ICS261: SSAD & Project

Team number	39
Project Title	EmoViz
Document	SSAD Project Concept Document
Creation date	24th August, 2014
Created By	Lasya Venneti, Sai Vamshi, Shubham Rathi, Vivek Ghaisas
Client	AksharSpeech (Contact point: Pranay Pratik)

Description

Our project aims to develop call center enterprise software application that helps analyse the caller's spontaneous reaction (angry, upset, dull) based on speech emotion recognition. Our product is a speech analytics engine that allows call centers to analyze real time audio from calls and educate the customer care representative about the instinctive reaction of the callee. We are constructing an interactive graphical representation of the audio data with the emotional analysis superposed on it, periodically communicating with a server to collect the audio data which generates some basic statistics and sends them to the server. Our product is easy to use, with a friendly user interface and simple integration to existing call center platforms. The UI includes emospeech representation of the audio, ability to span and play the audio and some other features which the client is yet to finalise.

Profile of Users

The end users for this tool would be call center employees who would be using this tool to differentiate the speech between two basic emotions (for now), Angry and Cool, as the employee might not be capable of handling hot headed customers. In that case, the call can be elevated to a higher official. The employee will be able to easily understand how the customer is reacting to the call and do the necessary. Features such as listening to specific portions of the call again, playing and pausing will be included for

the convenience of the call center's employees.

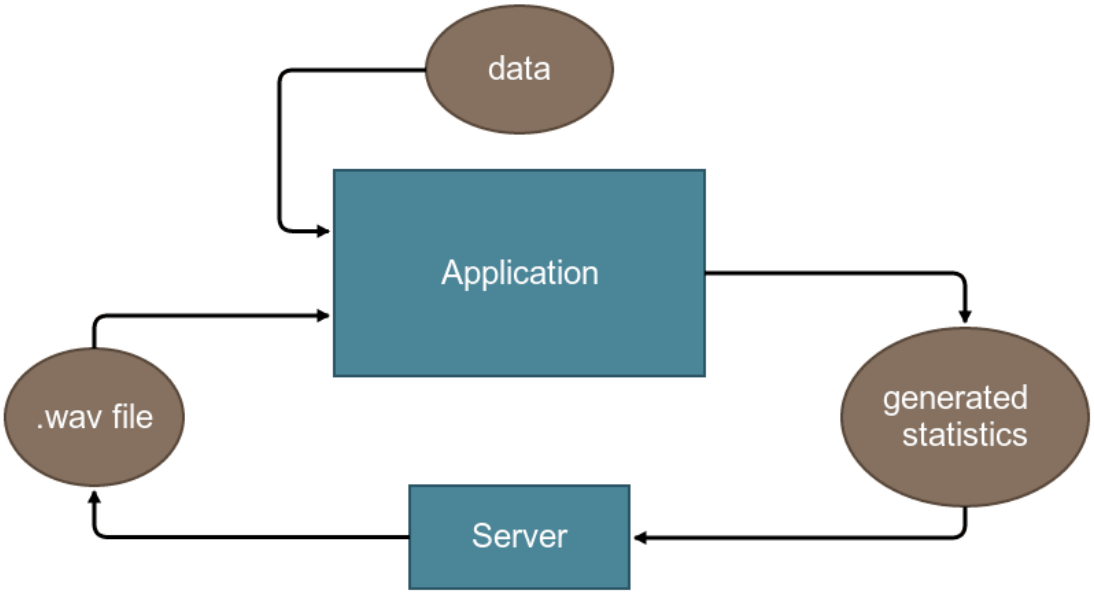
Usage Model and Diagrams (if any)

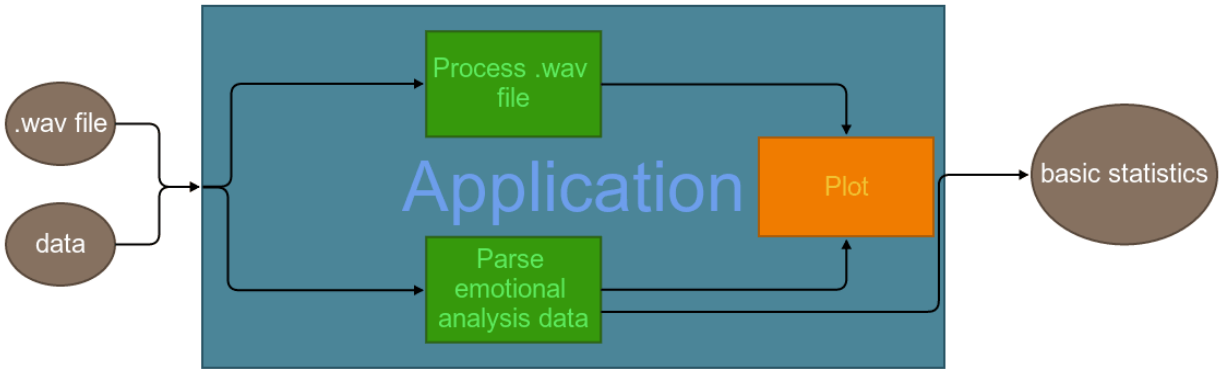
Essentially, there are 3 components in the application ecosystem, one of which we have to develop one (for now).

1. **Call centre sever (Input source 1)**
   The call centre's server generates fixed time WAV files periodically.

2. **AksharSpeech back-end (Input source 2)**
   AksharSpeech's back-end communicates with the server and runs a emospeech analysis on the WAV files. It generates data about what emotion was perceived during what time periods and passes it to the third component.

3. **UI combining both input sources**
   This is a cross platform (Windows, Linux, Mac) desktop application that provides an interactive visualisation of both input sources.




.wav file
data
Process .wav file
Application
Parse emotional analysis data
Plot
basic statistics