SERvis

Emotional Health Detection for Effective & Efficient Diagnosis

Group 29 (ME781)
Conceptual Design Report

Gerard Tony Thomas
180100044

Navjit Debnath
190010049

Pushpendra Yadav
190010073

Shashwat Trivedi
204010003

Shiven Barbare
190100110

Problem Objective

- Access to records of emotional indicators of patients can improve diagnosis and treatment strategies. Our project aims to develop a **software** that would accomplish this

- The highlight of our software is **emotion recognition** that is accomplished by analyzing user's speech with Deep Learning model

- We would partner with **healthcare professionals** for whom our product would be advantageous while also coming at very **nominal cost** since it would be incorporated into already existing systems

Problem Definition

Customer Requirement	An accurate detection of the range of emotions of an individual over any desired time interval. A potent tool for health experts to aid in diagnosis, treatment and health monitoring
Market Survey	*Emotion AI:* **Vokaturi** (software), **Good Vibrations Company** (elderly care, neural disorder detection), **Behavioral Signals** (AI mediated conversations), **MindCrafter App** (Mental Health Diagnosis & Consulting)
USP	**One-of-its-kind** arbitrary length tracking of patient's **emotional** offering **better in-depth analysis**
Protection of USP	Encryption of code and usage of secure servers to update application, filing patents and copyrighting relevant material
Barrier to Entry	Permitting and licenses required to establish production Rigid industry standards Patents and copyright issues from other startups/companies working in the field
Product & Operational Structure	Recored audio is received -> Model identifies deviations from standard conditions -> Model reports comprehensive data to health expert

Technology Landscape Assessment

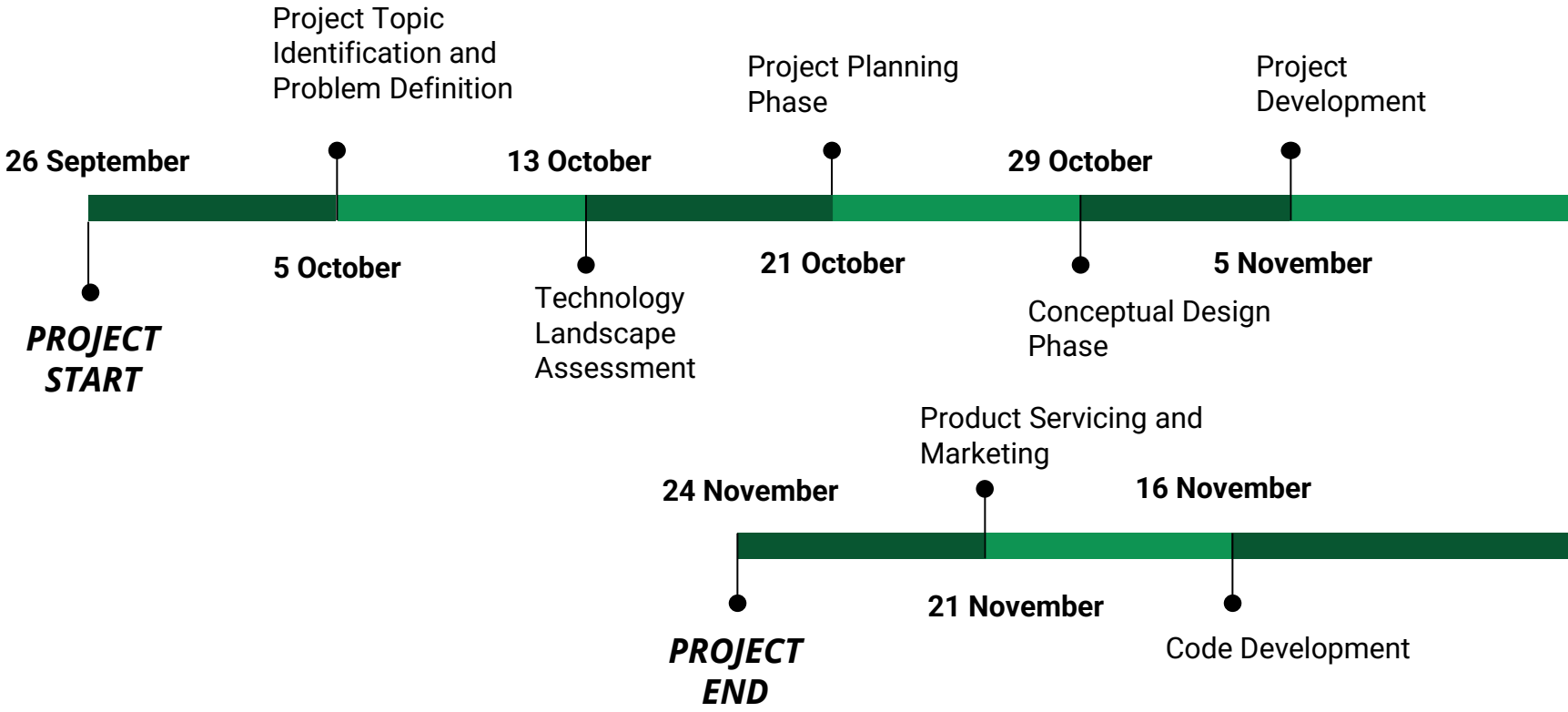
Published Literature

- Machine and Deep Learning Algorithms for Wearable Health Monitoring
- Wearable Sensors for Remote Health Monitoring
- Emotion Recognition from Speech : Tools and Challenges
- A Machine Learning Emotion Detection Platform to Support Affective Well Being
- DialogueRNN: An Attentive RNN for Emotion Detection in Conversations

Patents

- Emotion Recognition System
- Wearable Health Monitoring System
- System for monitoring health, wellness and fitness

Project Timeline



Roles and responsibilities (RASIC Chart)

Task	Gerard	Navjit	Puspendra	Shashwat	Shiven
Project Task Breakdown	S	I	I	R	I
Customer (End-user) requirements	I	I	I	R	I
Market Survey for available solutions (Similar products and services avialable in the market)	I	I	R	A	S
Key Differentiator (How are you different from the available solutions)	I	I	I	R	S
Your Unique Selling Point (USP)	R	I	I	S	C
How can you protect your USP	R	I	I	I	I
Barrier to entry both by you and others	I	I	R	I	C
Patents	I	R	I	S	A
Published literature	I	A	I	I	R
Project timeline	I	C	R	I	I
Roles and responsibilites (RASIC Chart)	I	I	I	R	I
Project Monitoring	I	R	C	I	I
ML Model selection	S	S	R	I	I
ML Dataset selection	S	I	R	S	S
High Level activity object and class diagrams (using uml)	I	I	I	R	I
Choice of language, OS, data set, algorithms and dependencies on external libraries	I	I	R	I	I
High level sequence diagram with type of user input and output	I	I	I	R	I
Variable Naming	I	I	I	I	R
Coding Style	I	I	I	I	R
Indentation style	I	I	I	I	R
Basic User interface	I	R	I	I	I
Output visualization	I	R	I	I	I
Error handling (inbuilt python function)	R	I	I	I	I
Auto document generation (Pydoc/Pycco, etc)	I	I	I	R	I
Version control using Github	I	I	I	I	R
Code testing (using unittest or another library)	R	S	I	I	I
Brochure development	I	S	I	S	R
Marketing presentation	S	I	S	R	I
1 minute marketing video	S	I	R	I	I
Demonstration presentation	R	S	I	I	S
User manual	S	R	I	S	I

R: Responsible: the person is responsible for this task
A: Approve: the person giving the approval
S: Supporting: the person giving support for completion of task
I: Inform: the person to inform about the task
C: Consulted: people who can act as expert in regard to the task
Rule: Only one A and one R possible per row

Conceptual Design

Audio Pre-processing

- MFCC, a highly researched and utilized feature is used to differentiate the gender and emotions
- 2-D matrix of Features is extracted from audio clips and converted into 1-D by taking row means and column means

Modelling

- Post feature extraction, CNN-1D is applied on 1D data frame
- Train-test data is split in the ratio of 3:1

Unit Testing

There is not a requirement for unit testing in our model since we ensure no discrepancies in audio pre-processing and we input only 2D matrix of audio extracted features in our model which can be handled smoothly by model

Conceptual Design (contd.)

Dataset

- TESS (Toronto Emotional Speech Set): 2 female speakers (young and old), **2800 audio files**, random words were spoken in 7 different emotions
- SAVEE (Surrey Audio-Visual Expressed Emotion): 4 male speakers, **480 audio files**, same sentences were spoken in 7 different emotions
- RAVDESS: **2452 audio files**, with 12 male speakers and 12 Female speakers, the lexical features (vocabulary) of the utterances are kept constant by speaking only 2 statements of equal lengths in 8 different emotions by all speakers
- CREMA-D (Crowd-Sourced Emotional Multimodal Actors Dataset): **7442 audio files**, 91 different speakers (48 male and 43 female between the ages of 20 and 74) of different races and ethnicities, different statements are spoken in 6 different emotions and 4 emotional levels (low, mid, high and unspecified)

https://www.kaggle.com/ejlok1/toronto-emotional-speech-set-tess
https://www.kaggle.com/ejlok1/surrey-audiovisual-expressed-emotion-savee
https://www.kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio
https://www.kaggle.com/ejlok1/cremad

Conceptual Design (contd.)