PRODUCT REQUIREMENT DOCUMENT

**PRODUCT REQUIREMENT DOCUMENT (PRD)**

Project name: Speech Analysis

Version: PRD 0.1

Author: Group 2

Date accepted: 27 March 2020

Launch date: 8 April 2020

Project owner: Kaushal Dabhi, Ameya Mote, Nanma Joseph, Yash Shetty

Priority:

## PURPOSE

To analyse speech for :

* Detect Speech vs silence
* Detecting male and female speakers
* To present analytic data for the android group to display

## DETAILED DESCRIPTION OF THE PRODUCT

Speech recognition being a non-biometric technology allows machine to understand the words the speaker pronounces. The core part of speech recognition is converting audio files into a 2X2 array . Speech thus being a defined as most natural ,intuitive means for human communication which not only express the thoughts and feelings but also articulates sound gestures for communication purposes . There are different forms of speech that can be categorised into various Dialects, Accents .More advanced level of speech signal level exists in the form of: Amplitude variation, Duration, Pitch, Timbre.One area where this technology can be used is customer or client experience.

# USE CASES :

1: Voice activity detection : here the segments in the waveform of audio where speech is present are recognized whereas the silent segments and non-speech are neglected .

2: Enhancing the speech :increasing the speech quality by removing the noise factor from the extracted segments .

3: Recognizing the speech: transforming the speech signal into text form.

4: Now from the text extract the natural speech making the speech sound more natural with emotions .

5: recognizing the speaker(speaker diarization ) : process of segmentation of speech signals into chunks belonging to different speakers and identifying who is speaking at a particular time .

6: separating audio source : this separates mixed or overlapped speech from noise.

7:modifying the speech: this is done by changing the emotion, tone and thus converting it to speech spoken by speaker

8:classification of emotional speech : recognizing and identifying different emotions like happy, sad,angry and anxiety

9: spotting the keyword: detecting specific keywords in entire speech utterance

## CONSUMER EXPERIENCE:

Fintech companies are well known for disrupting the banking industry and forever changing the way it operates. Now these organisations must embrace technologies like Artificial Intelligence (AI) and voice recognition to transform the way they communicate with their customers.

When consumers ring businesses it is more often than not because they have an urgent matter to resolve, particularly when they call their bank. However, people are often faced with lengthy hold times, exasperating hold music and a frustrating series of questions before they are even able to speak to a call centre agent. This only increases the customer’s sense of frustration.

These frustrations could also lead to issues being magnified more than needed. For example, Sainsbury’s Bank hit the headlines last summer when its customers faced hold times of several hours when they called to discuss why there had been a delay to their credit card statements being sent out. This led to consumers venting their anger on social media and publications like MoneySavingExpert covering the story. If these phone calls were dealt with swiftly and effectively, the bank would have been able to reassure their customers in a timely manner, and the problem would not have been amplified as much.But through the use of technology, fintech businesses and banks can make such issues a thing of the past. Artificial Intelligence and natural language Machine Learning (ML) are on the verge of transforming the fabric of voice-driven customer experience through pioneering speech recognition technology.Voice-driven customer service can complement staff and improve customer satisfaction. We’re now at a stage where the technology can understand speaking habits, conversational linguistics, dialects, idiosyncrasies, slang, foreign nationals’ accents, intonation, emphasis, intention and enunciation. It can even understand sentiment and recognises intent.

This means that the technology can be deployed in a number of ways to ensure that customer calls are dealt with quickly and effectively.

**EXAMPLE:**

For example, if a customer is clearly angry or a more human touch is needed, the tech solution can escalate the issue in real-time so that a human call centre agent can step in and take charge. By deploying such technology, the customer experience can be improved, and staff members can be freed up to deal with the more serious or complex issues.

Leading voice technologies, in essence, will enable businesses to automatically identify, understand and service their clients. It will help them improve the customer experience and reduce costs. FinTech companies have helped shape the future of banking, now they must embrace the future of customer communication.

# FEATURES:

These are the initial capabilities needed by the company’s solution to address the use cases and which are covered by the requirements of this document. The features described here are intended to give a high level view only.

### EXAMPLE:

## AUDIO RECOGNITION SERVICE:

This system will primarily be an audio recognition system for the purpose of identifying various video contents that can be viewed by the general public. This does not rule out the opportunity of using or including our music recognition service, as there are use cases where customers want to know both the video and music an end-user might be experiencing.

A recognition system based on video fingerprinting is not ruled out, but is beyond the scope of this document.

### REAL TIME RECOGNITION:

The Recognition Service needs to make searchable the reference fingerprints in near real-time, so that an application can sample the audio of a TV broadcast and successfully identify the program being watched.

### PRIVATE CONTENT :

Some content submitted to the Service is expected to be private or proprietary for a specific customer. The Service should not include this in the generally available database. Customers querying private content may want to access only their private content. They may also want to access the general Service.

### GENERALLY AVAILABLE CONTENT :

Most of the reference content placed online is expected to be part of the generally available solution. Unless specifically called out as private, reference content such as TV broadcasts, music, or Company’s video collection is expected to be generally available to customers.

### SCALE TO MILLIONS OF END-USERS:

We should expect our customers to make this Service available to millions of users. This could be smartphones, tablets, cable subscribers, TVs, and Set-Top-Boxes.

## BACKEND FUNCTIONALITY :

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Current Status** | **Phase** | **Description** |
| **Adding and managing Dataset** | Exists, | Phase 1 | -new datasets so the Machine can learn more. |
| **Adding and managing Code** | Exists, | Phase 1 | -new features could be added. |
| **Adding and managing Video** | Does not exist | Phase 1 | - |
| **Adding and managing GUI** | Exist, In rework in new backend | Phase 1 | -GUI to make customer experience good. |
| **Managing users** | Exist, should be reworked in new backend | Phase 1 | Blocking, enabling users, changing passwords, e-mails, usernames. |
| **Function for creating video playlists** | Doesn’t exist | Phase 2 | - |

## FRONTEND :

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Current Status** | **Phase** | **Description** |
| **Free Live streaming** | Doesn’t exist | Phase 1 | - |
| **Product Display** | Exist, needs design | Phase 1 | Simple display initially with future versions planned. |
| **Gallery Display** | Doesn’t exist. | Phase 1 | - |
| **Video Display** | Doesn’t exist | Phase 1 | - |
| **Poll Display** | Doesn’t exist | Phase 1 | - |
| **Carousel (For home page)** | Exist, needs design, may need fine tuning | Phase 1 | Carousel for the home page and the subpages, if applicable for a website. |
| **Hot or Not** | Exist, needs design | Phase 2 | .A customer review of any future feature planned. |
| **Personalization of content** | Doesn’t exist | Phase 2 | The ability to have the GUI in different langsuages. |
| **Facebook Login** | Doesn’t exist | Phase 1 | The ability for the user to login in our sites with his/her Facebook account. |
| **Google+ Login** | Doesn’t exist | Phase 1 | - |
| **Own Comments** | Exist, need design | Phase 1 | The ability for the users to comment on content. |
| **TV Schedule Listing** | Doesn’t exist | Phase 1 | - |
| **Function for creating video playlists** | Doesn’t exist | Phase 2 | -. |
| **Login Access** | Exist partially | Phase 1 | Login allows access for users to comment under articles/videos and to upload content. |

## Phases of Development

|  |  |
| --- | --- |
| **Phase** | **Tasks** |
| **Phase 1** | * Creation of Dataset. * Preprocessing the dataset * Statistical Analysis |
| **Phase 2** | * Functioning ML Code * Training and testing on the data * Development of GUI. |
| **Phase 3** | * To be detailed on later stage. |

## OTHER DETAILS AND DEVELOPMENTS :

1. Login (with facebook,google+) will allow the user to comment and to upload pictures/videos (as well as participate in contests).
2. Accents on homepage will be automated with the TV Schedule (where the editors can pick accents).
3. Face 2 Face polls should be automatically exported
4. To implement video playlist feature with refresh on clip change.

## VISUAL PRESENTATION :

To be available on later stage

## DETAILED DESCRIPTION OF NEW FEATURES:

**Facebook / Google+ Login** – the feature will allow the users to login with their Facebook or Google+ account. They will not need to register in our web sites, but can easily become registered with few clicks.

**Facebook / Google+ Comments** – the feature will allow the users who are not registered in our web site to comment. This will improve social media and search visibility.

**PERSONALIZATION**

The ability of the users to choose which content to display. It will be based on profile information, when the user is logged in or will be set through cookie, when not logged. Personalization would include: display of articles by selected by the user sections, only the top level sections will be included. We will display the latest articles from the selected categories, but not more than two per one category. The user must select 3 sections or more at least per site.

**Live Stream** – will require login. Login will be on dedicated page.