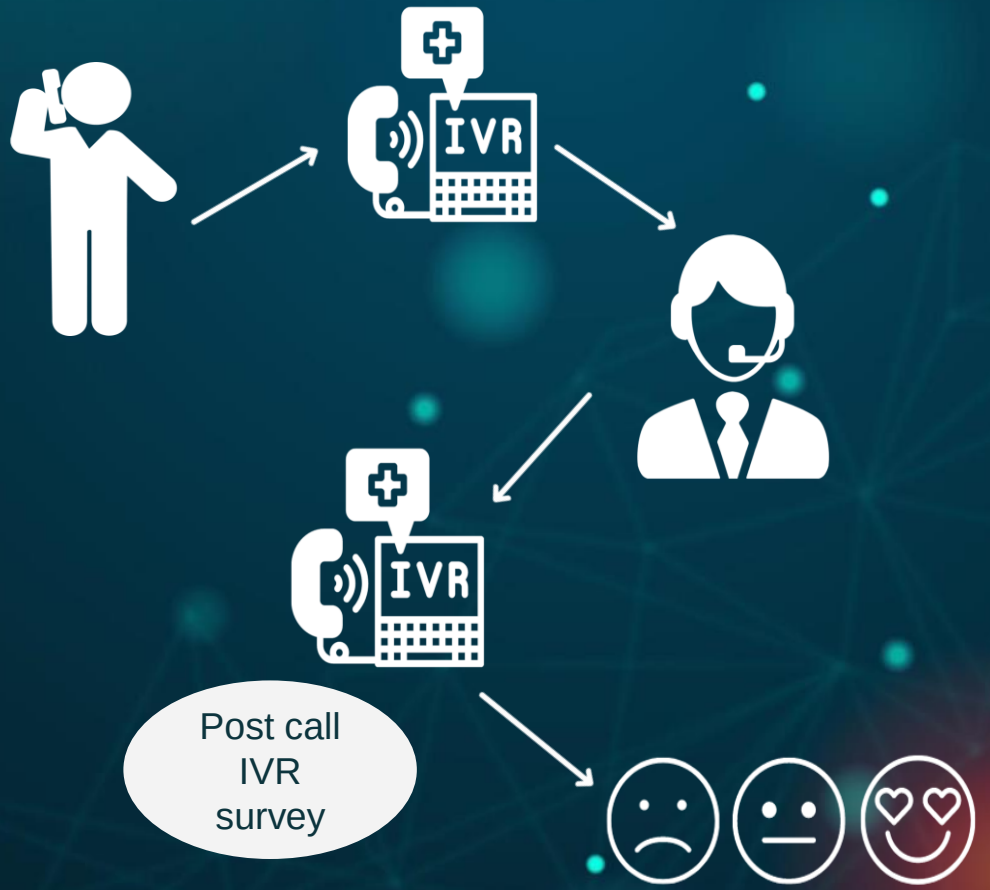
>$1.3 Trillion

for

256 Billion

customer care calls per annum”

~ Analytics Insights



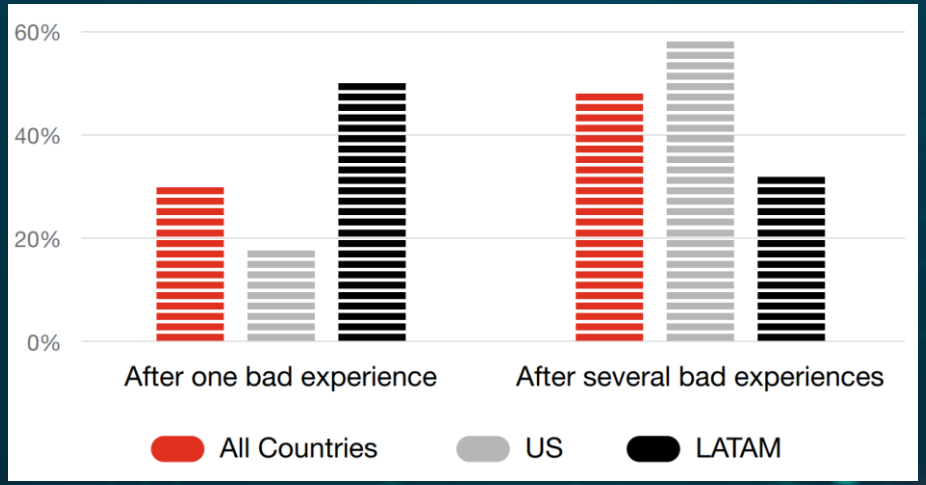
CODANIANS

Post Call IVR Survey

- Easy to Execute
- Low Cost
- Easy to Get Agent Level Scores

- Low Response Rates
- Low Volumes of Qualitative Feedback
- Fail to Depict the Sentiments of Customers
- Difficult to Assess True Feedback

Leaving Ratio of Customers

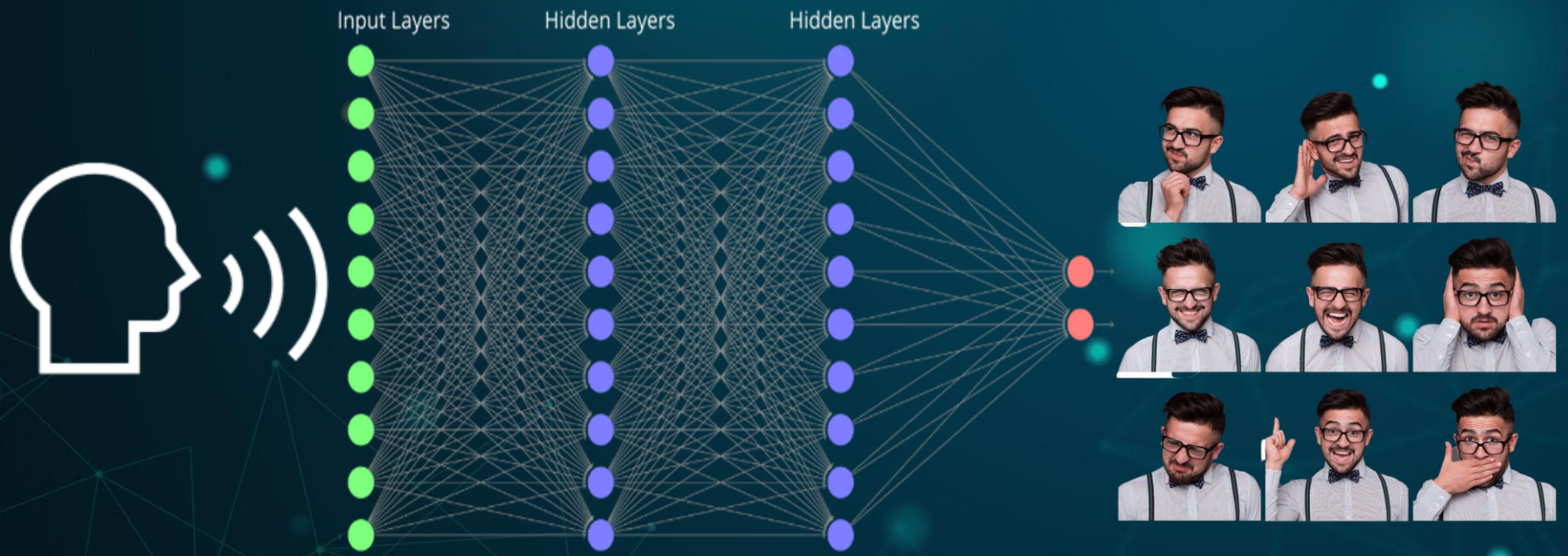


https://www.pwc.com/future-of-cx

CODANIANS

How to Do it?

CODANIANS



Speech Signal Deep Neural Network Emotions

CODANIANS

BENEFITS

Cost Effectiveness

Customer Satisfaction

Brand Reputation

Increased ROI

Customer Rentention

CODANIANS

METHODOLOGY

Data Sets

01

RAVDESS, SAVEE, CREMA-D, TESS, Speech Accent Archieve

Features Extraction

03

Zero Crossing Rate, Chroma_stft, MFCC, RMS value, MelSpectogram

Model Evaluation

05

Confusion Matrix, Classification Report

CODANIANS

02 Exploratory Data Analysis

Drawing Insights from Data

04 Model Training

Convolutional Neural Network

06 Emotion Prediction

Predicted Emotions from Speech Accent Archive Dataset


pandas
SciPy
K
CODANIANS

DATA SETS

RAVDESS Ryerson Audio-Visual Database of Emotional Speech and Song

TESS Toronto Emotional Speech Set

SAVEE Surrey Audio-Visual Expressed Emotion

CREMA-D Crowd-sourced Emotional Mutimodal Actors Dataset

ACCENT Speech Accent Archive

CODANIANS

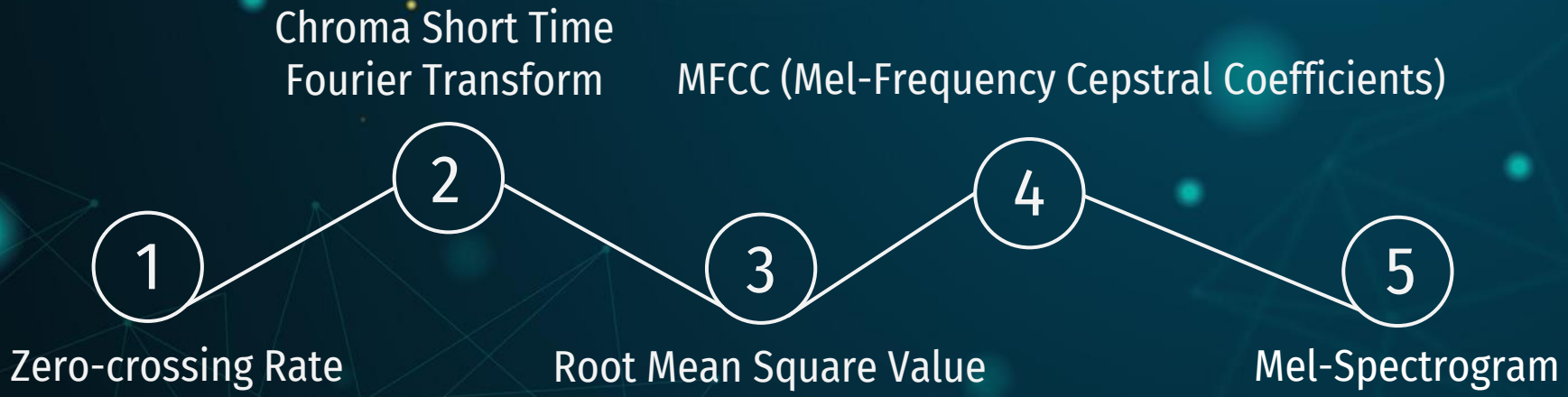
DATA AUGMENTATION

The process to augment or enhance the data with some modification in the given data

CODANIANS

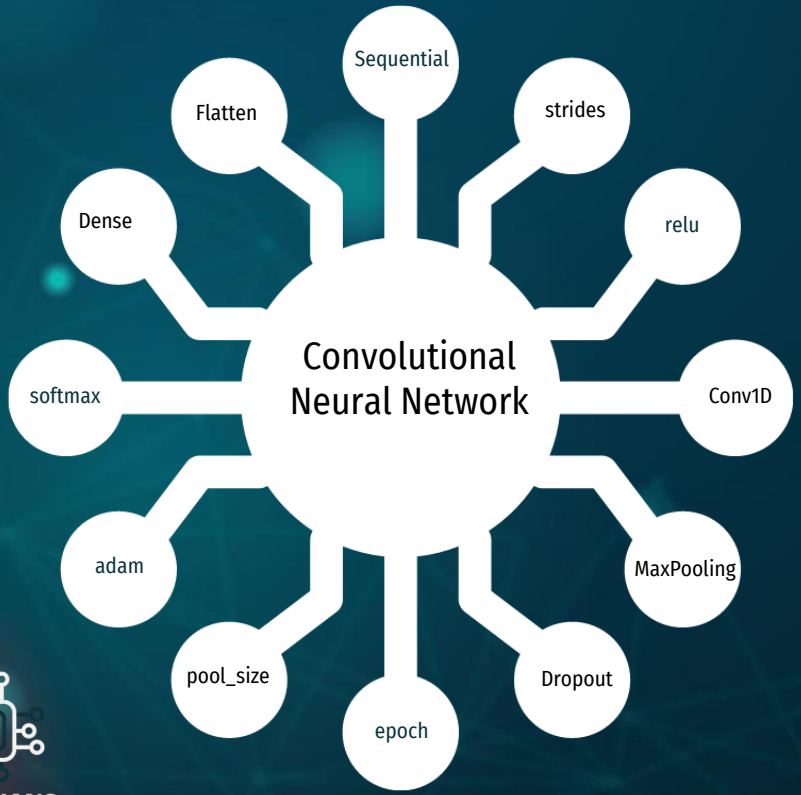
FEATURE EXTRACTION

The process of retrieving information from audio signal is called features extraction.



CODANIANS

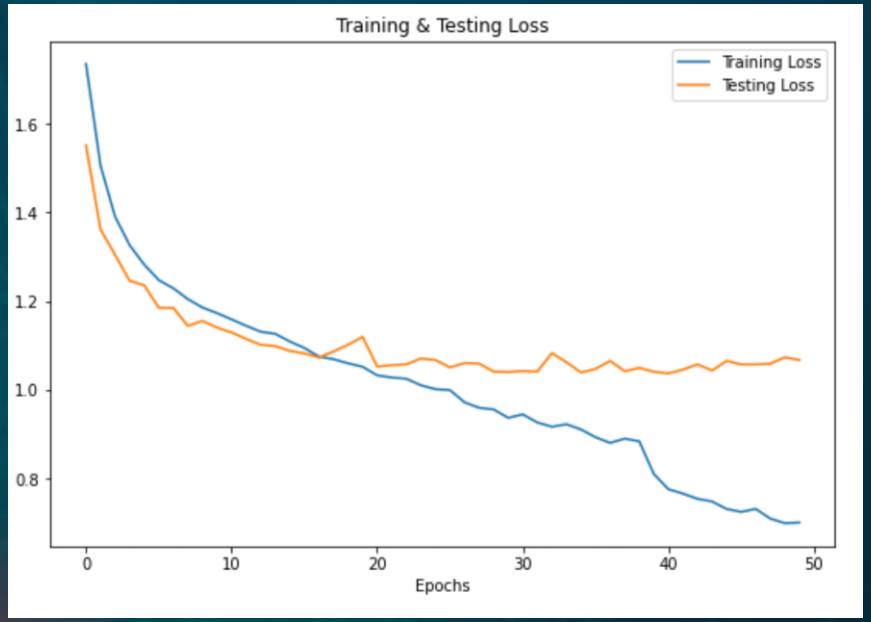
MODEL TRAINING

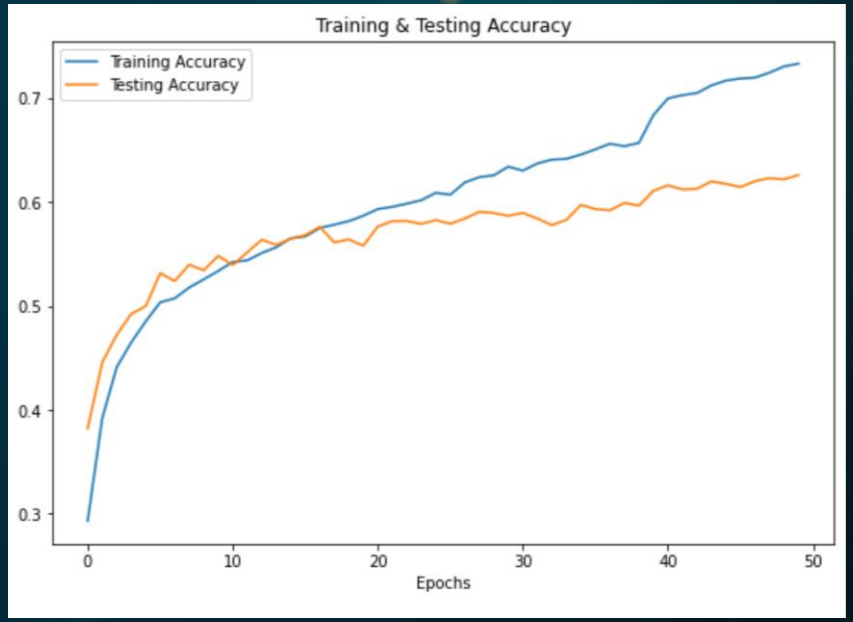



```
Model: "sequential_7"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d_35 (Conv1D)          (None, 162, 256)          1536

 max_pooling1d_35 (MaxPoolin  (None, 81, 256)          0
 g1D)

 conv1d_36 (Conv1D)          (None, 81, 256)           327936

 max_pooling1d_36 (MaxPoolin  (None, 41, 256)          0
 g1D)

 conv1d_37 (Conv1D)          (None, 41, 128)           163968

 max_pooling1d_37 (MaxPoolin  (None, 21, 128)          0
 g1D)

 dropout_21 (Dropout)        (None, 21, 128)           0

 conv1d_38 (Conv1D)          (None, 21, 128)           82048

 max_pooling1d_38 (MaxPoolin  (None, 11, 128)          0
 g1D)

 dropout_22 (Dropout)        (None, 11, 128)           0

 conv1d_39 (Conv1D)          (None, 11, 64)            41024

 max_pooling1d_39 (MaxPoolin  (None, 6, 64)            0
 g1D)

 flatten_7 (Flatten)         (None, 384)               0

 dense_14 (Dense)            (None, 32)                12320

 dropout_23 (Dropout)        (None, 32)                0

 dense_15 (Dense)            (None, 8)                 264

=================================================================
Total params: 629,096
Trainable params: 629,096
Non-trainable params: 0
_________________________________________________________________
```

CODANIANS

LOSS AND ACCURACY





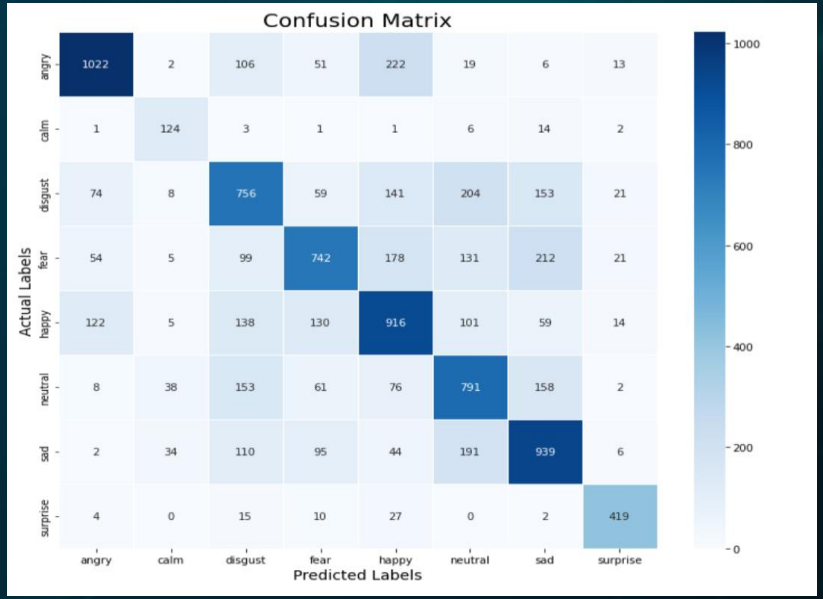
CODANIANS

MODEL EVALUATION

Classification Report

	precision	recall	f1-score	support
angry	0.79	0.71	0.75	1441
calm	0.57	0.82	0.67	152
disgust	0.55	0.53	0.54	1416
fear	0.65	0.51	0.57	1442
happy	0.57	0.62	0.59	1485
neutral	0.55	0.61	0.58	1287
sad	0.61	0.66	0.63	1421
surprise	0.84	0.88	0.86	477
accuracy			0.63	9121
macro avg	0.64	0.67	0.65	9121
weighted avg	0.63	0.63	0.63	9121

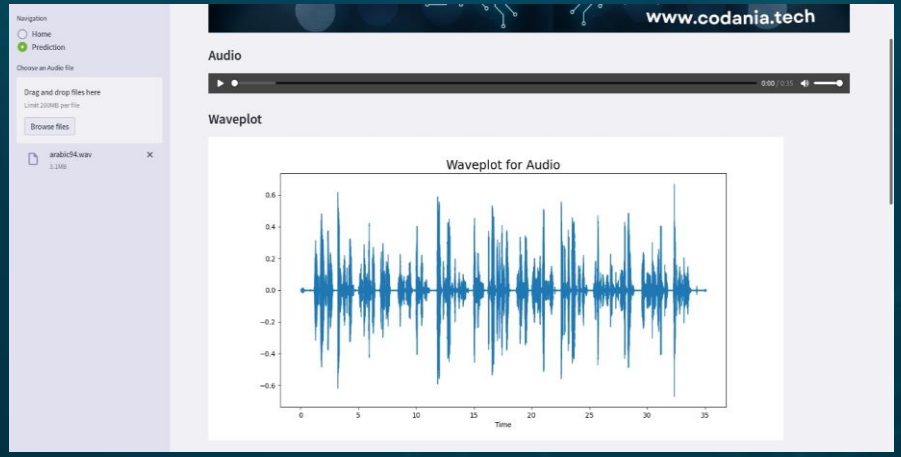
Confusion Matrix

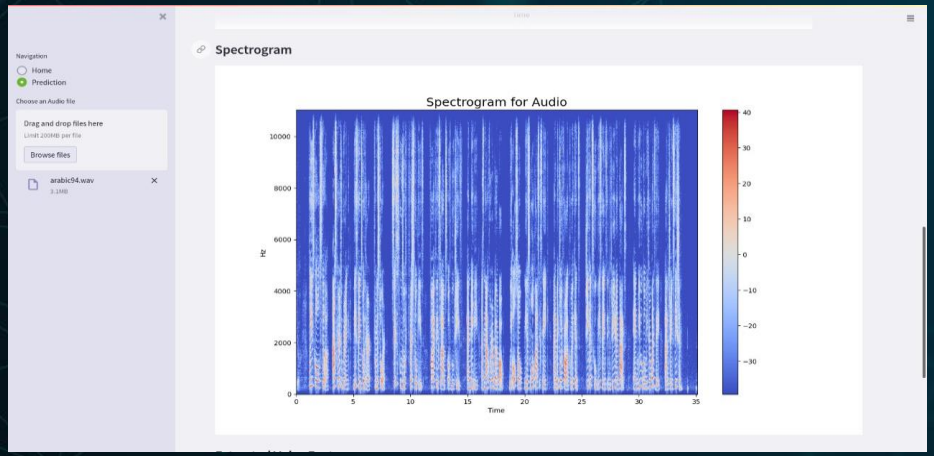


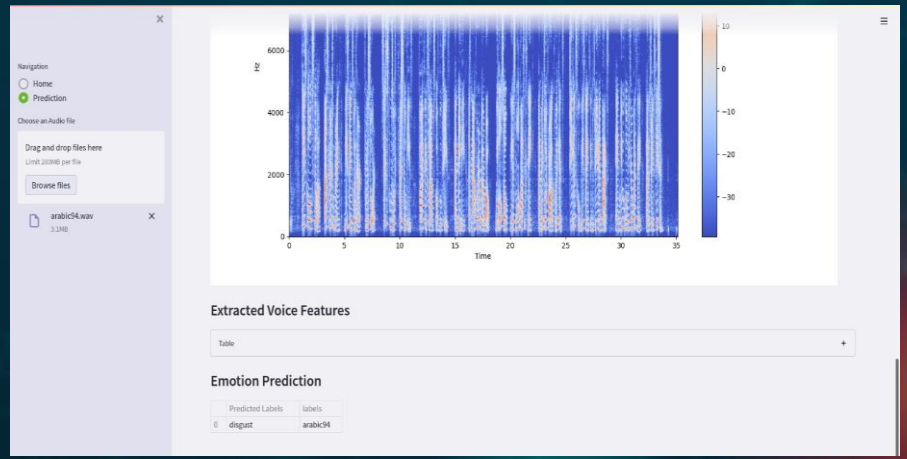
CODANIANS

PROTOTYPE


Navigation
Home
Prediction
Choose an Audio file
Drag and drop files here
Limit 200MB per file
Browse files
Speech Emotion Recognition
Shows Customers Emotions!
www.codania.tech


www.codania.tech
Audio
Waveplot
Waveplot for Audio
Time


Spectrogram
Spectrogram for Audio
Hz
Time


Hz
Time
Extracted Voice Features
Table
Emotion Prediction
Predicted Labels
labels
disgust
arabic94

FUTURE PROSPECTS

SER
Speech
Emotion
Recognition

Combined with Conversational AI, SER is a perfect model for Voice Chat bots

CONVESATIONAL AI CHATBOTS

Model can be trained on local languages in Pakistan

SER IN LOCAL LANGUAGES

TARGETED MARKETING

Customized offers based on customers' speech emotions

SER IN HEALTH CARE

In psychiatry wards Emotion Recognition can be used to assess patients mental status

CODANIANS