

Emotion Recognition of Call Center Conversations

Robert **Bosch** Engineering and Business Solutions Private Limited



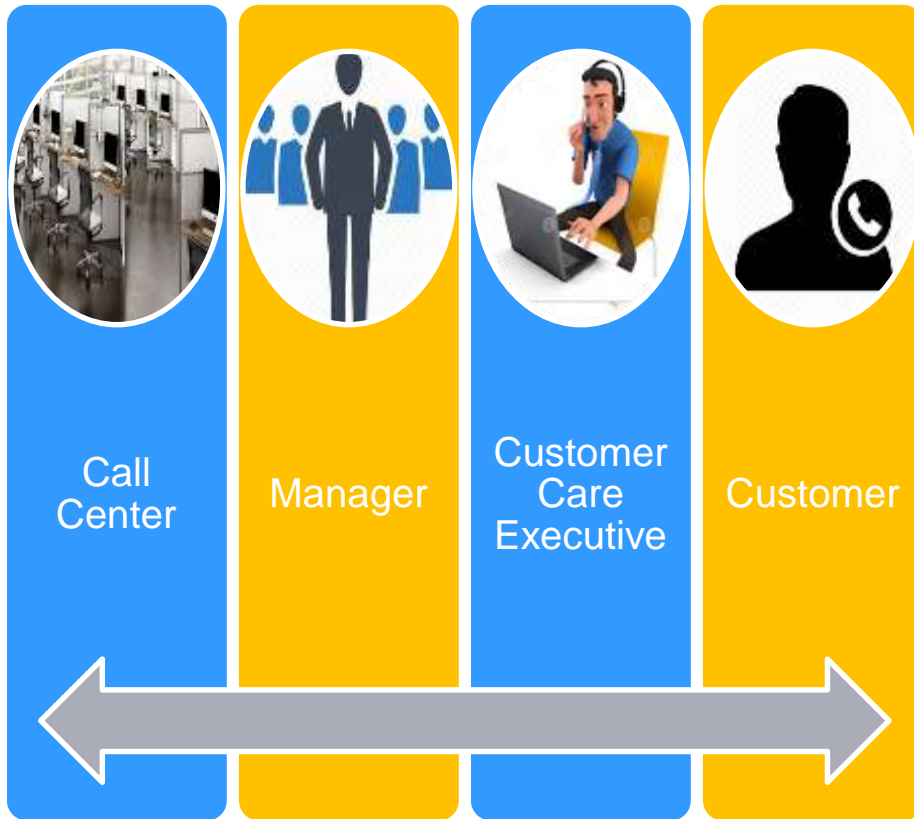
BOSCH

Agenda

- 1 • Introduction
- 2 • Problem Definition
- 3 • Solution Overview
- 4 • Why Consider Emotions in Conversations ?
- 5 • Key Features
- 6 • Work Flow
- 7 • Results
- 8 • Technology
- 9 • Benefits
- 10 • Future Scope of Work



Introduction

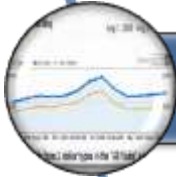


- ❖ Today Call Centers are most prominent means of providing service to customers
- ❖ Call Centre Executives provide product information and helpline to existing customers as well as promote and advertise products through telephonic conversations
- ❖ **Managers at Call Centers manage the call centre executives and grade them based on the success of audio calls recorded**

Problem Definition



No Scientific Tool to Grade Customer Care Executives Based on Historical/Recorded Data



Unable to Evaluate Product Advertisement Success Ratio based on recorded calls

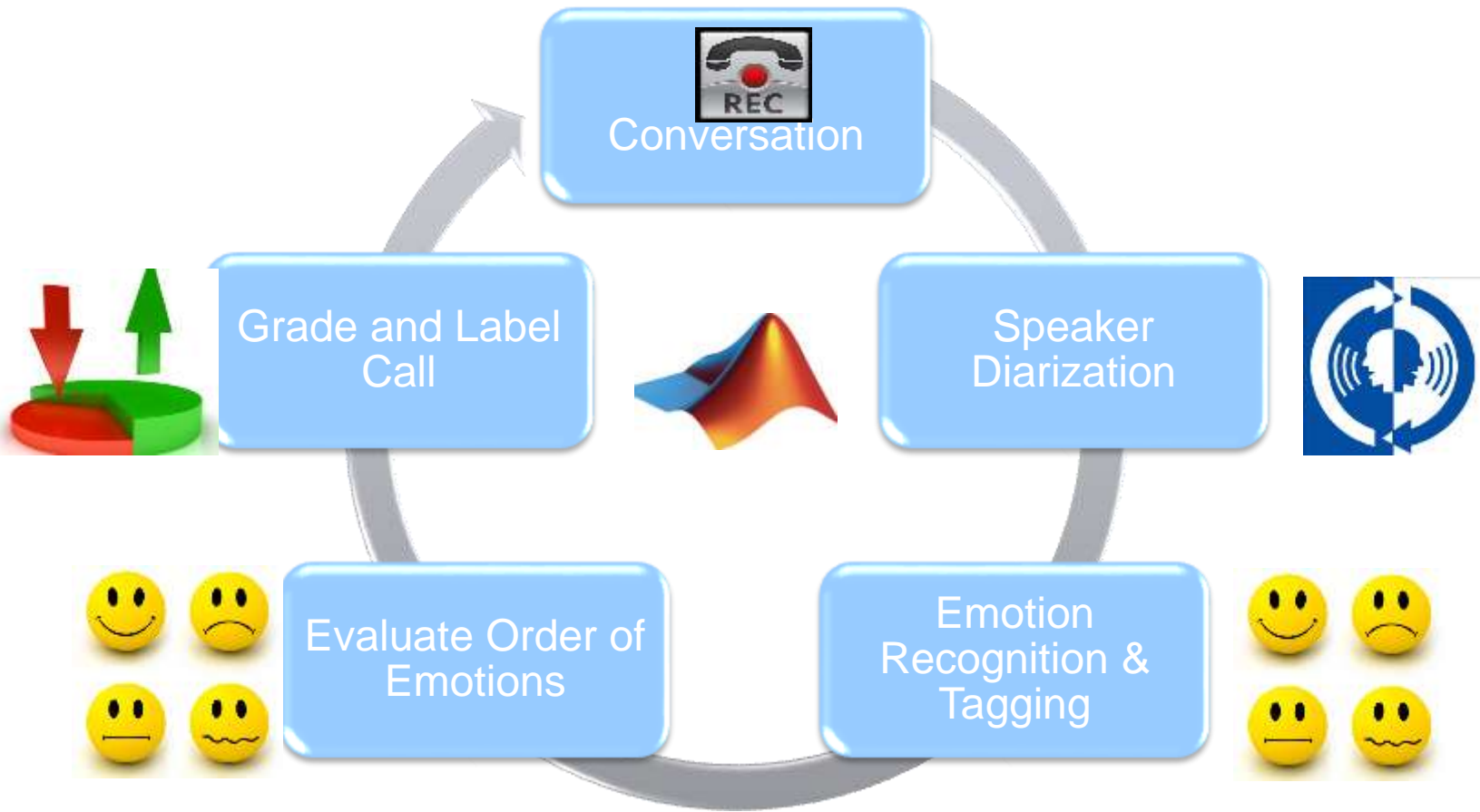


Managers have to manually generate daily/monthly evaluation of quality of calls



No Tool uses Emotion Recognition to process and analyze the recorded audio calls

Solution Overview



Why Consider Emotions in Conversations ?

Acoustic Features –Pitch,Fo

- Arguement
- Question
- Reply



Mel Frequency(MFCC)

- Anger
- Sadness
- Disgust



Speaker Diarization

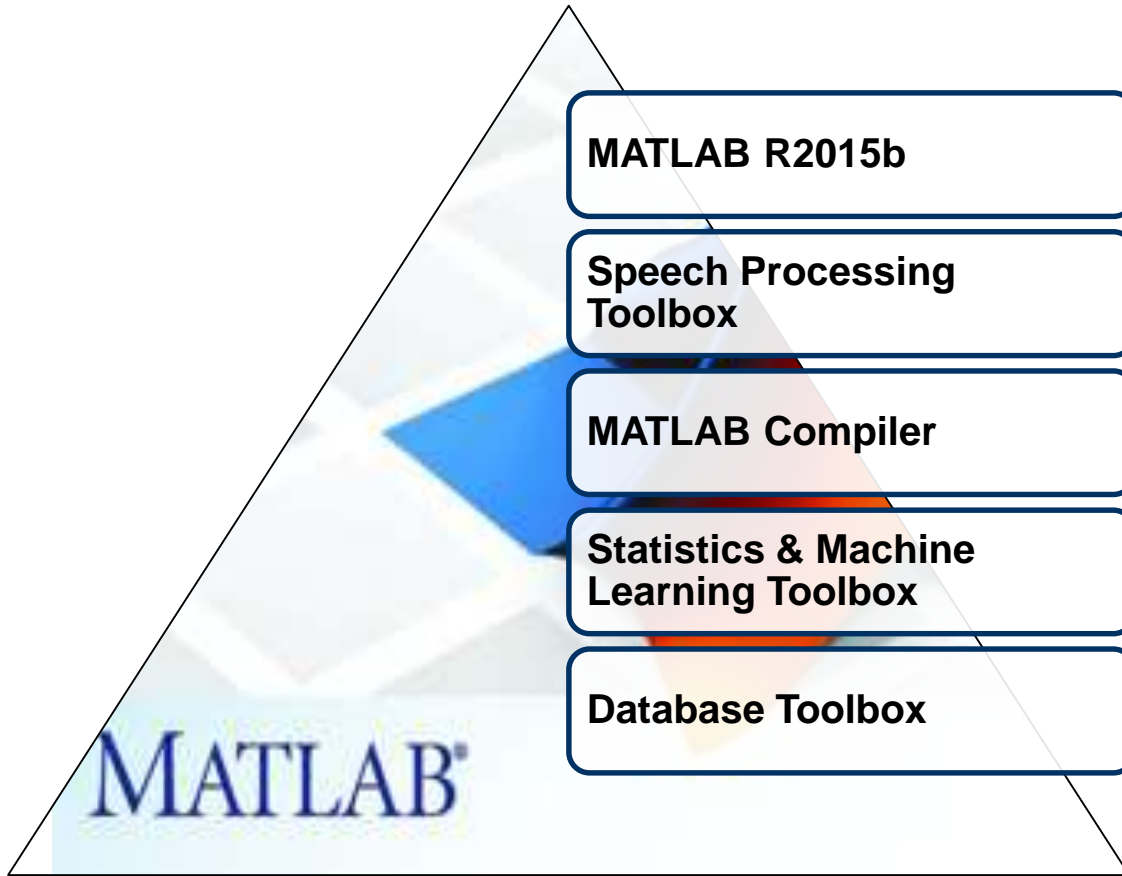
Machine Learning

Data Analytics



BOSCH

Technology



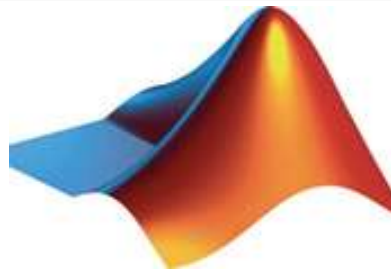
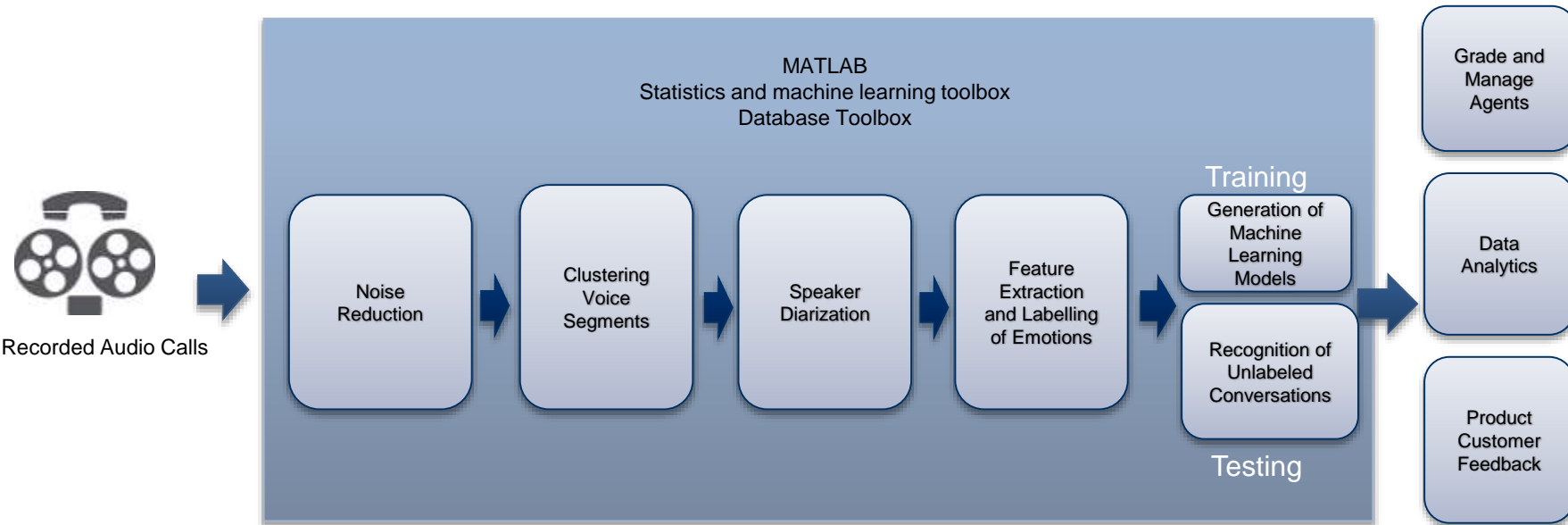
Key Features

- ❖ Data Analytics about Product Advertisement Success Ratio based on the call which were in agreement with Customer Care Executive
- ❖ Daily analytics about percentage of calls which are able to convey service information to customers with and without feedback from analyzing the order of emotions and acts in the conversation
- ❖ Grading the Customer Care Executives by Processing the monthly/yearly recorded audio calls

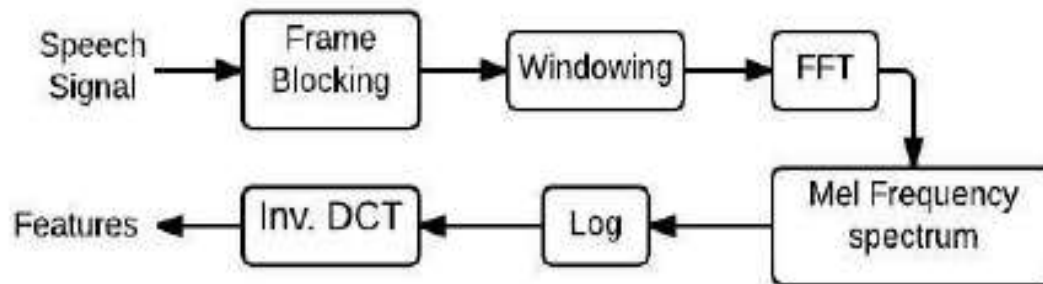


Emotion Recognition of Call Center Conversations

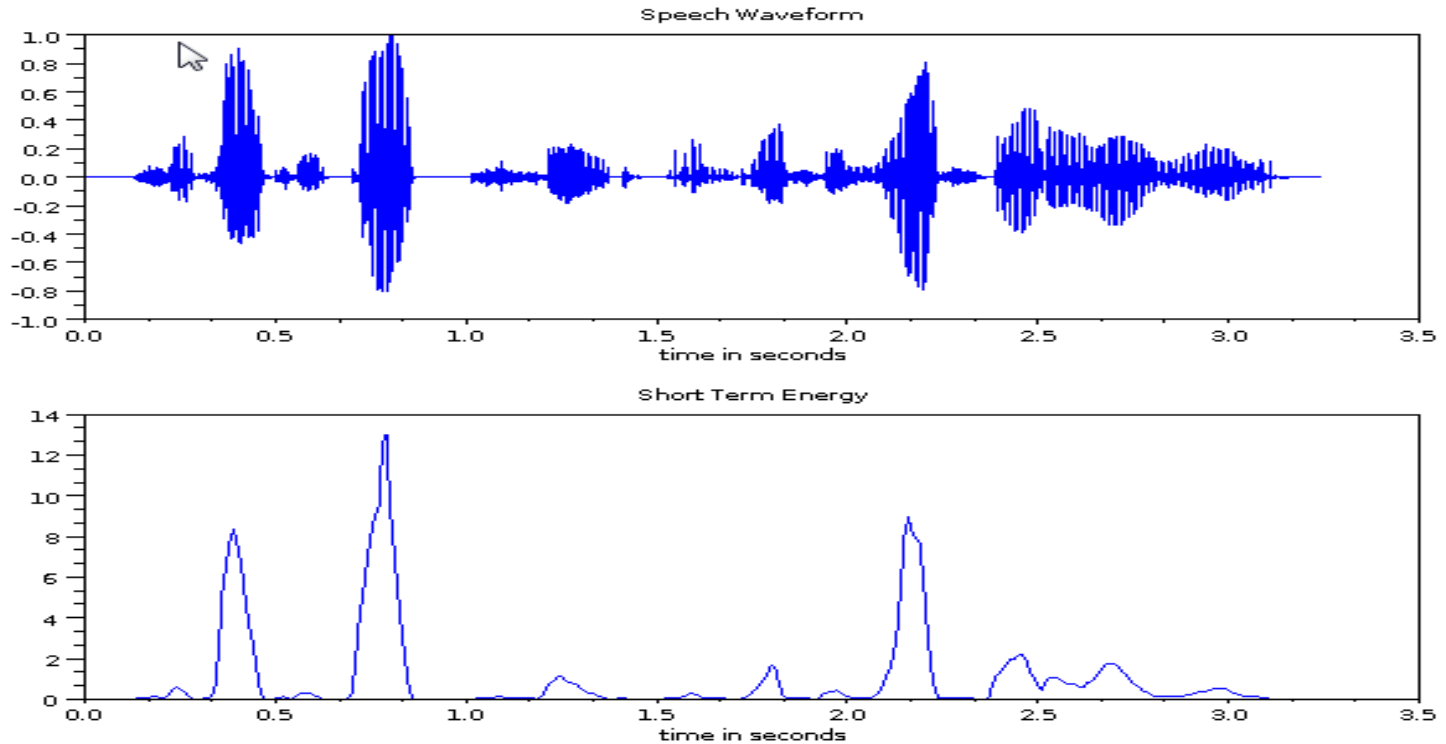
Workflow



MFCC Features



Speech Detection



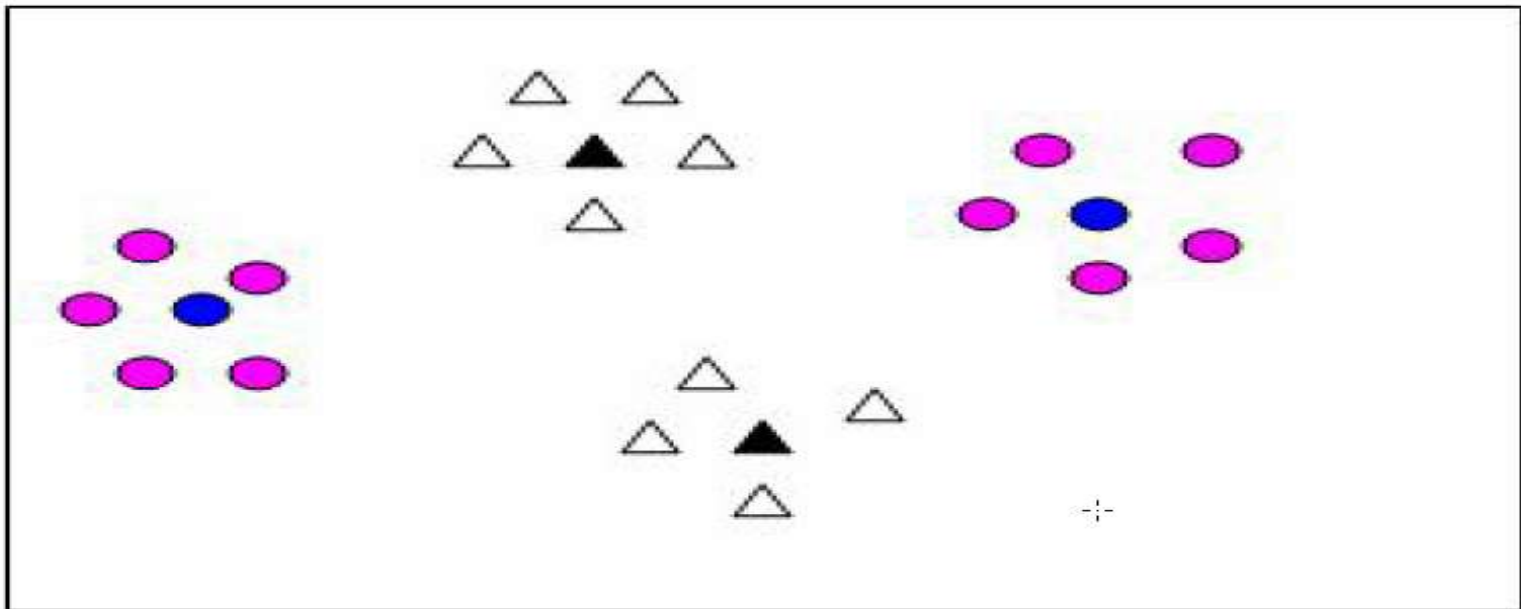
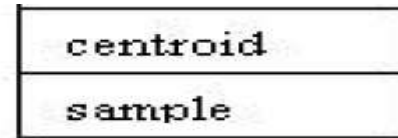
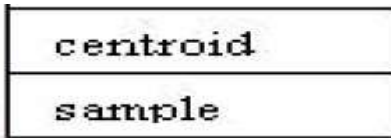
Figure_1: Short term energy contour for the spech signal



Vector Quantization

Call Center Executive

Customer



Emotion Recognition of Call Center Conversations

Results - Comparison of Speech Diarization of 2 Speakers with Manual Annotation

Segment	Time1	Time2	Result	Actual	Segment	Time1	Time2	Result	Actual
1	0.0	0.5	1	1	26	12.5	13.0	1	1
2	0.5	1.0	1	1	27	13.0	13.5	1	1
3	1.0	1.5	1	1	28	13.5	14.0	1	1
4	1.5	2.0	1	1	29	14.0	14.5	1	1
5	2.0	2.5	1	1	30	14.5	15.0	1	1
6	2.5	3.0	2	2	31	15.0	15.5	2	2
7	3.0	3.5	2	2	32	15.5	16.0	1	2
8	3.5	4.0	2	2	33	16.0	16.5	2	2
9	4.0	4.5	2	2	34	16.5	17.0	2	2
10	4.5	5.0	2	2	35	17.0	17.5	2	2
11	5.0	5.5	2	2	36	17.5	18.0	1	1
12	5.5	6.0	2	2	37	18.0	18.5	1	1
13	6.0	6.5	2	2	38	18.5	19.0	1	1
14	6.5	7.0	1	2	39	19.0	19.5	1	1
15	7.0	7.5	2	2	40	19.5	20.0	1	1
16	7.5	8.0	1	2	41	20.0	20.5	1	1
17	8.0	8.5	2	2	42	20.5	21.0	1	1
18	8.5	9.0	1	2	43	21.0	21.5	1	1
19	9.0	9.5	2	2	44	21.5	22.0	1	1
20	9.5	10.0	2	2	45	22.0	22.5	2	2
21	10.0	10.5	2	2	46	22.5	23.0	1	2
22	10.5	11.0	1	1	47	23.0	23.5	1	1
23	11.0	11.5	1	1	48	23.5	24.0	1	1
24	11.5	12.0	1	1	49	24.0	24.5	1	1
25	12.0	12.5	2	1	50	24.5	25.0	1	1

The audio file is also manually diarized to check for accuracy.

The number of correctly labeled segments are 44.

Therefore, Accuracy = (Number of correctly labelled segments)/(Total number of segments) =

$$44/50 = 88 \%$$



BOSCH

Results – Emotion Recognition

→ Confusion Matrix with 83% overall recognition rate for Berlin Emotion Speech Database

	Recognized Emotion						
Labeled Emotion	A	E	F	L	N	T	W
A	68.5	12.7	2.6	1.8	2.7	8.4	3.3
E	12.8	84.7	2.1	0.3	0.0	0.1	0.0
F	1.8	0.1	95.4	0.2	2.0	0.4	0.1
L	6.3	6.7	6.3	73.5	6.1	0.9	0.2
N	10.1	11.8	7.9	1.2	68.0	0.5	0.5
T	10.4	0.9	1.0	0.1	1.9	79.6	6.1
W	5.9	10.1	2.8	2.1	2.2	1.8	75.1



Recognition Rates for Real Time Data

Distribution of tags in data: percentages of emotions in training data

Utterances	All		English		Non-English	
	#	%	#	%	#	%
Neutral	8178	53.82%	6614	53.57%	1536	51.46%
Anger	2877	19.09%	2385	19.32%	505	16.92%
Sympathy	4089	33.04%	4058	32.87	944	31.62%
Turns	15144		13057	86.21%	2985	19.71%

Confusion matrix after improved annotation

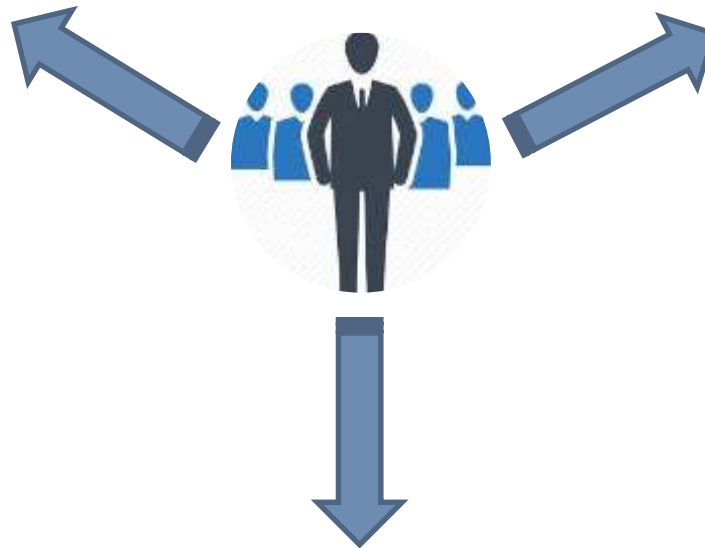
Emotion	n	a	s
n	61.26%	20.34%	18.39%
a	19.57%	70.75%	9.67%
s	18.21%	2.73%	79.05%



Benefits

Evaluate Product
Advertisement based on
the Success Ratio of Calls

Grade Customer Care
Executives based on the
Historical/Recorded
Data



Daily/Monthly Automatic evaluation of Quality of Calls

Future Scope of Work



Detect Forgery of Emotions from Phone Call

Introduce Algorithms for Real Time Processing

E-Learning for detecting the state of learner



Emotion Recognition of Call Center Conversations



BOSCH