SUBMITTED BY :

AMEYA MOTE

NANMA JOSEPH

KAUSHAL DABHI

YASH SHETTY

MARKET REQUIREMENT DOCUMENT

**MАRКЕТ REQUIREMENT DOCUMENT (MRD)**

Approved Product Name: Speech Analysis

Version: MRD 0.1

Author: Nanma Joseph , Yash Shetty , Kaushal Dabhi , Ameya Mote

Date Initiated: 23rd March 2020

Last Updated:

**Table of Contents :**

**1. EXECUTIVE SUMMARY :**

The project is literally focused to analyse on the various forms of speech based on different areas like gender recognition from the speech such as female and male speakers, detecting speech and silence and then providing the analytical data for the android group to display.

**Gender recognition –detecting male / female speakers:**

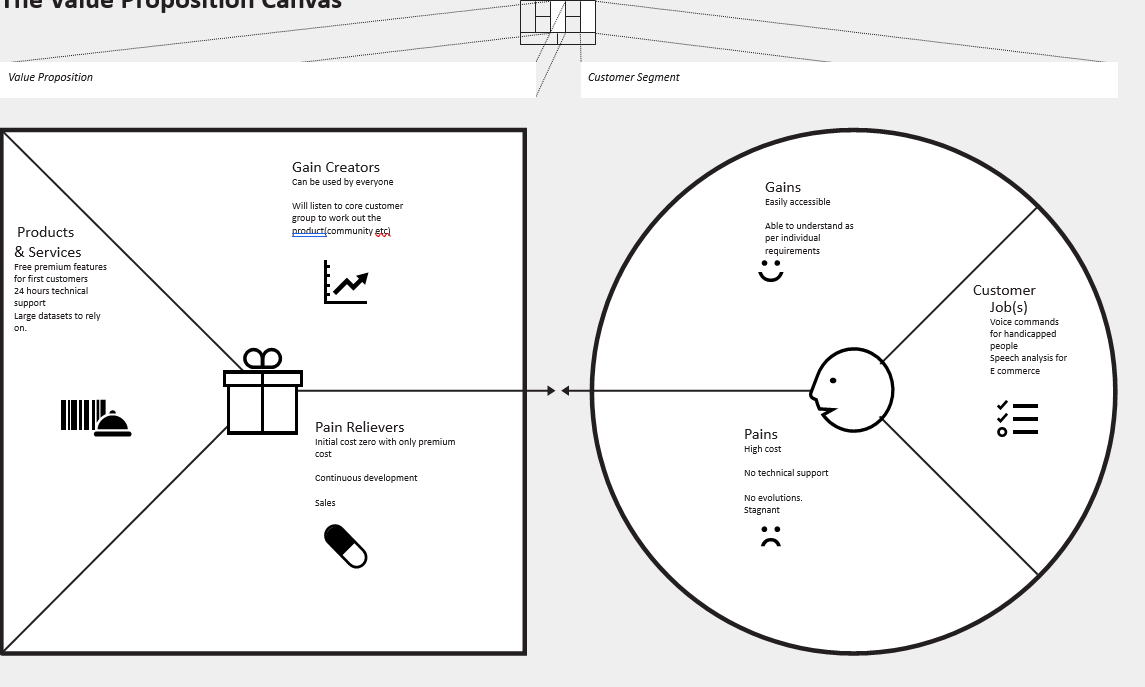
Gender classification using speech analysis literally predicts the speaker gender by analysing various parameters of voice sample. The speech signal carries both the information to the communicator and also contents the information regarding the speaker.The non-linguistic characteristic of the speaker provides help to differentiate male and female speaker. The different features such as power spectrum density, frequency determined at maximum power carries the speaker details . The features of the speaker could be tracked with frequency characteristic of vocal tract and also with the variation in excitation. gender recognition can be classified based on two sub categories such as gender identification and gender verification. With identification(1:N matching) a particular speaker is compared with database of N known speakers and the best match is returned as the decision for the recognition Whereas the verification task involves firstly making the decision whether the  given sample of voice is produced from the claimed speaker. Here when the unknown speaker voice sample is compared to the claimed speaker and if the similarity degree exceeds a particular threshold frequency then the speaker is accepted else the speaker is rejected.

**Detecting silence from speech:**

Silence being a non perceptual audio signal , the energy level contained in the silence is comparatively low . Therefore an energy thresholding could extract the segments of silence. There also exist different audio segments that can be classified as silence if the energy level is used alone, like low energy music . There exists a ZCR(zero crossing rate) for silence which is comparatively low when compared to other audio types . Hence by comparing the energy and ZCR, we can predict the improvement in accuracy of detecting silence .

**2. SOLVING THE PROBLEM :**

**2.1 VALUE PROPOSITION CANVAS :**



**2.4. USE CASES:**

* To eliminate incidences of customer fraud and non-agent compliance :

The speech analytics have the capability to analyse different call recordings inorder to identify compliance issues such as warn the agents , redirecting the call to the manager for corrective action since knowing every call is getting listened ,checked for privacy policies , bestselling practises so as to maintain reputation and financial issues that could rise from any areas .

* To achieve insights into customers and competitors:

When talking with the agents the customer gives a lot of information like about the usage of the products , comparing the prices, the efficiency of sales and this is how to stack up against the  competition .

* To generate more sales :

Understanding the factors that drove people to buy the product .

* Improving customer journey and overall experience
* Driving employee engagement and coaching.
* Recognizing and correcting the processes.
* Decreasing customer effort.
* Reducing the effort of employees.
* Agent monitoring.
* Monitoring the quality.

**2.5. HIGH-LEVEL STORY : CUSTOMER/ CONSUMER EXPERIENCE:**

Speech analysis is used mostly in telephone applications such as travel booking and information, financial account information, customer service call routing . For future ,automatic speech recognition proves to be an important factor in m-business and m- government applications . Here a 4 layered study design was created inorder to understand the consumer expectations and views :

1. **Testing of a service prototype:**

Here the stage provided three purposes. Firstly it had helped in order to test the basic service functionality i.e operations involved in speech technology .The respondents were requested to try out a service prototype in order for booking the appointment with local health care which includes different speech recognition elements . Second it had been served as an input for direct feedback : respondents were told how was their impression on the service during the phone call and also their independent responses had also been recorded. Third it helped to improve user experience thus enabled users to act accordingly with the technology and its applications .

**(2) Conducting questionnaire survey:**

Here the users were asked to fill in an email questionnaire based on the service provided. The questions included how ease was the service of speech recognition system,Sustainability of automatic speech recognition system and even the other feedbacks which they would like to give .

(**3) Focusing on the group discussions:**

Here they were to gain qualitative data on the acceptability of service , take user suggestions for further improvements and how they can use these suggestions for future implementations .

(**4)** **Conducting interviews and meeting :**

With developers of the product which was included with formal group interview for finalising and the final evaluation of the product development.

This survey was important so that he company would be able to obtain brief data on how well the product was useful for different users . the group discussions were done to study the consumer view points were the users had an opportunity to come up with their questions and comments and thus provide it to the service developers

**3.**  **BUSINESS OPPURTUNITY AND JUSTIFICATION :**

**3.1 INDUSTRY TRENDS THAT CREATE DEMAND FOR THE PRODUCT :**

The speech along with voice recognition market is expected to grow with CAGR of 20% from 2019 to 2024 . The future of speech recognition market have great further opportunities in enterprise, consumer, banking, insurance and also even in educational industries .The major demand in the market is for speech –based biometric system have indeed a growing demand with voice authentication in mobile banking applications .

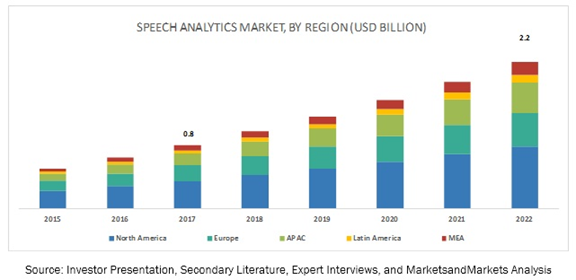
Market Segment by type:

* Speech Recognition
* Voice recognition

Market segment by application:

* Government
* Retail
* Medical
* Military
* Legal
* Education
* Other

**3.2.  SAM/TAM AND GROWTH RATE OF TARGEED MARKET :**



Link : <https://www.marketsandmarkets.com/Market-Reports/speech-analytics-market-17297779.html>

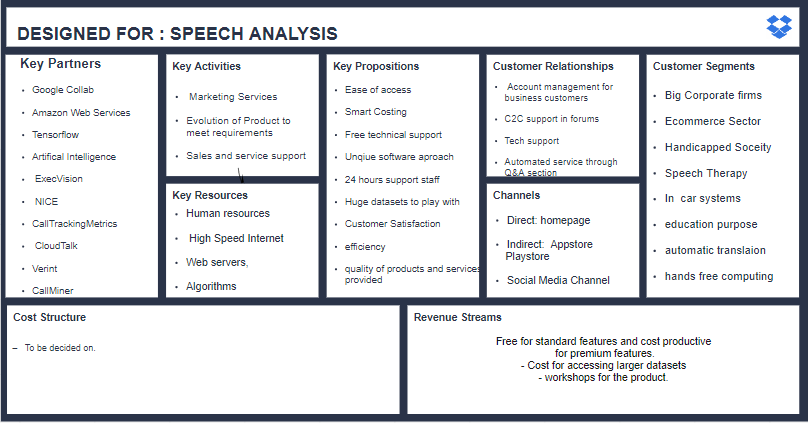
**Assumptions:**

1.  Many companies need a strategic map to showcase their capabilities and to have a linear flow of conversation across all departments. They need to adopt much more changes in and give credibility provided by speech analysis.

2. E-commerce sectors having grown so fast, are in lookout for news unique ways for customer interaction. As such Speech analysis provides them with a perfect data to plan their future goals.

 As you can see historically there has been an ever increasing trend in the speech analytics market across each country. Such factors drove our imagination in coming together and making this product.

**3.3. BUSINESS MODEL:**



**4. What we need to build**

**4.1. Key product features and benefits, in priority order, and form it needs to take.**

Audacity - Audacity is a free and open-source digital audio editor and recording application software, available for Windows, macOS, Linux, and other Unix-like operating systems

R Studio- RStudio is an integrated development environment for R, a programming language for statistical computing and graphics.

Jupyter Notebook - Project Jupyter is a nonprofit organization created to "develop open-source software, open-standards, and services for interactive computing across dozens of programming languages" .

**4.1.1. Must Haves**

Dataset- A Dataset comprising of female and male voice speaker across different ages so that the Machine can learn to predict the gender of the Speaker.

ML Programming Code – A code preferably written in Python to recognise gender should be made available on Jupyter Notebook.

Different voice analysis on which basis, graph or a comparison could be formulated for male and female gender.

**4.1.2. Nice to Haves.**

* A Code for distinguishing voice from silence would be a huge bonus.
* The number of samples could be increased so the code could be tuned more perfectly.
* GUI could be incorporated so that the user does not face difficulty in whatsoever operation he/she intends to perform.

**4.2. Customer interaction with key features.**

* The product needs to properly define the gender of the speaker and should predict it even with any voice modulation. This would corroborate to the perfectness of the code and delight the customer.
* A proper pilot opening should be done with audiences from the core customer base so as to catch their eye right from the start.

**4.3. Support for brand promise**

* Pledge to support customer support for future and past versions of the product.
* Product manual should be easy to understand.
* Experiences sales support to ensure the customer is satisfied while buying the product.

**4.4. Performance benchmark and other system requirements**

* A Portable product
* Simple operations so that it’s easy to understand.
* A high-grade CPU to withstand heavy computations that would come with the ML code.

4.5. Cross-vertical synergies and opportunities to consolidate products

So far there has been no indication of any other commercial product being available in the market as this and would certainly be the first in this genre .

**4.6. Additional requirements by country or region**

Language- The GUI interface should have the option of having different languages so that it could be used by anyone from any part of the world.

**4.7. Target pricing and costs**

Since this the first of it’s kind, there is no previous examples to base the price on. So it should be balanced so that the customers do not feel it’s too costly. As such 10 Euros is perfect for such a product at this moment of time.

4.8. Internal and external testing and interoperability requirements

Code- The code should be checked and rechecked properly for any syntatical or logical mistakes.

Dataset- The dataset should be pre-processed to remove any bias or bugs.

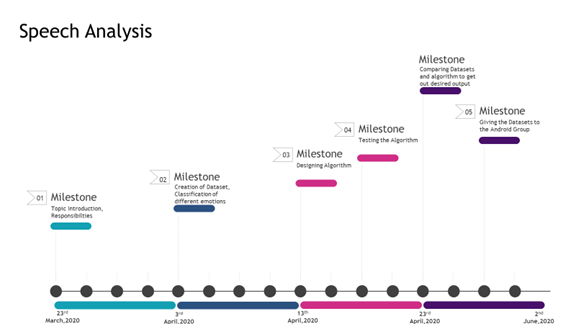
Proper identification- the customer should check if the product is performing as per specifications.

**5.** **WHEN WE NEED IT :**

**5.1 PRODUCT DELIVERY DATE , BUSINESS CYCLE DEADLINES :**

The Database needs to needs to be made ready for the next design cycle no later than June 2nd. This is done for the Android Department to have sufficient timeline to design a Graphical User interface to display the data. The Database needs to be cleaned and refined with exact labels so as to not have any bad version of the output.

**5.2 REQUESTED TIMELINE :**



**5.3 REQUESTED PRODUCT LAUNCH DATE :**

We want the group to be launched in the Holiday Season of 2020, since it’s the time of festivities and a lot of people are looking out for new things to do in the Holiday Season. So we expect, Oct 20th 2020 to be the perfect launch date.

**5.4 ANTICIPATED EVOLUTIONARY DEVELOPMENT OF THE PRODUCT :**

The initial product will just be the version 1.0 . Depending on how well the product is received by customers and their feedbacks, the product will be built on it and futures versions will be developed.

**6. DEPLOYMENT :**

**6.1: DISTRIBUTION METHOD**

Indirect method of Distribution will be used in this aspect since, the products needs to be put on platforms like Google Store, for IOS we need to put it up on AppStore.

**6.2 KEY CHANNELS :**

Since our project is a software based product, it will be released solely on the internet, initially on the Appstores, then after a certain point of time, will be available on our own website.

**6.3 EARLY PARTNERS:**

No scope of having early partners.

**6.4 DEVELOPER PROGRAMS:**

The introduction of 3rd party programmers would give us a customer based view of the product and a totally new direction not initially explored. This will also help us to approach our product in a practical way rather than emotional way.

**6.5 SALES TRAINING OUTLINE :**

- Training sales support with the whole idea of the product.

- Training the customer support department with the bugs and scope of the products.

- Android Programming specialization.

- Digital Media Marketing since totally Internet based.

- SEO Optimization training needs to be imparted.

**6.6 GO TO MARKET OUTLINE:**

* + Reach out to the core customer group.
  + Generate Target revenue for next Version development.
  + Capture the market during the Holiday Season.

**6.7 DEMONSTRATION CAPABILITIES:**

Participation in Technical Conferences to demonstrate it’s capabilities. Holding out workshops and training modules for the sales and support staff. This all will be done during the development phase of the algorithm so as to familiarize the staff with the functions and peripherals.

7 **RISK ASSESMENT (HIGH,MED,LOW,PROBABLITY) AND WHY:**

**7.1 MARKET AND COMPETITIVE RISKS :**

* Low Risk: Not able to meet the market deadline, no finished product.
* Medium Risk:Incapability of handling large loads on the server of very few datasets to rely on.
* High Risk: Less understanding of the product by the customers and as a result low sales due to less demands.

**7.2 INTERNAL CHALLENGES :**

* Less similar products in the Market,thus cannot compare enough to give out a refined product.
* Less experience in the business field.
* Young market for such products as a result of no similar products.
* Difficulty in procuring the required output due to external factors present.
* Since we are growing, No huge investment due to which machinery cost bearing difficult.
* No exact datasets to be relied on.

**8 . REVIEWERS AND APPROVERS :**

**8.1 REVIEWED BY SALES(REGIONALLY) , SELECTED CROSS-VERTICAL LEADS**

* Nanma Joseph
* Kaushal Dabhi
* Yash Shetty
* Ameya Mote

**8.2. FINAL APPROVED BY TBD :**

* Alexander Ilev

**9 APPENDICES** :

**9.1 REFERENCE DOCUMENTS:**

* <https://www.researchgate.net/publication/325554736_Speech_synthesis_systems_Disadvantages_and_limitations>
* <http://jorgenboge.wikidot.com/how-to-analyze-a-speech>
* <https://www.marketsandmarkets.com/Market-Reports/speech-analytics-market-17297779.html>
* <http://www.ub.edu/lexdialgramarxius/theapplicationofspeech.pdf>
* https://medium.com/@TEAM.BULLS/application-of-speech-recognition-technology-in-speech-related-disabilities-an-analysis-and-5b102da57f86