MINOR PROJECT-2

SYNOPSIS

ON

Emotion Detection using Facial Expressions and Bio Signals

Submitted By

Abhishek Bisht	Diksha Mehta	Janhvi Joshi	Shweta Rawat
R610217002	R110217050	R110217066	R103217106

Under the guidance of

Mr. Pankaj Badoni

Assistant Professor- SG

Department of Virtualization

Cloud Computing & Virtualization Technology

Department of Virtualization

School of Computer Science

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES

Dehradun-248007

Aug-Nov, 2019

Index

Chapter 1: Abstract

Emotion Recognition is one of the traits that humans carry naturally but even then in some cases we need computers to intervene, judge and identify the emotions of humans. This emotion detection can be used in fields like Usability testing, Lie Detection testing and others. In this study we are trying to catch the emotions of humans using their facial expressions, voice signals, and physiological signals like EEG, EMG and GSR by implementing well trained Deep Learning and Machine Learning models by conducting a deep study on the features that show major activity on emotion change. Later we aim to optimize the accuracy of the models and detect and recognize the correct emotion of the user in real-time.

Chapter 2: Introduction

Emotions can be described as a set of predictable responses that humans tend to have to various situations. These responses often have accompanying physiological signals, for example, faster heartbeat, change in skin temperature, change in pitch of voice and so on. When studied carefully, these changes can yield useful information that can be applied in various fields and domains. A major part of the study of humans includes the study of their emotions.

Hence emotion recognition, a process of recognizing human emotions, has become quite popular recently. The input taken can be from many sources - facial expressions taken from a video, voice signals from audio, and physiological signals or biosignals from various sensor devices. In this project, we will be focusing on facial expressions, voice signals from audio and biometric signals. Biometric signals include GSR, EEG, EMG amongst others.

While facial and voice signals can be manipulated to some extent, biometric signals are produced subconsciously in humans and are difficult to manipulate.

In this project, we will be following a two-dimensional emotion system as shown in the figure below. The two axes represent *valence*, describing the range of negative to positive emotions,

and *arousal*, which describes an active to passive scale. Low valence and low arousal indicate an unpleasant feeling which describes emotions like sadness and depression.

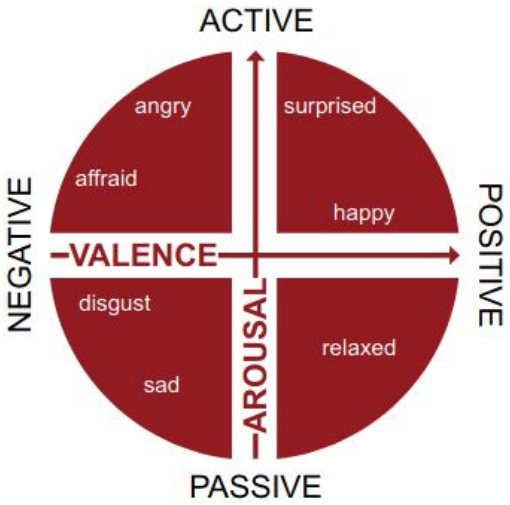


Figure 1

Following are the parameters that we will be using to assess emotions:

1. *Facial Expressions:* Each muscular movement of the face corresponds to an emotion displayed on the face. The study of the change of key features in a human's face to capture what he/she is feeling is known as Facial Coding. This process is made easier through computer-based facial expressions analysis. It provides for two-dimensional emotion measurement, that is, measurement of both arousal and valence.

   Facial recognition using a webcam with a backend of deep learning models is an easy and most accurate way of two-dimensional emotion measurement.

2. *Galvanic Skin Response (GSR):* Sweat glands are quite sensitive to our emotional state. Emotional arousal (either positive or negative) leads to an increased release of sweat in our hands. The changes in the conductance of the skin are too small to be noticed by humans, so a GSR sensor is used to detect these changes. The device is unable to classify whether a person is feeling happy or excited or sad but can tell whether the arousal is activated or deactivated.

3. *Speech and Voice:* The speech and voice of a human varies with changing emotions. Features like pitch, formant or cepstrum can be detected and extracted from this variation. Being the fastest and most natural method of communication between humans, speech has a direct relation with emotions.

4. *Electromyography (EMG):* EMG records and evaluates the electrical activity of skeletal muscles. Emotional responses lead to the activation of the sympathetic nervous system which causes contraction and expansion of muscles. While contracting, muscles create some potential which can be measured from the skin. This method requires the attachment of electrodes to the skin. It provides for one-dimensional valence measurements.

5. *Electroencephalogram (EEG):* This is the study of measuring electrical changes happening on the brain's surface. It indicates the amount of brain activity that directly corresponds to the amount of engagement of the person. Both valence and arousal measurements are possible. High frontal activation corresponds to positive emotions while negative emotions cause a right frontal activation.

For the purpose of our project, we will be using facial coding and speech for detecting two-dimensional valence and arousal measurements. For further accuracy and better results, we will also be measuring one-dimensional biosignals - GSR, EMG and EEG.

One of the many applications of real-time emotion detection using both single and two-dimensional signals is *usability*. The traditional usability testing of software involves making users complete some task and then give feedback about their experience. This often leads to inaccurate responses because we rely too much on individual and conscious answers. Biometrics provides a way to measure real-time and subconscious feelings of the users which are guaranteed to be unbiased. Other applications may include lie detection, security systems, driver fatigue monitoring and so on.

Chapter 3: Problem Statement

In recent years, emotion detection and recognition has become one of the emerging areas of study which finds its application in a wide variety of fields including Artificial Intelligence.

However, efficient systems for emotion detection and recognition are not readily available and are also hard to build. Nevertheless, more and more research is being conducted in this emerging field to apply it in real-world scenarios. It is expected that the demand for such systems will only grow in the near future. Hence, sooner or later, efficient systems for emotion detection and recognition will become a necessity.

Chapter 4: Literature

1. **Emotion Recognition from Physiological Signal Analysis [1]**: A Review: As stated here, we are following the two-dimensional approach to classifying emotions. This paper also gives an overview of the various methods that have been used in past research for emotion recognition. Features needed to extract for each biosignal has also been provided. Benefits, limitations and applications of each method has been discussed with great detail.
2. **Alone versus In-a-group: A Comparative Analysis of Facial Affect Recognition [7]**: In this paper study is conducted to know about the emotion variance in people when they are in individual settings versus in group settings. Three models have been made, one with the data of individual settings, one with the group setting and the last one with a combined dataset.
3. **OpenFace: An open source facial behavior analysis toolkit [8]**: this paper presents Open Face, an open source tool for facial behavior analysis. This tool can be used by computer vision, machine learning, and affective computing community alike. The tool is capable of detecting facial landmark, estimate head pose, recognize facial action unit and estimate eye gaze.

4. **AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups [9]:** This paper presents a dataset made in 2 social contexts i.e. when people are watching videos individually and when they watch it in a group. Here they showed 16 short and 4 longs videos and took their data consisting of GSR, EEG and EMG. Then they also classify and pre-process the data they have collected and make a dataset out of it.
5. **DEAP: A Database for Emotion Analysis using Physiological Signals [11]:** DEAP is a multimodal dataset for human affective states analysis. It consists of data gathered from 32 participants and includes physiological signals such as EEG and GSR. The stimuli for the participants were 40 one minute long music videos. The participants annotated the videos while they were played and the physiological signals were correspondingly recorded.
6. **Overcoming unavailability of data [10]:** Dataset unavailability is a major issue research as the result varies with the type of dataset we have. In [10] we found Data augmentation, a technique for increasing size of training data. This is a good alternative when we are short of data. [10] highlighted some techniques like Noise injection, Shifting time, Changing pitch, etc for increasing the size of input data.
7. **Survey on Speech Emotion Recognition- Features, Classification Schemas and Databases:** In [12] we studied local and global speech features and how they contribute in recognizing emotions in speech. Apart from both local and global features there are Cepstral features(MFCC, LPC, etc), Qualitative features like voice quality, harshness etc and Continuous features like pitch, energy, etc. We also found some good speech datasets that are available for research purposes. Also, [12] highlighted classification schemes, processing of speech signals and how we can apply Machine learning models to achieve good accuracy.

Chapter 5: Objective

1. Obtain relevant datasets for training and testing of our models to be built.
2. Build a model to recognize emotions from facial expressions using image processing techniques.
3. Build a model to recognize the emotions in a speech sample.
4. Build a model to recognize the emotions from body signal received from the various sensors such as GSR, EEG etc.
5. Compare and Contrast the results provided by our models.

Chapter 6: Methodology

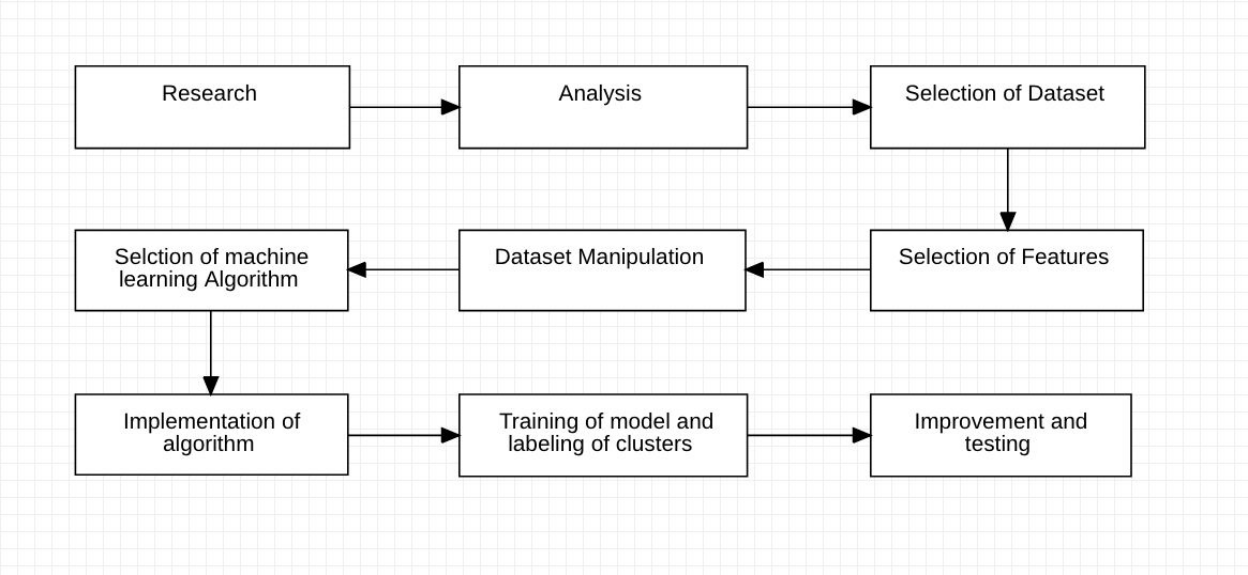
The methodology that we will follow for this project is waterfall model.

1. A deep study of the various signals that the human body emits that have a direct relation with emotions.
2. Classifying the signals into one and two dimensional. Analyzing the correlation of these signals so as to combine them to produce better results.
3. Selection of Data Sets for facial expressions and the selected biosignals.
4. Selection of features of each of the signals selected.
5. Dataset manipulations which include cleaning and preprocessing of data.
6. Selection and implementation of algorithms:
    a. For the purpose of emotion detection through facial expressions, computer vision techniques shall be used along with Deep Neural architecture with convolutional, pooling and fully connected layers for classification of emotions.
    b. The convolutional neural network shall be used to classify emotions. Silent segments need to be removed from the speech in the preprocessing step as it

doesn't contribute in recognizing the emotions followed by feature extraction. Our focus will be mainly on Cepstral features like MFCC, LPC,etc.

c. The sensor reading for the specific body signal will be extracted from the datasets along with the annotations for those signals and will be used to train the ML/DL model.

7. Training of the models and labeling of clusters
8. Improvement and testing



Chapter 7: System Requirement

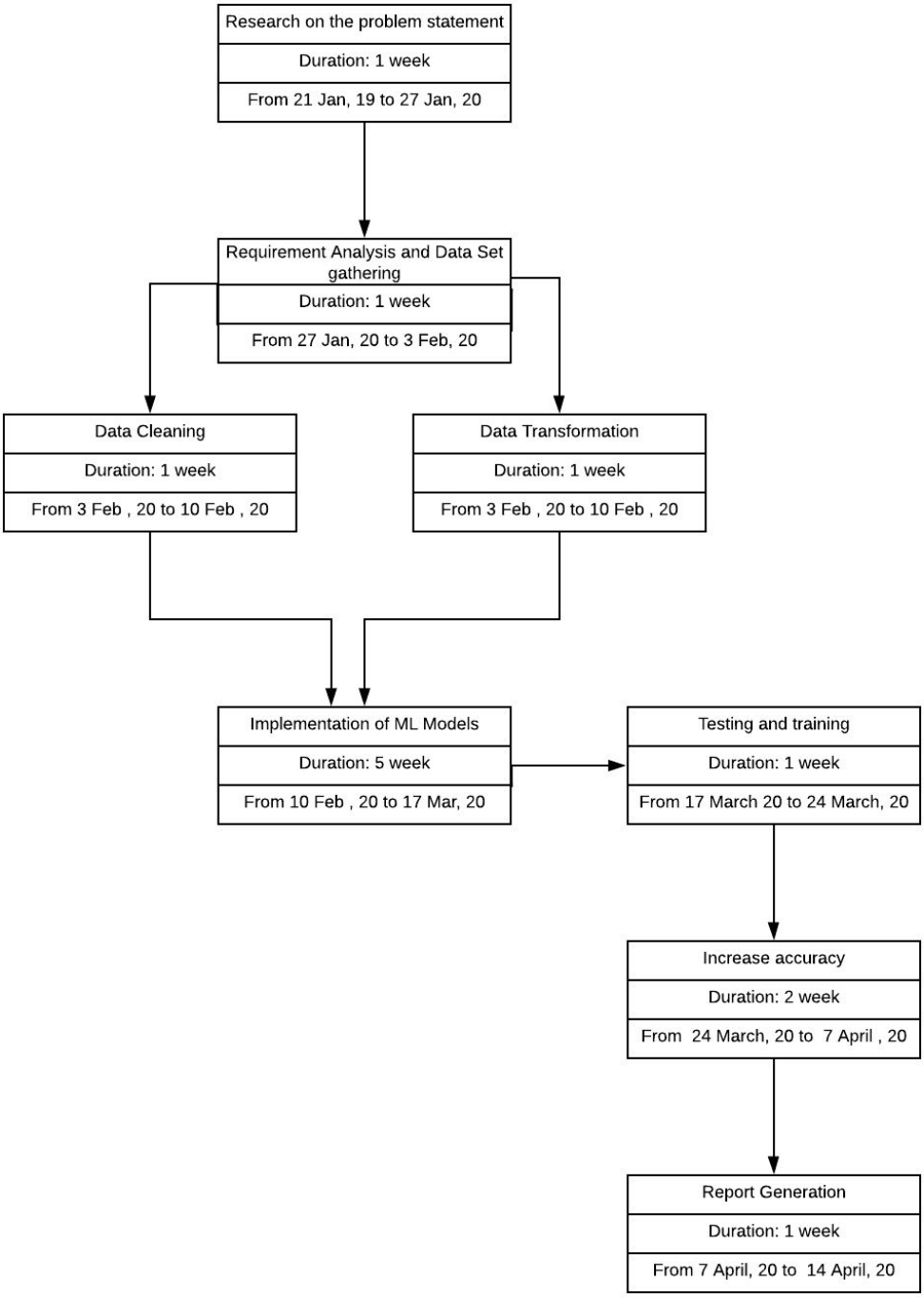
1. Software Requirements

Operating System	:	Windows 10/8/7 (32-bit or 64-bit)/ Linux
Software	:	Text Editor
Compiler	:	GCC

2. Hardware Requirements

Processor	:	Dual Core 2.7 GHz or better
RAM	:	512 MB or higher
Disk Space	:	512 MB

Chapter 8: PERT Chart



Chapter 9: References

[1]https://reader.elsevier.com/reader/sd/pii/S157106611930009X?token=951A30F8DECA01E82B1F53DDBC4AFE2DC8814A39A2153FD5B14B60F6CFA74DC634C45CD4B3672D20C1648EF9288E58F7

[2] https://imotions.com/blog/what-is-biometric-research/

[3] https://imotions.com/biosensor/fea-facial-expression-analysis/

[4] Ayadi, Moataz & Kamel, Mohamed S. & Karray, Fakhri. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition. 44. 572-587. 10.1016/j.patcog.2010.09.020.

[5] https://imotions.com/blog/facial-action-coding-system/

[6] Kerkeni, Leila & Serrestou, Youssef & Mbarki, Mohamed & Raoof, Kosai & Mahjoub, Mohamed. (2018). Speech Emotion Recognition: Methods and Cases Study. 175-182. 10.5220/0006611601750182.

[7] https://dl.acm.org/doi/10.1145/2964284.2967276

[8] https://ieeexplore.ieee.org/abstract/document/7477553

[9] AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups Juan Abdon Miranda-Correa, Mojtaba Khomami Abadi, Nicu Sebe, Ioannis Patras (Submitted on 2 Feb 2017 (v1), last revised 13 Apr 2017 (this version, v3))

[10] https://medium.com/@makcedward/data-augmentation-for-audio-76912b01fdf6

[11] "DEAP: A Database for Emotion Analysis using Physiological Signals (PDF)", *S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras*, IEEE Transaction on Affective Computing, Special Issue on Naturalistic Affect Resources for System Building and Evaluation, *in press*