Customer Satisfaction Based on Voice Signals

by

Siyamamkela Bomela

Thesis presented in fulfillment of the requirements for the degree of Honours in Computer Science at the University of the Western Cape

Supervisor: Reg Dodds
Co-supervisor: Mehrdad Ghaziasgar

November 2017

Contents

Declaration 4
Abstract 5

1 User Requirements Document 6
1.1 Introduction 6
1.2 Users' view of the problem 6
1.3 Description of the problem 7
1.4 Softwares' Limitations 7

2 Requirements Analysis Document 8
2.1 Introduction 8
2.2 Designers interpretation of the users requirements 8
2.3 Related Work 9
2.3.1 First Solution 9
2.3.2 Second Solution 10
2.3.3 Third Solution 10
2.3.4 Best Solution 10
2.4 Testing 11

3 Conclusion **12**

References **13**

Declaration

I,**Siyamamkela Bomela**, declare that this thesis *"Customer Satisfaction Based on Voice Signals"* is my own work, that it has not been submitted before for any degree or assessment at any other university, and that all the sources I have used or quoted have been indicated and acknowledged by means of complete references.

Signature: ..... Date: .....

Abstract

This paper focuses on the requirements that should be implemented for the above mentioned system, the requirements recorded on the first chapter are solely based on the users' perspective. The user will be expected to clearly define what problems' they have been facing that the system will be expected to solve, This will help to clearly outline what the systems capabilities will be, and what its limitations should be.

The first chapter will not require any formal knowledge of IT/Computer Science as it is meant for the users. The programmers point of view and expectations for the system will be disregarded. The implementation and design process will not be discussed.

The second chapter focuses on the programmers view. The programmer will be expected to analyze the requirements the user had in the first chapter. The programmer will define the problem in their own view. After this process the programmer will be expected to come up with a solution for this problem. After this process the programmer will discuss related work then finally come up with an implementation plan for creating the system.

The purpose of this document is to assist in the software engineering phase, It will assist the developers and designers to know what is required by the user.

Chapter 1

User Requirements Document

1.1 Introduction

The following are ideas of how the user wants the system to function, what its expected outcomes are. After the users view a short and long description of the problem will be given and the expected limitation of the software.

1.2 Users' view of the problem

Emotion recognition is the ability to tell how a person is feeling based on the response on their face or from the tone of their voice, The aim of the project is to determine the satisfaction of a customer, whether the customer was happy or not with the service that he was offered. The customer satisfaction can be classified into three different categories; positive(which could be happy, or excited), neutral(No emotion detected) or negative(sad, angry, dissapointed).

The user requires a system that will be viable for all customers, whether a sale is done in the store or telephonically. The system should be able to determine what the overall emotions the customer went through during the sale.

The will also be a matter of accuracy of the results, as people have different signals of expressing their emotions. The system will have to be able to accommodate any user.

1.3 Description of the problem

The user require a system that will be able to determine whether s customer is happy with the service they have obtained. Numerous systems have been implemented for such a service, but these systems are mostly based on facial expressions. They use a camera to look at the changes in the customers face to determine emotion. The limitations of such systems are that the person has to be in the same room as the camera to determine the satisfaction.

Organizations such as call centers rely on voice to voice interaction to provide services or make sales, which means the system of facial expression cannot be used. The main problem that this paper aims to to tackle is to create a system that will be able to work for both kind of scenarios mentioned above, the system should accommodate sales made in store and sales that are made through call center sales agents.

The proposed system should be able to determine how satisfied a customer by a service. This will be done by observing the changes that the customer experiences as they interact with the sales person and how their emotions improve as they buy the product .

1.4 Softwares' Limitations

The software can only be implemented on one customer per time. The system will only be used in an environment with minimum people and should be in environment where there is no noise and no busy people moving around constantly.

Chapter 2

Requirements Analysis Document

2.1 Introduction

The Requirements analysis document focuses on the designers point of view. This document will look at how the system is expected to function, what tools will be required for the implementation of the system and what factors should be considered in order to ensure the system functions fully.

The designer will also be expected to analyze the user requirements document in order to determine if any data is required, data that will assist in improving the functionality of the system or data that is required in order for the system to be usable.

2.2 Designers interpretation of the users requirements

Given the above user requirements, the system created will have to determine emotions through the customers voice, When a customer makes a sale the voice will be analyzed to determine their emotion. This system will be able to accommodate the requirements the user had.

The user will have to be in an environment with minimum noise in order for the system to clearly hear their voice, A more detailed

approach of the project can be seen as the steps given below:

- A high quality microphone: This will be used to get high quality audio samples from the user to be used when determining their emotions
- A pre-existing database of emotion recognition audio: This will be used for comparing the customers sample audio to that similar in the database
- Python programming language: The system will be built using this language

These are the basics needed in order to implement the project. The system will be designed to work on computers.

2.3 Related Work

2.3.1 First Solution

A paper written by Te Won Lee, Jiucang Hao, Kwokleung Chan and Oh-Wook Kwon from the university of Carlifonia in San Diego, which also focused on Emotion Recognition by Speech Signals was implemented successfully with a high accuracy.

While designing the system, they noticed that the most important aspect that improved the results was the pitch and energy that the user had while using the microphone. They used a hybrid o Quadratic Discriminant analysis algorithm, Support vector machines and MFCC as the base feauture. Their chosen databases was the SUSAS database and the AIBO database. The SUSAS database is text-independent, while the AIBO database is speaker independent.

The SUSAS database when tested came back with an accuracy of 96.3 percent for stressed and neutral emotions, whilst the AIBO database provided a 42.3 percent accuracy. (Kwon, Hao, Chan & Lee, 2003)

2.3.2 Second Solution

The Indian Institute of Technology Kharagpur wrote a paper on Emotion Recognition through speech. They used two different databases, one was acted emotion speech and the second one was real emotion speech samples. The project used the Guassian Mixture Model. The average classification accuracy obtained was 95 percent. (Rao et al., 2003)

2.3.3 Third Solution

The Hewlett-Packard laboratories wrote a paper on a system called Recognition of Emotions in Interactive voice system. The system was designed to determine when a customer gets annoyed or angry by the Interactive Voice Response system, it would transfer the customer to a real human call center agent.

HP used two databases, one was obtained from actors which imagined a scenario and acted it out. The second database used was obtained from the linguistic department in the University of Pennsylvania.

They used Support Vector Machines and achieved an overall 77 percent accuracy. (Yacoub, Simske, Lin & Burns , 1990)

2.3.4 Best Solution

All the solutions provided excellent results, most of the excellence of the results mostly depended on which database they use and the method they chose to use. From all these solutions it is quite clear that in order to get great accuracy, one should use a hybrid for the classifiers and obtain a reliable database.

The best solution is the first solution. Their system was implemented the following way:

1. Feature Extraction: They selected the log energies, pitch and MFCCs as their base features, then selected their frame shift to be 10 miliseconds.
2. Feature Selection: Identifies features and properties of speech that are important for distinguishing different emotions.

3. Classification: In this stage they used their chosen algorithms to process the audio to obtain the emotion outcome

(Kwon, Hao, Chan & Lee, 2003)

The reason for choosing this as the best solution is because it is more inline with the system the user requires and this system was tested for different emotions.

2.4 Testing

To determine whether the system is functioning properly, the following can be looked at

- Test i the system recognizes audio when the customer speaks
- Test if it can differentiate between angry, sad or neutral
- Test if it gives an output.

These are the basics tests that the system should pass to be seen as functional.

Chapter 3

Conclusion

The Customer Satisfaction Based on Voice Signals system appears to require some pre-existing data before it can be implemented. It will be a challenge to complete as it has to be tailored for different customers with different voice tones and different accents. Although this seems difficult it is can be achieved within the time frame given

Based on the steps given ahead the system will require a lot of research before it can be done and the correct database will have to be found, but besides this the project can be developed to a functioning system.

References

[1] Oh-Wook Kwon, Kwokleung Chan, Jiucang Hao, Te-Won Lee, Emotion Recognition by Speech Signals ,Institute for Neural Computation University of California, San Diego, USA

[2] Sherif Yacoub, Steve Simske, Xiaofan Lin, John Burns , Recognition of Emotions in Interactive Voice Response Systems , HP Laboratories Palo Alto

[3] K. Sreenivasa Rao, Tummala Pavan Kumar, Kusam Anusha, Bathina Leela, Ingilela Bhavana and Singavarapu V.S.K. Gowtham, Emotion Recognition from Speech, School of Information Technology, Indian Institute of Technology Kharagpur