VISVESWARAIAH TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belgavi-590018

A PROJECT REPORT

On

"ANALYZING PEOPLE'S BEHAVIOR FROM MEDIA"

A Dissertation Submitted in partial fulfillment of the requirement for the degreeof

BACHELOR OF ENGINEERING

In

COMPUTER SCIENCE & ENGINEERING

Submitted by

Pradaap.S.S	**(1RG16CS067)**
Sangamesh	**(1RG16CS083)**
T.N.Varsha	**(1RG16CS103)**
Vikram Sri	**(1RG16CS108)**

Under The Guidance of

Mrs. Soniya Komal.V

Assoc. Professor, Dept. of CSE

RGIT, Bengaluru-32

Department of Computer Science & Engineering

RAJIV GANDHI INSTITUTE OFTECHNOLOGY

Cholanagar, R.T.Nagar Post, Bengaluru-560032

2019-2020

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

(Affiliated to Visveswaraiah Technological University)

Cholanagar, R.T.Nagar Post, Bengaluru-560032

Department of Computer Science & Engineering

CERTIFICATE

This is to certify that the Project Report titled **"Analyzing people's behavior through media"** is a bonafide work carried out by **Mr. Pradaap S S (USN 1RG16CS067), Mr. Sangamesh (USN 1RG16CS083), Ms. T N Varsha (USN 1RG16CS103) and Mr. Vikram Sri (USN 1RG16CS108)** in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science and Engineering** of the **Visveswaraiah Technological University, Belgavi**, during the year **2019-2020.** It is certified that all corrections/suggestions given for Internal Assessment have been incorporated in the report. This project report has been approved as it satisfies the academic requirements in respect of project work (15CSP85) prescribed for the said degree.

eSignature of Guide	eSignature of HOD	eSignature of Principal
Mrs. Soniya Komal.V	**Dr. Manikandan J**	**Dr. Nagaraj.A.M**
Assoc. Professor	Head Of Department	Principal
Dept. of CSE, RGIT, Bengaluru	Dept. of CSE, RGIT, Bengaluru	RGIT, Bengaluru

External Viva

Name of the Examiners	**Signature with date**
1.	
2.	

VISVESWARAIAH TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belgavi-590018

RAJIV GANDHI INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DECLARATION

We hereby declare that the project work entitled **"Analyzing people's behavior through media"** submitted to the **Visveswaraiah Technological University, Belgavi** during the academic year **2019-2020**, is record of an original work done by us under the guidance of **Mrs. Soniya Komal.V,** Assistant Professor, Department of Computer Science and Engineering, RGIT, Bengaluru in the partial fulfillment of requirements for the award of the degree of **Bachelor of Engineering** in **Computer Science & Engineering.** The results embodied in this project have not been submitted to any other University or Institute for award of any degree or diploma.

Pradaap S S **(1RG16CS067)**

Sangamesh **(1RG16CS083)**

T N Varsha **(1RG16CS103)**

Vikram Sri **(1RG16CS108)**

ACKNOWLEDGEMENT

We take this opportunity to thank our college **Rajiv Gandhi Institute of Technology, Bengaluru** for providing us withanopportunity to carry out this project work.

We express our gratitude to **Dr. D G Anand**, Rector, RGIT, Bengaluru and to **Dr. Nagaraj.A.M**, Principal, RGIT, Bengaluru for providing the resources and support without which the completion of this project would have been a difficult task.

We extend our sincere thanks to **Dr. Manikandan J**, Associate Professor and Head, Department of Computer Science and Engineering, RGIT, Bengaluru, for being a pillar of support and encouraging us in the face of all adversities.

We would like to acknowledge the thorough guidance and support extended towards us by **Mrs. Kavyashree**, Assistant Professor, Dept. of CSE, RGIT, Bengaluru, **Mrs. Pushpalatha,** Assistant Professor, Dept. of CSE, RGIT, Bengaluru, **Mrs. Pragathi,** Assistant Professor, Dept. of CSE, RGIT, Bengaluru and **Mrs. Soniya Komal.V** Assistant Professor, Dept. of CSE, RGIT, Bengaluru. Their incessant encouragement and valuable technical support have been of immense help. Their guidance gave us the environment to enhance our knowledge and skills and to reach the pinnacle with sheer determination, dedication and hard work.

We also want to extend our thanks to the entire faculty and support staff of the Department of Computer Science and Engineering, RGIT, Bengaluru, who have encouraged us throughout the course of the Bachelor's Degree. We want to thank our family for always being there with full support and for providing us with a safe haven to conduct and complete our project. We are ever grateful to them for helping us in these stressful times.

Lastly, we want to acknowledge all the helpful insights given to us by all our friends during the course of this project.

Pradaap.S.S (1RG16CS067)

Sangamesh (1RG16CS083)

T N Varsha (1RG16CS103)

Vikram Sri (1RG16CS108)

ABSTRACT

In this world of 7 billion people, around 3.2 billion people suffer from behavioral disorder, depression, anxiety, mentally stressed, etc, which in turn affects one's efficiency to work as well as their health. In this era of technology and management we need to learn how to use technology how to manage and monitor our health. This can be done by using monitoring one's activity which they do the most. According to a survey people are more active on social networking sites (SNS) and media. The use or addiction of media is increasing these days especially among younger generation. These days' media is used for social and professional interaction among people. These activities can be recorded and can be analyzed to judge predict their behavior. This can be done by using SNS data. But if someone is not socially active then in this situation, we can use another approach which is media. In this paper we are focusing on non-social people, because a lot of papers and projects are present on SNS analysis. We are aiming to identify and analyze people's behavior, which could be used for various purposes like screening and also to overcome various health and mental issues

CONTENTS

LIST OF FIGURES / ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Overview

Emotion plays a significant role in daily interpersonal human interactions. This is essential to our rational as well as intelligent decisions. It helps us to match and understand the feelings of others by conveying our feelings and giving feedback to others. Research has revealed the powerful role that emotion plays in shaping human social interaction. Emotional displays convey considerable information about the mental state of an individual.

Contemporary modern world has witnessed the widespread emergence of online social media and similar technologies. Peoples' behavior over different social network platform has become an interesting topic of research. In this study, we investigate whether people express analogous identity over different platforms and analysis of different social platform usage contributes to reveal more of a person.

We analyze people's usage pattern in three major arenas, the most widely audio, video and text. We extract linguistic features and infer personality traits from all the three of these modules. Our study reveals differential relationship between personality traits. Combination of these three profiles gives an extensive view of a user's interest and sensitivity.

In our daily life, we go through different situations and develop feeling about it. Emotion is a strong feeling about human’s situation or relation with others. These feelings and express Emotion are expressed as facial expression. The primary emotion levels are of six types namely; Love, Joy, Anger, Sadness, Fear and Surprise. Human expresses emotion in different ways including facial expression, speech, gestures/actions and written text. As the technology progresses, the internet is now commonly used on PCs, tablets, and smartphones.

This generates a huge amount of data, especially textual data. It has become impossible to manually analyze all the data for a specific purpose. New research directions have emerged from automatic data analysis like automatic emotion analysis. Emotion analysis has attracted researcher's attention because of its applications in different fields. For example, security agencies can track emails/messages/blogs etc. and detect suspicious activities. The business communities nowadays prefer to use emotional marketing. In emotional marketing, they try to stimulate the customers' emotions to buy the products or services. There are many applications of detecting the emotion of the persons like in the interface with robots, audio surveillance, web-based E-learning, commercial applications, clinical studies, entertainment, banking, call centers, cardboard systems, computer games, etc.

1.2 Motivation

Psychometric test is a standard and traditional test which contains multiple choice questions. The user is asked to choose one option out of many. This test is be used identify candidate's personality, characteristics and attitude. It is difficult to identify hidden aspect of a candidate which can be only by face to face interview. Another traditional approach is face to face interview which involves extra man power and time. All these issues can be resolved through this project.

The idea behind creating this project was to build a machine learning model that could detect emotions from the speech, text and video we have with each other all the time. Nowadays personalization is something that is needed in all the things we experience every day. So why not have a emotion detector that will gauge your emotions and in the future recommend you different things based on your mood.

1.3 Problem Identification

Social Media users are the potential viewers of news provided by mass media, but their interest on particular topic is not considered. Time component is not considered while collecting the topic related information. Topics are not ranked according to their popularity by taking into account their prevalence in both news media and social media. The quality of news recommender system is not good. The hidden topics could not be uncovered.

1.4 Scope

This low-cost system can be used for gathering and finding the useful information from social media sites. Since the main aim of this project is to develop a system which can easily and effectively find top news related to a particular topic, it can help the news channels to improve the coverage and provide quality news to users. Also, it can aid the news providers by providing feedbacks about the topics that has been discontinued by the news media but are still being discussed among the general people.

1.5 Objective and Methodology

The main objective of this project is to analyze people's behavior from media either live (for not so active social media users) or uploaded media by splitting media into its different constituents like video, audio and text. Analyzing these parts separately for behavior analysis in order to extract as accurate outcome as possible.

The waterfall method which is considered as the traditional software development method has been used in this project. It's a rigid linear model that consists of sequential phases (requirements, design, implementation, verification, maintenance) in which distinct goals are accomplished. Each phase must be 100% complete before the next phase can start, and traditionally there is no process for going back to modify the project or direction.

1.6 Existing System

In Existing System, project aimed to study analyze people's behavior through separate platforms for each of text, speech and images only. The existing model gives an option to upload a picture and it will analyze and predict your emotion like Happy, Angry, Neutral, surprise, sad and fear. Also, through a form of questionnaire based on the responses the emotion will be detected.

Disadvantages:

- After processing the image, the result displays a mixed of emotions even a person is completely happy or sad. Also, it is not using face detection algorithm. Thus, when giving an image of any object or place as input, it is still showing emotions in percentage even when the image does not have any face in it.
- In questionnaire responses the users can manipulate their responses according to their likes thus decreasing the accuracy. A platform for

analyzing video data with accurate results was not possible in the previously designed model.

1.7 Proposed System

The proposed system will analyze and predict the user behavior. For this purpose, the system will take user's data from media. The collected data is then analyzed and classified.

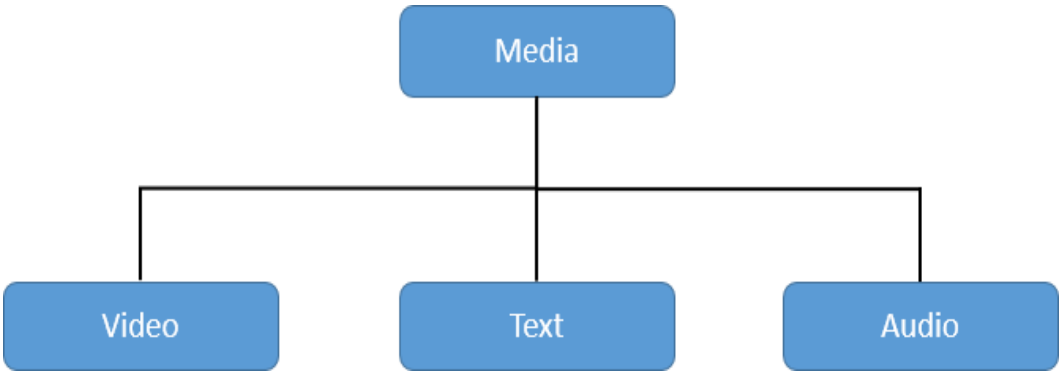


Figure 1.1 Block diagram

The system will work in three steps:

1. Text
2. Video
3. Audio

Text: the data will be analyzed for the sentiments present in it and then sentiment classification will take place. For this purpose, we will use python library TextBlob which will detect the sentiments present in text/post by analyzing the words present in it with the predefined word in the library.

Video: We will use Support Vector Machine (SVM) algorithm. Initially the relevant regions are targeted and the unwanted regions are ignored, then required features will be extracted and then SVM will analyze and classify the sentiments. A support vector machine takes the data points and outputs the hyperplane (which in two dimensions it's simply a line) that best separates the data. This approach is used for displaying classified output by using trained data and test data.

Audio: Audio can be used to analyze sentiments just like text and media is used. In audio module voice is extracted from the video and then features were extracted using by plotting its waveform and spectrogram. audio goes for spectrogram.

The extracted audio is segmented and then classification is done by using CNN algorithm. Through this emotion is analyzed. The next step involves organizing the audio files. Each audio file has a unique identifier at the 6th position of the file name which can be used to determine the emotion the audio file consists. We used Librosa library in Python to process and extract features from the audio files. We have used Librosa library in Python to process and extract features from the audio files. Librosa is a python package for music and audio analysis. It provides the building blocks necessary to create music information retrieval systems. Using the librosa library we were able to extract features i.e. MFCC(Mel Frequency Cepstral Coefficient). MFCCs are a feature widely used in automatic speech and speaker recognition

Advantages:

- The input is divided into separate modules and each separate module is analyzed with the algorithms specialized specific for that module.
- The highest optimal emotion is highly accurate as each module's optimalemotion is compared ; expected outcome v/s actual outcome.
- The input can be either live or stored video, thus making this designed model an on- spot emotion detector.

1.8 Outcome of the Project

In this project we are taking media file as input then we separate frames, voice and text from media. Then analyzing each separated part for persons sentimental characteristics. There are separate modules for analyzing each part using best algorithms. Also, checking and comparing the module's output with expected output for detecting highest optimal output.

- **Highest optimal emotion detected**: After each module is divided into respective audio, video and text, each module after processed with specialized give individual module inputs out of which the highest optimal dominant emotion is detected.
- **Efficient and quick outputs:** For the non- social media users, simple recorded videos are accepted as input to this model and the emotion is detected within short span ofseconds.
- **CSV File**: The outcomes of each module and in last we represent them module in csv file against there input file which gives overall insight about the characteristics of person's in that file.
- **Graphs**: Each of the csv filed outcomes are plotted on graph for accuracy checks,
- expected v/s actual outcome.

1.9 Report Organization

Chapter 1:

This chapter gives the overall description about the project. It gives the overview of the proposed project work. It tries to answer why this project is needed in current scenario and what are various motivation factors that motivated to implement this project. This chapter also points out the limitations in the existing systems and tells how these limitations cam be overcome by using this project.

Chapter 2:

This chapter gives details about various base papers that are related to the proposed project work. It shows how various activities related to the project were carried out at different point of time. It gives a short introduction to each base paper, talks about their shortcomings and tells how this project can overcome those shortcomings.

Chapter 3:

This chapter introduces the system analysis process. It gives brief idea whether this project should be done or not based on various feasibility study. It gives the summary of various feasibility studies that were carried out and shows the advantages of doing this project. At the same time, it also gives the overview of various functional and non-functional requirements of the system.

Chapter 4:

This chapter talks about various hardware and software tools that are necessary in order to implement this project. It provides details of software and languages that will be used and also lists the minimum requirements needed to run the project.

Chapter 5:

This chapter shows the detailed design of the architecture, components, modules, interfaces, and data for the proposed system to satisfy specified requirements. It shows various standard UML diagrams that are needed to design the system. It provides a visualization of how the data will flow among various components of the system.

Chapter 6:

This chapter shows the implementation of the structure created during architectural design and the results of system analysis to construct system elements that

meet the stakeholder requirements and system requirements developed in the early life cycle phases. It shows the segment of programming code that is used in order to implement this project.

Chapter 7:

This chapter shows the various test results produced by the system. Various kinds of test are performed for each part of the system and as well as the whole system. It shows various pre-defined test cases and result of running these test cases on the system. It provides the comparison of expected output and the actual output produced by system based on which bugs are identified and eliminated.

Chapter 8:

This chapter shows various screenshots of the system. It also shows how data processing happens at various stages of the system and the final output is also displayed. And it also shows the outer interface design of the system.

1.10 Introduction Summary

This Introduction chapter gives the overview about the project and gives the short description of the proposed project work. It tries to answer why this project is needed in current scenario and what are various scopes and advantages of this project. This chapter also points out the limitations in the existing systems and tells how these limitations cam be overcome by using this project.

CHAPTER 2
LITERATURE SURVEY

2.1 Related Work

2.1.1 Reference Papers

A journal or book of original entry is the place where journal entries are recorded before they are posted to the ledger accounts. A journal is a record of all the transactions a company has recorded. A research paper is a piece of academic writing based on its author's original research on a particular topic, and the analysis and interpretation of the research findings. It can be either a term paper, a master's thesis or a doctoral dissertation.

2.1.2 Social media analysis

Human beings are social creatures. We need the companionship of others to thrive in life, and the strength of our connections has a huge impact on our mental health and happiness. Being socially connected to others can ease stress, anxiety, and depression, boost self-worth, provide comfort and joy, prevent loneliness, and even add years to your life. On the flip side, lacking strong social connections can pose a serious risk to your mental and emotional health.In today's world, many of us rely on social media platforms such as Facebook, Twitter, Snapchat, YouTube, and Instagram to find and connect with each other. While each has its benefits, it's important to remember that social media can never be a replacement for real-world human connection. It requires in-person contact with others to trigger the hormones that alleviate stress and make you feel happier, healthier, and more positive. Ironically for a technology that's designed to bring people closer together, spending too much time engaging with social media can actually make you feel more lonely and isolated—and exacerbate mental health problems such as anxiety and depression.

2.1.3 Social Network Analysis

Sentiment analysis is contextual mining of text which identifies and extracts subjective information in source material, and helping a business to understand the social sentiment of their brand, product or service while monitoring online conversations. However, analysis of social media streams is usually restricted to just basic sentiment analysis and count based metrics. This is akin to just scratching the surface and missing out on those high value insights that are waiting to be discovered. With the recent advances in deep learning, the ability of algorithms to analyze media has improved considerably.

2.2 A multi-level predictive model for detecting social network users with depression

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
A multi- level predictive model for detecting social media network users with depression	2018	-Akkapo Wongko bla P. -Miguel A Vdillo. -Vasa Crucin.	Depressio n predictive model to find hidden pattern and relationshi p between life satisfactio n and depression social network users	It grades people behaviour area wise so that we can know strong and weak areas.	They are comparing life satisfaction and depression for different users. But it is impossible to identify satisfaction criteria for different users.

The large volume of data generated on social network platforms enables us to detect hidden patterns in data and obtain new insights. This work aims to (a) explore the relationship between life satisfaction and depression in social network users, using Facebook as an example, and (b) develop a multilevel predictive model to detect users with depression. They trained a set of predictive models on datasets from myPersonality project including 2,085 participants who took the Satisfaction with Life Scale and 614 users who submitted the Centre for Epidemiological Study Depression (CES-D) scale. The resulting multilevel model establishes a negative correlation between life satisfaction and depression, and it can also improve the accuracy of a predictive model using depressive labels alone.

2.3 A video object extraction algorithm based on depth map for multi-view video

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTA-GE
A video extractio n algorith m based on depth view map for multi- view video	2018	-Zhou Xailia ng, -Jiang Gangv i, Fu Songv in,Yo meri,S hao Feng,P eng Zongj u,Li Fucui	The video based emotion based recognition efforts are based on extracting emotional states	Features like video chat and conferencing are included for enhancing relationship management	Practical problems like length of the window segment

Video sequences have the rich texture information in practical applications, which makes the extraction of the semantic objects of interest more difficult. This paper presents a video object extraction algorithm based on depth map for multi-view video coding in three- dimensional video system. First of all, gradient operators are used to roughly segment color image into flat and texture regions with threshold, so object contours are extracted, while The OTSU algorithm is used to distinguish backgrounds and foregrounds in the color image, which can fill the pixels of semantic objects. At the same time, inter-frame difference is taken into account, which joins the moving objects into foregrounds, and extracts the interested region with morphological operations. Finally, object of block level is obtained though combination of operators outlined above and block-process though threshold. Compared with the existing algorithms, the proposed algorithm does not adopt popular clustering scheme but joins the OTSU algorithm, thus it can effectively avoid lots of computational complexity which the clustering algorithm brings. Experimental results show that the proposed algorithm can not only extract accurately the semantic objects, but also reduce the computational complexity.

2.4 Automatic sentiment detection in naturalistic audio

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAG E
Automati c sentiment detection in naturalisti c audio	2017	Lakshmish Kaushik, Abhijeet Sangwan, John H.L Hansen	Keyword Spotting is used where text-based sentiment classifier is automatically determine the most useful sentiment bearing keywords rather than analyzing each word in the transcribed text.	Addressing the traditional robustness problems of ASR (accent, noise, etc.) can have significant impact of performance.	Increasing the size of the keyword list can potentially increase the number of false alarms (i.e., the system falsely detects a keyword that is not actually present in the audio).

Audio sentiment analysis using automatic speech recognition is an emerging research area where opinion or sentiment exhibited by a speaker is detected from natural audio. It is relatively under-explored when compared to text based sentiment detection. Extracting speaker sentiment from natural audio sources is a challenging problem. Generic methods for sentiment extraction generally use transcripts from a speech recognition system, and process the transcript using text-based sentiment classifiers. In this study, we show that this baseline system is sub-optimal for audio sentiment extraction. Alternatively, new architecture using keyword spotting (KWS) is proposed for sentiment detection. In the new architecture, a text-based sentiment classifier is utilized to automatically determine the most useful and discriminative sentiment-bearing keyword terms, which are then used as a term list for KWS. In order to obtain a compact yet discriminative sentiment term list, iterative feature optimization for maximum entropy sentiment model is proposed to reduce model complexity while maintaining effective classification accuracy.

2.5 Video extraction and tracking of moving targets with a camera

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Video extraction and tracking of moving targets with a camera	2017	Adrien Nègre, Dann Laneu ville	It uses background subtraction approach to obtain a binary mask of the moving pixels by means of a clustering step.	It should add functionality of tracking multiple objects simultaneously.	It tracks the object Using background Subtraction approach. So, it fails when there are multiple objects in the video.

Cameras are nowadays widely used in maritime surveillance applications to detect small non-cooperative boats. This paper presents a simple and efficient video extractor based upon the background subtraction approach to obtain a binary mask of the moving pixels which are then segmented into targets by means of a clustering step. For the first step, they use a robust Kalman filter to obtain the binary mask with an Expectation / Maximization step to update the model's parameters. Then, an adaptive bandwidth mean shift approach is used to obtain the final clusters corresponding to the detected objects, comprising real targets and false alarms. Finally, a multitarget tracking algorithm fed by the bounding box centers corresponding to the detected objects is used to track the different targets in the image plane and to filter out the false alarms. Real results in a maritime surveillance context are presented.

2.6 Emotion recognition from speech signal

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Emotion recogniti -on from speech signal	2017	Esther Ramdinmawii , Abhijit Mohanta	Basic emotions like Anger, Happy etc are analyzed from emotional speech signals..	Signal processing methods are used for obtaining the production features from these signals.	Voice should be clear. If there is noise in the background, then its Efficiency decreases. So, it has to be done in a silent environment.

Emotion recognition is a rapidly growing research domain in recent years. Unlike humans, machines lack the abilities to perceive and show emotions. But human-computer interaction can be improved by automated emotions recognition, thereby reducing the need of human intervention. In this paper, four basic emotions (Anger, Happy, Fear and Neutral) are analyzed from emotional speech signals. Signal processing methods are used for obtaining the production features from these signals. Source feature the instantaneous fundamental frequency (F0), system features the formants and dominant frequencies, zero-crossing rate (ZCR), and the combined features signal energy are used for the analyses. Distinct differences are observed between high-arousal emotions (Anger and Happy) and Neutral emotion. Results indicate overlap between Anger and Happy emotions. But distinct differences are observed in the features for Happy/Anger and Fear, and between Happy and Anger emotions which is otherwise a challenging problem.

2.7 Sentiment analysis from social networking sites (SNS) using machine learning approach for the measurement of depression

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurem ent of depression	2017	Anees Ul Hassan ; Jamil Hussain ; Musarrat Hussain ; Muhammad Sadiq ; Sung young Lee	Find the depression by extracting emotions from the text, using emotion theories, machine learning techniques, and NLP.	Specialized algorithms are used to detect the most accurate emotion from text	Here it is only Analyzing the text from social media. People don't share their view using text only, but also from photos & videos.

The advent of different social networking sites has enabled anyone to easily create, express, and share their ideas, thoughts, opinions, and feelings about anything with millions of other people around the world. With the advancement of technology, mini computers and smartphones have come to human pockets and now it is very easy to share your idea about anything on social media platforms like Facebook, twitter, Wikipedia, LinkedIn, Google+, Instagram etc. Due to the tremendous growth in population and communication technologies during the last decade, use of social networks is on the rise and they are being used for many different purposes. One such service for which their use may be explored is an analysis of users post to diagnosis depression. IN this paper, they present how to find the depression level of a person by observing and extracting emotions from the text, using emotion theories, machine learning techniques, and natural language processing techniques on different social media platforms.

2.8 Object recognition in videos by sequential frame extraction using CNN

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Object recognition in videos by sequential frame extraction using CNN	2017	Meghajit ,Mazumdar , V Sarasvathi, Akshay Kumar	In this approach, a method to develop an interactive application in order to detect objects from videos is proposed.	Further improved and complexity to get faster results.	The accuracy Percentage is 77% that is not bad, but Processing is not very fast.

In this paper, a method to develop an interactive application in order to detect objects from videos is proposed. The application is able to classify the video according to a particular genre. Also, upon user input, it is also able to detect the particular object being shown at that instant on the screen. A sequential frame extraction method of videos and also deep learning approach of Convolutional Neural Networks along with Fully Connected Neural Networks is used for this task. The method gives good accuracy of average 77 percent.

2.9 Predicting depression levels using social media posts

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Predicting depression levels using social media posts	2016	Phooi Seng Li-Minn Ang Chien Shing Ooi	Mining the User generated Content using Support Vector Machines (SVM) and Naïve Bayes	User can also share their views in comments.	Here it is only taking user's post

This paper proposes an audio-visual emotion recognition system that uses a mixture of rule-based and machine learning techniques to improve the recognition efficacy in the audio and video paths. The visual path is designed using the Bidirectional Principal Component Analysis (BDPCA) and Least-Square Linear Discriminant Analysis (LSLDA) for dimensionality reduction and class discrimination. The extracted visual features are passed into a newly designed Optimized Kernel-Laplacian Radial Basis Function (OKL-RBF) neural classifier. The audio path is designed using a combination of input prosodic features (pitch, log-energy, zero crossing rates and Teager energy operator) and spectral features (Mel-scale frequency cepstral coefficients). The extracted audio features are passed into an audio feature level fusion module that uses a set of rules to determine the most likely emotion contained in the audio signal. An audio visual fusion module fuses outputs from both paths. The performances of the proposed audio path, visual path, and the final system are evaluated on standard databases. Experiment results and comparisons reveal the good performance of the proposed system.

2.10 Twitter, MySpace, Digg: Unsupervised sentiment analysis insocial media

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Twitter, MySpace, Digg:Unsupervised Sentiment Analysis in Social Media	2012	Georgios Paltoglou Mike Thelwall	It uses an unsupervised, lexicon-based classifier that estimates the level of emotional valence in text in order to make a prediction, explicitly designed to address the issue of sentiment	Even the polarity(i.e. , predicting whether a Subjective text is Negatively or Positively oriented)	It fails to identify informal abbreviations like "m8" for "mate") used on social media,only polarity is shown.

Sentiment analysis is a growing area of research with significant applications in both industry and academia. Most of the proposed solutions are centered around supervised, machine learning approaches and review-oriented datasets. In this article, we focus on the more common informal textual communication on the Web, such as online discussions, tweets and social network comments and propose an intuitive, less domain-specific, unsupervised, lexicon-based approach that estimates the level of emotional intensity contained in text in order to make a prediction. Our approach can be applied to, and is tested in, two different but complementary contexts: subjectivity detection and polarity classification. Extensive experiments were carried onthree real-world datasets, extracted from online social Web sites and annotated by human evaluators, against state- of-the-art supervised approaches. The results demonstrate that the proposed algorithm, even though unsupervised, outperforms machine learning solutions in the majority of cases, overall presenting a very robust and reliable solution for sentiment analysis of informal communication on the Web.

2.11 Emotion recognition from text using semantic andseparablemixture model

NAME	YEAR	AUTHOR	FEATURE	ADAVANTAGE	DISADVANTAGE
Emotion recognitio n from text using semantic labels and separable mixture models	2006	Chung-hsien Wu Ze-jing Chuang Yu-chung Lin	A separable mixture model (SMM) is adopted to estimate the similarity between an input sentence and the EARs(emotion association rules) of each emotional state.	A separate model is adopted based on emotion association rules(EARs)	Due to the ambiguities in EGRs(emo tion generation rules), they adopted only happy, unhappy, and neutral emotional states for evaluation

This study presents a novel approach to automatic emotion recognition from text. First, emotion generation rules are manually deduced from psychology to represent the conditions for generating emotion. Based on the ERGs the emotional state of sentence can be represented as a sequence of semantic labels (SLs) and attributes (ATTs). SL are defined as the domain independent features while ATTs are domain dependent. The emotional association rules (EARs) represented SLs and ATTs for each emotion are automatically derived from the sentences in an emotional text corpus using the a-priori algorithm. Finally, a SMM is adopted to estimate the similarity between an input sentence and the EARs of each emotional state. Since some features defined in this approach are domain-dependent, a dialog system focusing on the daily expressions is constructed and only three emotional states happy, unhappy and neutral are considered for performance evaluation. According to the results of the experiments given the domain corpus, the proposed approach is promising**.**

2.12 Literature Survey Summary

This chapter gives details about various base papers that are related to the proposed project work. It shows how various activities related to the project were carried out at different point of time. It gives a short introduction to each base paper, talks about their shortcomings and tells how this project can overcome those shortcomings.

CHAPTER 3

SYSTEM ANALYSIS

3.1 Introduction to System Analysis

- **System:** A system is an orderly group of interdependent components linked together according to a plan to achieve a specific objective. Its main characteristics are organization, interaction, interdependence, integration and a centralobjective.

- **Analysis:** Analysis is a detailed study of the various operations performed by a system and their relationships within and outside of the system. One aspect of analysis is defining the boundaries of the system and determining whether or not a candidate system should consider other related systems. During analysis data are collected on the available files decision points and transactions handled by the present system. This involves gathering information and using structured tools for analysis.

- **System Analysis:** System analysis and design are the application of the system approach to problem solving generally using computers. To reconstruct a system the analyst must consider its elements output and inputs, processors, controls, feedback and environment.

3.2 Feasibility Study

Feasibility is the determination of whether or not a project is worth doing. The process followed in making this determination is called feasibility Study. This type of study if a project can and should be taken. In the conduct of the feasibility study, the analyst will usually consider seven distinct, but inter-related types of feasibility.

3.2.1 Technical Feasibility

This is considered with specifying equipment and software that will successfully satisfy the user requirement the technical needs of the system may vary considerably but might include

- The facility to produce outputs in a giventime.
- Response time under certain conditions.
- Ability to process a certain column of transaction at a particular speed.

3.2.2 Economic Feasibility

Economic analysis is the most frequently used technique for evaluating the effectiveness of a proposed system. More commonly known as cost / benefit analysis. The procedure is to determine the benefits and savings are expected form a proposed system and a compare them with costs. It benefits outweigh costs; a decision is taken to design and implement the system will have to be made if it is to have a chance of being approved. There is an ongoing effort that improves in accuracy at each phase of the system life cycle.

3.2.3 Operational Feasibility

It is mainly related to human organization and political aspects. These points are considered are

- What changes will be brought with the system?
- What organizational structures are distributed?
- What new skills will be required?
- Do the existing system staff members have these skills?
- If not, can they be trained in the course oftime?

3.3 Functional Requirements

Various functional modules that can be implemented by the system will be:

Record video

If user wants to give a live video for the analysis.

Linking Video

User can upload the previously recorded video of them.

Processing Data

The data is processed by the system.

Get the results

After the processing is successful the user will get the optimal emotion as the result.

3.4 Non-Functional Requirements

Various non-functional modules that can be implemented by the system will be:

Security

- The system provides privacy to confidential customer information.
- The system must automatically log out all customers after a period ofinactivity.
- The system should not leave any cookies on the customer's computer containing the user's password.
- The system's back-end servers shall only be accessible toauthenticated administrators.

Reliability

- The system provides storage of all databases on redundant computerswith automatic switchover.
- The overall stability ofthe system depends on the stability ofcontainer and its underlying operating system.

Availability

- The system should be available at all times, meaning the user can access it using a web browser, only restricted by the down time of the server on which the system runs.

3.5 System Analysis Summary

This chapter introduces the system analysis process. It gives brief idea whether this project should be done or not based on various feasibility study. It gives the summary of various feasibility studies that were carried out and shows the advantages of doing this project. At the same time, it also gives the overview of various functional and non-functional requirements of the system.

CHAPTER 4

REQUIREMENT ANALYSIS

4.1 Functional Requirements

A functional requirement defines a function of a system or its component, where a function is described as a specification of behavior between outputs and inputs:

- System should be able to take the input as media
- System should be able to analyze and output the behavioral status.

4.2 Non-Functional Requirements

A non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors.

- Accuracy
- Speed
- Security
- Consistency

4.3 System Requirements

- Window 8 and above
- 4GB RAM
- Hard Disk: 200GB
- Camera
- Microphone
- Internet access

4.4 Software Requirements

• Spyder

Spyder is a powerful scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection and beautiful visualization capabilities of a scientific package. Furthermore, Spyder offers built in integration with many popular scientific packages, including NumPy, SciPy, Pandas, IPython, QtConsole, Matplotlib, SymPy, and more. Beyond its many built-in features, Spyder's abilities can be extended even further via first- and third-party plugins. Spyder can also be used as a PyQt5 extension library, allowing you to build upon its functionality and embed its components, such as the interactive console or advanced editor, in your own software.

• Python

Python is an interpreter, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding; make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python language itself is managed by the Python Software Foundation, who offers a reference implementation of Python, called, CPython, under an open source license. You can even download the Python source code, here. Besides the Python implementation itself being open source, many open source projects make use of Python, and Python has many libraries available for developers under open source licenses.

• Tensorflow

TensorFlow is an open source software library for high performance numerical computation. Its flexible architecture allows easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices. Originally developed by researchers and engineers from the Google Brain team within Google's AI organization, it comes with strong support for machine learning and deep learning and the flexible numerical computation core is used across many other scientific domains.

Unlike other numerical libraries intended for use in Deep Learning like Theano, TensorFlow was designed for use both in research and development and in production systems, not least RankBrain in Google search and the fun DeepDream project. It can run on single CPU systems, GPUs and large scale distributed systems of hundreds of machines. We are using it's library for mathematical analysis of data collected from the sns.

4.5 Requirement Analysis Summary

This chapter talks about various hardware and software tools that are necessary in order to implement this project. It provides details of software and languages that will be used and also lists the minimum requirements needed to run the project.

CHAPTER 5

SYSTEM DESIGN

Systems design is the process of defining elements of a system like modules, architecture, components and their interfaces and data for a system based on the specified requirements. It is the process of defining, developing and designing systems which satisfies the specific needs and requirements of a business or organization.

5.1 System architecture

The system architecture: An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

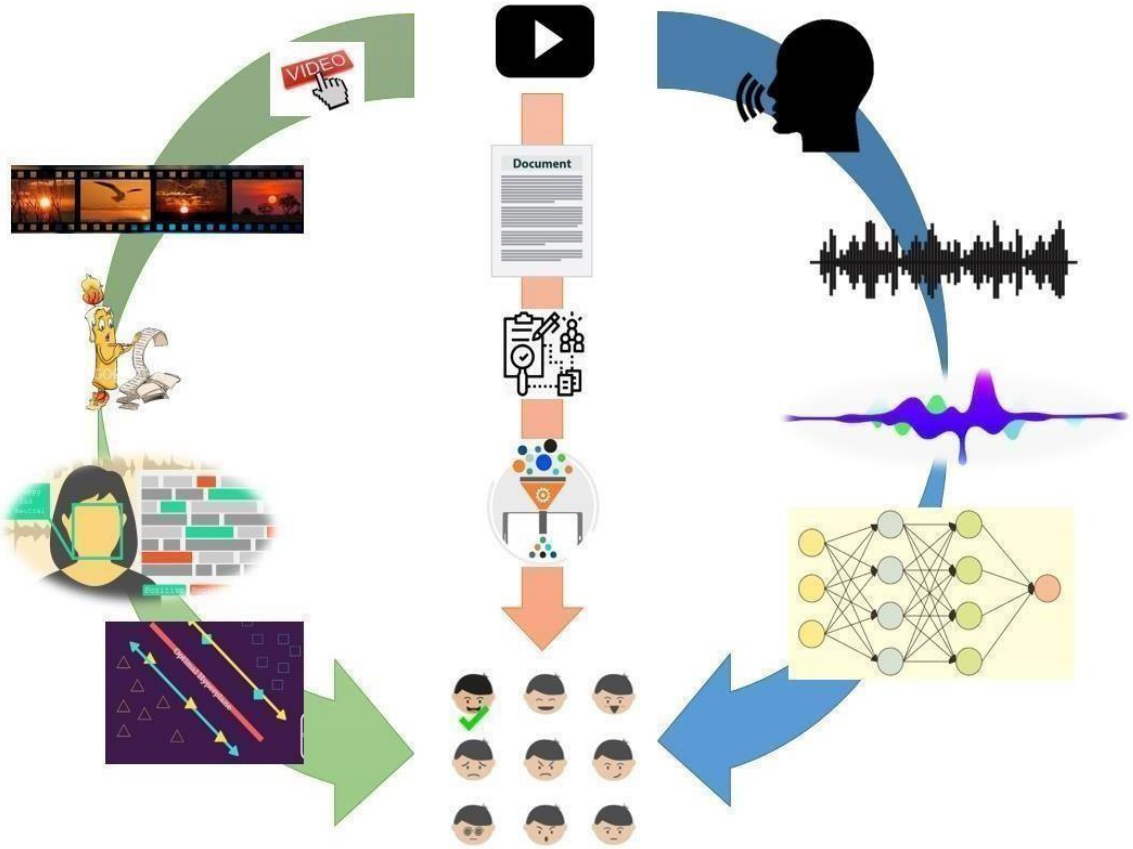


Figure 5.1 - System architecture

In our system figure 5.1 media analysis, data is collected then processing is done. After processing, a module is generated, which then compared to predefined module and hit percentage is calculated. After combining and comparing the hit percentage of media, outcome can be determined. The system will work in three steps:

1. Text
2. Video
3. Audio

Text: the data will be analyzed for the sentiments present in it and then sentiment classification will take place. For this purpose, we will use python library TextBlob which will detect the sentiments present in text/post by analyzing the words present in it with the predefined word in the library.

Video: This project works by first detecting a face in each frame of a video. The face region and the landmark coordinates found and then used to detect the probabilities of each prototypic emotion (happiness, sadness, anger, fear, surprise and disgust, plus the neutral expression). The face detection is based on a Cascade Detector (Viola-Jones algorithm). Its trained detector model is loaded from the folder `models'. The emotion detection is based on a Support Vector Machine (SVM) using the RBF kernel and trained from labelled images obtained from two openly available (for non-commercial use) datasets. The bank of Gabor filters is constructed using the functions from the scikit-image package instead of the OpenCV library. The reason is because OpenCV's implementations only produces the real part of the filter response, and the solution used in this project requires both the real and imaginary parts of the Gabor responses. Nonetheless, the application of the filter uses convolution, which is performed using OpenCV. This application is the most computationally intensive task (convolution of 32 filters for each frame of a video), which causes the processing to run with a very low frame rate.

Audio: Audio can be used to analyze sentiments just like text and media is used. In audio module voice is extracted from the video and then features were extracted using by plotting its waveform and spectrogram. audio goes for spectrogram. The extracted audio is segmented and then classification is done by using CNN algorithm. Through this emotion is analyzed. The next step involves organizing the audio files. Each audio file has a unique identifier at the 6th position of the file name which can be used to determine the emotion the audio file consists.

5.2 Gant chart

5.2.1 Gantt Chart for Phase One

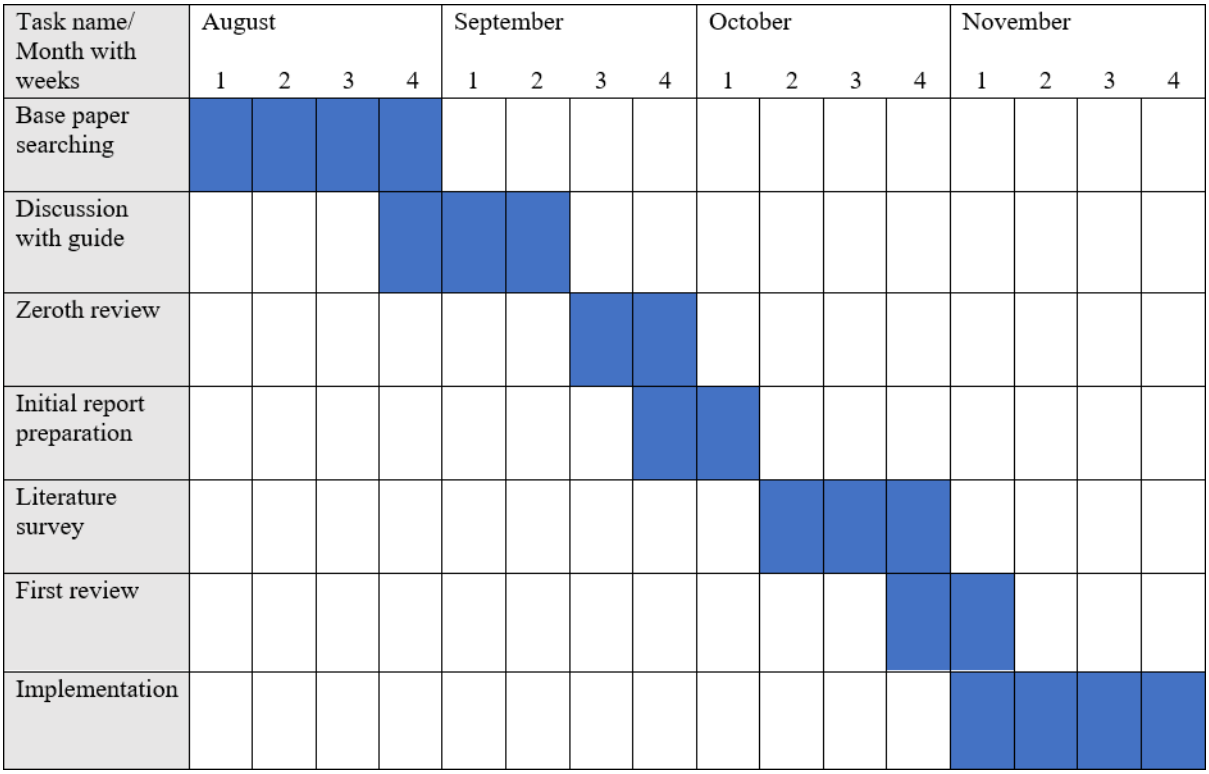


Figure 5.2 Gantt chart For Phase One

5.2.2 Gantt Chart for Phase Two

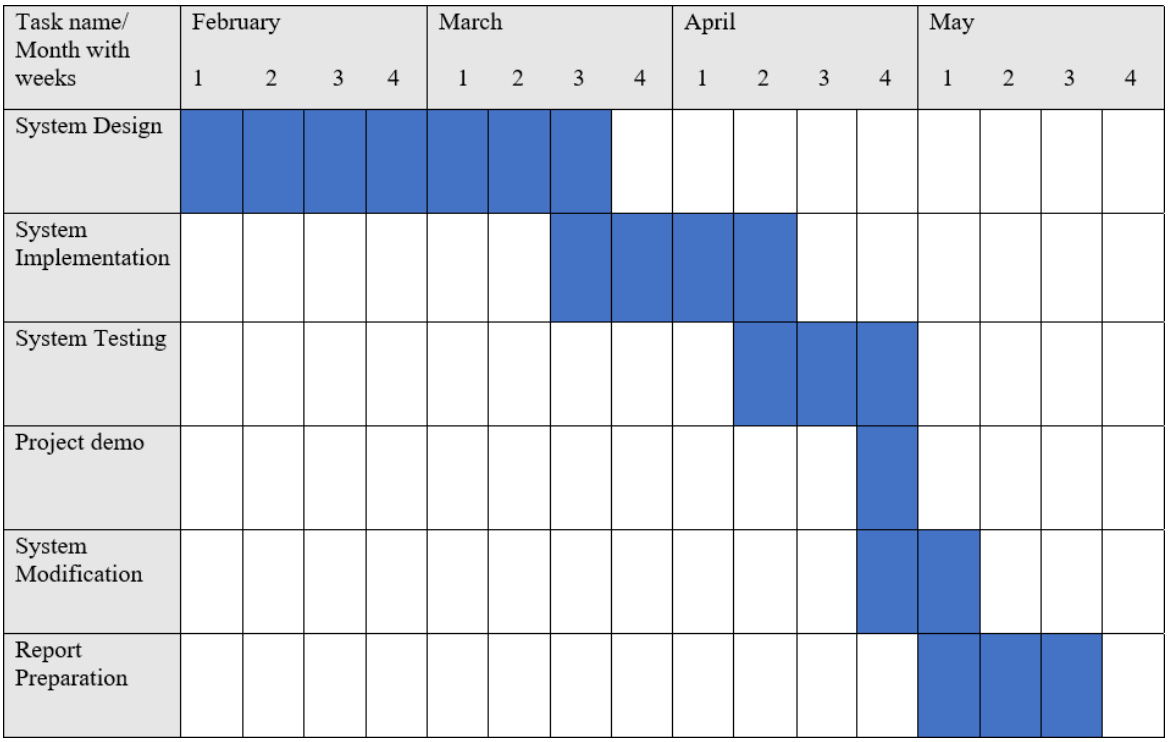


Figure 5.3 Gantt chart For Phase Two

5.3 Life cycle model

The agile model is applied for the software development process in our project. Agile means relating to or denoting a method of project management, used especially for software development, that is characterized by the division of tasks into short phases of work and frequent reassessment and adaptation of plans.

phases involved in our project:

1. Recognizing hardware and software requirements
2. Developing and working on architecture diagram, class diagram, use case designand sequence diagram
3. developing the software required for execution and working on GUI
4. Testing with various media file and checking if model classifies them correctly
5. Deployment On Windows OS and using Python 3.7 along with Anaconda Spyder, and jupyter including the Librosa and the TextBlob libraries of python.
6. Review and test of whether accuracy desired has beenachieved.

Repetitively, this process iterates to improve accuracy ofclassification

Figure 5.4 Agile Model

5.4 Data flow diagram

DFD graphically representing the functions, or processes, which capture, manipulate, store, and distribute data between a system and its environment and between components of a

system. The visual representation makes it a good communication tool between User and System designer. Structure of DFD allows starting from a broad overview and expand it to a hierarchy of detailed diagrams. DFD has often been used due to the following reasons:

- Logical information flow of the system
- Determination of physical system construction requirements
- Establishment of manual and automated systems requirements

Figure 4.2 shows Data Flow Diagram of the proposed system. In this diagram the overview of the working architecture is shown. As we can see that the input of our system is media. Media can be an type of video. The input can be recorded from any device which has camera. It could be a web cam or a mobile device.

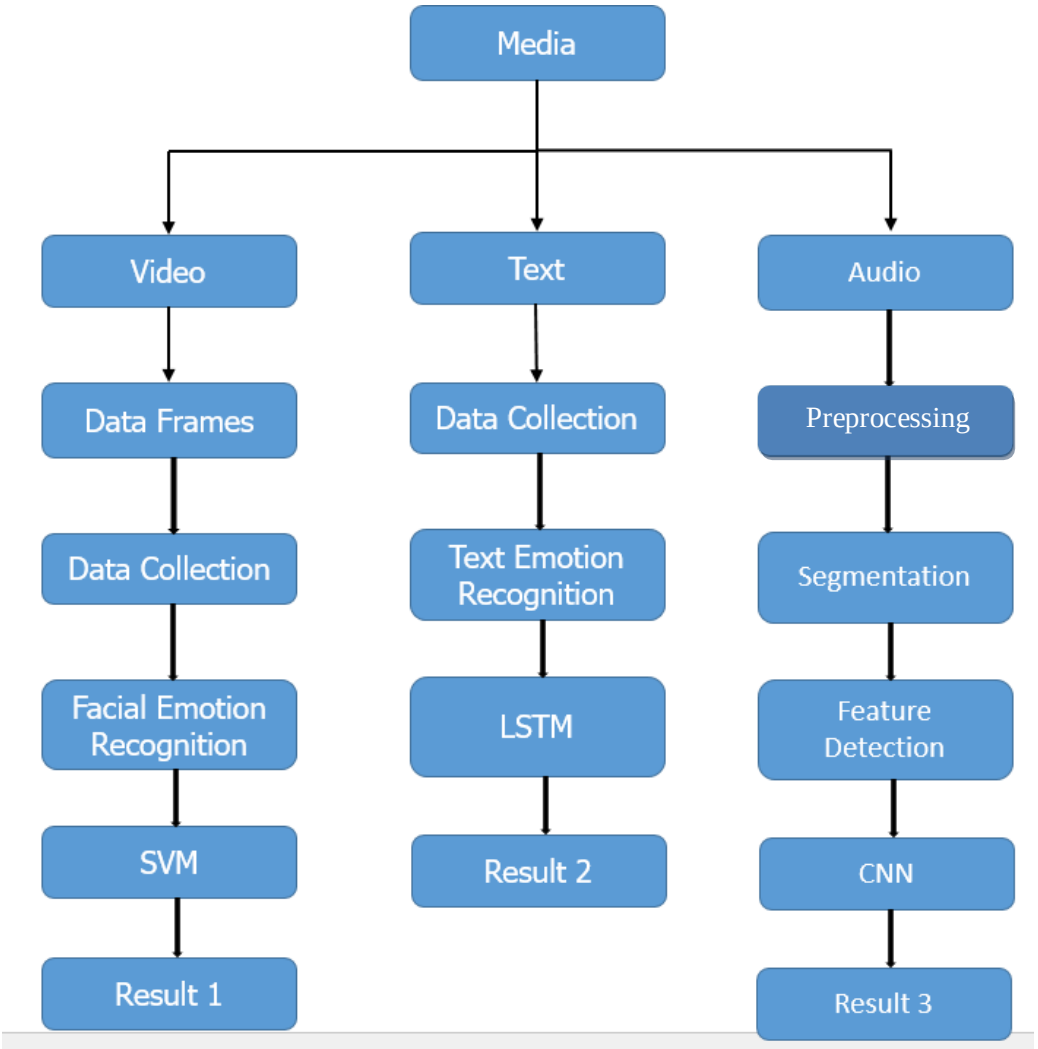


Figure 5.5 – Data flow diagram

Further input is divided into 3 major modules. Audio, video and Text. These modules take different input and give emotion as an output.

5.4.1 Video

- This module can take input as live data or stored video.
- It extracts video without voice from the media.
- Then the voiceless video is converted into data frames.
- These data frames are analyzed and emotion is detected from each module.
- With the emotion the optimality of the emotion is also recorded.
- The average optimality of same emotion keeps updating till the video runs.
- The emotion with maximum optimality is displayed as output.

5.4.2 Text

- This is the second method of analyzing emotion from a video.
- In this module, audio is extracted from the video.
- Then the audio is converted into a text by using google api.
- Now the text sentiment analyses is done by using text.
- Useless words like punctuations, nouns, etc are removed from the text. What remains are the words with can be used to detect emotions.
- Each word has an emotion associated with it. Our job is to detect that emotion.
- Each emotion has some optimal value attached with it. That value is recoded every time an emotion is detected.
- The average of optimal value of same emotion is taken.
- The emotion with maximum optimality wins.

5.4.3 Audio

- This is the third module and another way of detecting emotion from a video.
- This is one of the best ways of emotion analysis.
- In this module we focus mostly on the pitch, power, frequency, etc.
- By using the properties of the voice emotion is detected.
- Analyzing Personal Mood on a daily basis by using Conversational Data

- Using Mood Analysis for Custom Suggestions in Third Party Services
- Analyzation of Shopper Experience from Customer Conversation
- Automation of Customer Care Call Rating
- Monitoring of Mental Health Status of Patients in Hospitals for Long Periods ofTime
- This is a minimalistic UI which doesn't require opening WebApp every time, and automatically keeps track of progress in background and displays in a track bar.

5.5 Use case diagram

Use case diagram refers to the behaviors, set of action that some system should or can perform in collaboration with one or more external users of the system (actors). Emotion recognition is the process of identifying human emotion. This is both something that humans do automatically but computational methodologies have also been developed. Humans show universal consistency in recognizing emotions but also show a great deal of variability between individuals in their abilities. This has been a major topic of study in psychology.

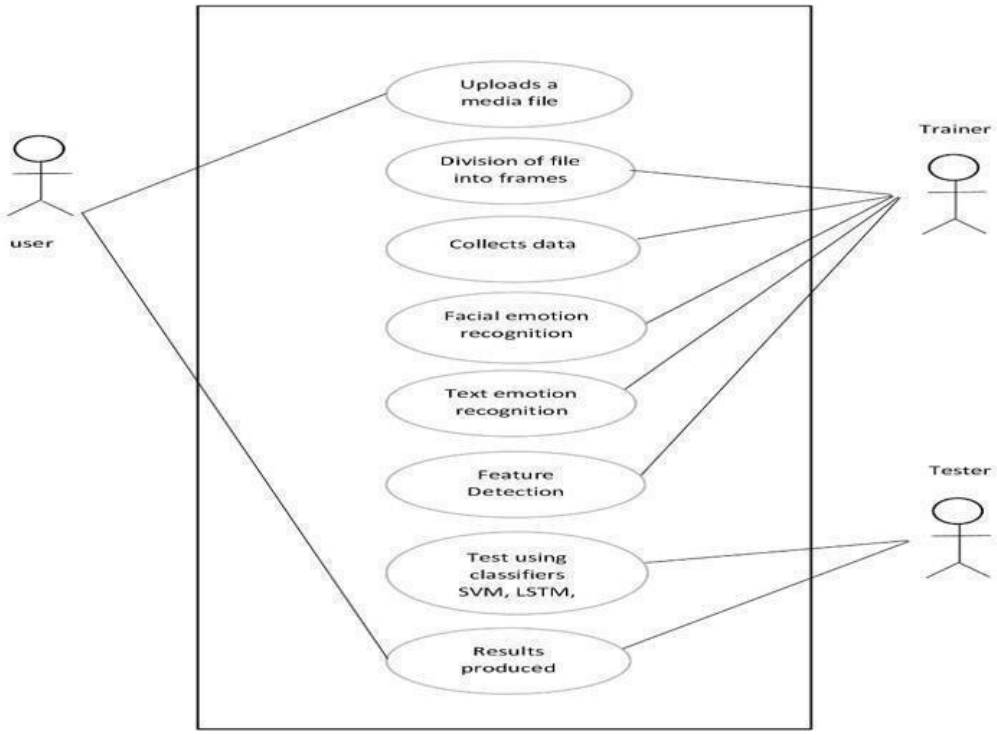


Figure 5.6 Use case diagram

The task of emotion recognition often involves the analysis of human expressions in multimodal forms such as texts, audio, or video. Different emotion types are detected

through the integration of information from facial expressions, body movement and gestures, and speech. The technology is said to contribute in the emergence of the so-called emotional or emotive Internet.

5.6 Sequence diagram

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios. A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.

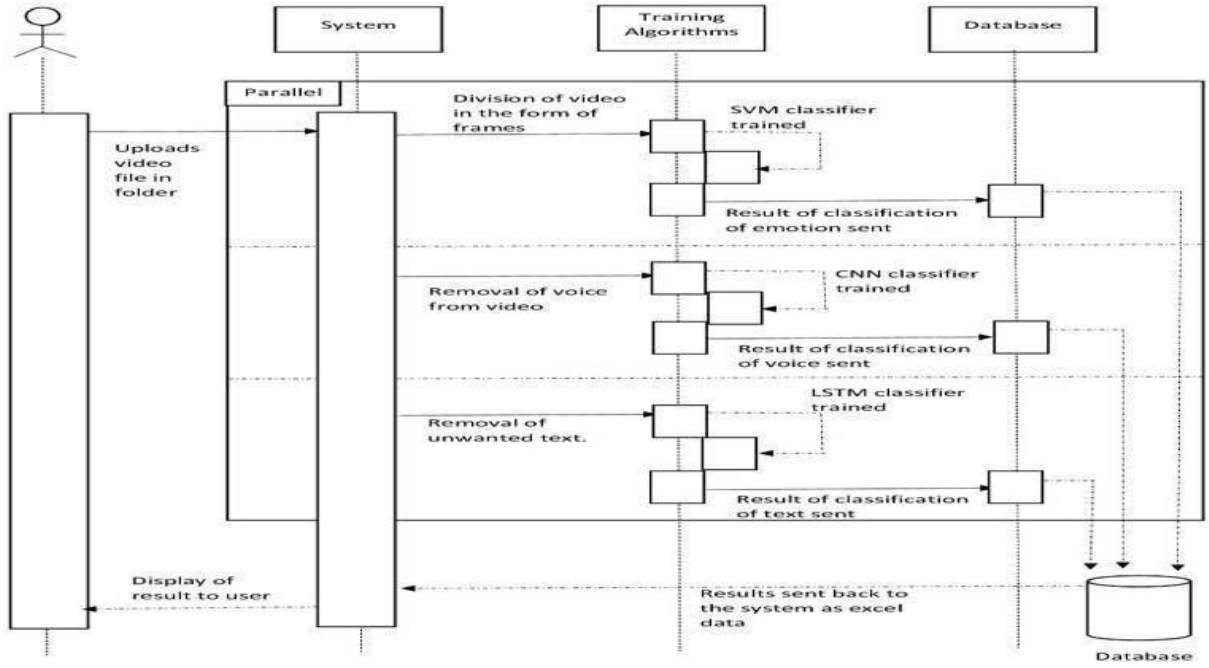


Figure 5.7 Sequence Diagram

5.7 Class diagram

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects.

Purpose of Class Diagrams

1. Shows static structure of classifiers in a system.
2. Diagram provides basic notation for other structure diagrams prescribed by UML.
3. Helpful for developers and other team members.
4. Business Analysts can use class diagrams to model systems from business perspective.

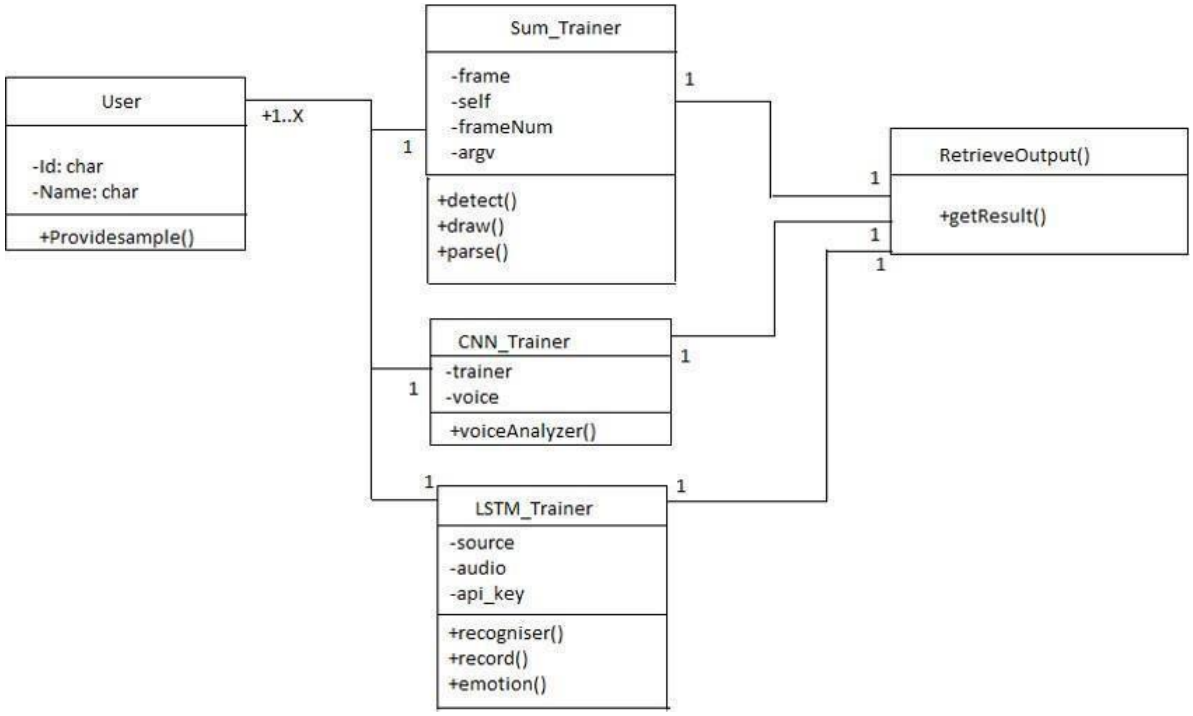


Figure 5.8 Class diagram

5.8 System design summary

This chapter shows the detailed design of the architecture, components, modules, interfaces, and data for the proposed system to satisfy specified requirements. It shows various standard UML diagrams that are needed to design the system. It provides a visualization of how the data will flow among various components of the system.

CHAPTER 6

SYSTEM IMPLIMENTATION

System Implementation uses the structure created during architectural design and the results of system analysis to construct system elements that meet the stakeholder requirements and system requirements developed in the early life cycle phases. These system elements are then integrated to form intermediate aggregates and finally the complete system-of-interest (SoI). Implementation is the process that actually yields the lowest-level system elements in the system hierarchy (system breakdown structure). System elements are made, bought, or reused. Production involves the hardware fabrication processes of forming, removing, joining, and finishing, the software realization processes of coding and testing, or the operational procedures development processes for operators' roles.

6.1 Modular Description

Modular design, or "modularity in design", is a design approach that subdivides a system into smaller parts called modules or skids, that can be independently created and then used in different systems. A modular system can be characterized by functional partitioning into discrete scalable, reusable modules; rigorous use of well-defined modular interfaces; and making use of industry standards for interfaces.

Besides reduction in cost (due to less customization, and shorter learning time), and flexibility in design, modularity offers other benefits such as augmentation (adding new solution by merely plugging in a new module), and exclusion.

Our project consists of three main modules and those modules are:

1. Video
2. Text
3. Audio

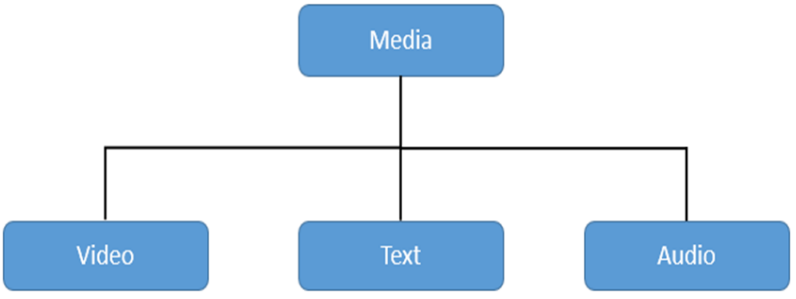


Figure 6.1 Major module

6.1.1 Text

The data will be analysed for the sentiments present in it and then sentiment classification will take place. For this purpose we will use python library TextBlob which will detect the sentiments present in text/post by analysing the words present in it with the predefined word in the library.

6.1.2 Video

This project works by first detecting a face in each frame of a video. The face region and the landmark coordinates found and then used to detect the probabilities of each prototypic emotion (happiness, sadness, anger, fear, surprise and disgust, plus the neutral expression). The face detection is based on a Cascade Detector (Viola-Jones algorithm). The facial landmarks are located by using a deformable model adjusted to the face region previously found with the Cascade. The code uses the implementation available from the dlib library. Its trained detector model is loaded from the folder `models`.The emotion detection is based on a Support Vector Machine (SVM) using the RBF kernel and trained from labelled images obtained from two openly available (for non commercial use) datasets. The code uses the SVM implementation from the Scikit-Learn package. The training and the detection is based on the responses of a bank of applied to the face region on the frame image. Only the responses under the coordinates of the facial landmarks are used. The bank of Gabor filters is implemented in file `gabor.py`. The trained detector model is also loaded from the folder `models`.The bank of Gabor filters is constructed using the functions from the scikit-image package instead of the OpenCV library. The reason is because OpenCV's implementations only produces the real part of the filter response, and the solution used in this project requires both the real and imaginary parts of the Gabor responses. Nonetheless, the application of the filter uses convolution, which is performed using OpenCV. This application is the most computational intensive task (convolution of 32 filters for each frame of a video), which causes the processing to run with a very low frame rate.

6.1.3 Audio

Audio can be used to analyze sentiments just like text and media is used. In audio module voice is extracted from the video and then features were extracted using by plotting its waveform and spectrogram. audio goes for spectrogram. The extracted audio is segmented and then classification is done by using CNN algorithm. Through this emotion is analysed. The next step involves organizing the audio files. Each audio file has a unique identifier at the 6th position of the file name which can be used to determine the

emotion the audio file consists. We used Librosa library in Python to process and extract features from the audio files. We have used Librosa library in Python to process and extract features from the audio files. Librosa is a python package for music and audio analysis. It provides the building blocks necessary to create music information retrieval systems. Using the librosa library we were able to extract features i.e MFCC(Mel Frequency Cepstral Coefficient). MFCCs are a feature widely used in automatic speech and speaker recognition.

6.2 Programming Code

Video module source code

```
import sys
import argparse
import cv2
import numpy as np
from collections import OrderedDict
from datetime import datetime, timedelta
from faces import FaceDetector
from data import FaceData
from gabor import GaborBank
from emotions import EmotionsDetector
count=0
val_0=0
val_1=0
val_2=0
val_3=0
val_4=0
val_5=0
val_6=0
#-----------------------------------------
```

```
class VideoData:
    def init (self):
        self._faceDet = FaceDetector()
        self._bank = GaborBank()
        self._emotionsDet = EmotionsDetector()
        self._face = FaceData()
        self._emotions = OrderedDict()
    #----------------------------------------
    def detect(self, frame):
        ret, face = self._faceDet.detect(frame)
        if ret:
            self._face = face
            # Crop just the face region
            frame, face = face.crop(frame)
            # Filter it with the Gabor bank
            responses = self._bank.filter(frame)
            # Detect the prototypic emotions based on the filter responses
            self._emotions = self._emotionsDet.detect(face, responses)
            return True
        else:
            self._face = None
            return False

    #----------------------------------------
    def draw(self, frame):
        # Font settings
        font = cv2.FONT_HERSHEY_SIMPLEX
        scale = 0.5
```

```
thick = 1
glow = 3 * thick
# Color settings
black = (0, 0, 0)
white = (255, 255, 255)
yellow = (0, 255, 255)
red = (0, 0, 255)
empty = True
# Plot the face landmarks and face distance
x = 5
y = 0
w = int(frame.shape[1]* 0.2)
try:
   face = self._face
   empty = face.isEmpty()
   face.draw(frame)
except:
   pass
# Plot the emotionprobabilities
global count
global   val_0
global   val_1
global   val_2
global   val_3
global   val_4
global   val_5
global val_6
try:
```

```
    emotions = self._emotions
    if empty:
        labels = []
        values = []
        #count=0
    else:
        labels = list(emotions.keys())
        values = list(emotions.values())
        bigger = labels[values.index(max(values))]
        #d1[bigger]=max(values)
        count=count+1
        val_0=val_0+values[0]
        val_1=val_1+values[1]
        val_2=val_2+values[2]
        val_3=val_3+values[3]
        val_4=val_4+values[4]
        val_5=val_5+values[5]
        val_6=val_6+values[6]
        print(count)
        print("Neutral ",val_0)
        print("Happiness ",val_1)
        print("Sadness ",val_2)
        print("Anger ",val_3)
        print("Fear ",val_4)
        print("Surprise ",val_5)
        print("Disgust ",val_6)
        #print (bigger)
        #print (max(values))
```

```
        # Draw the header
        text = 'emotions'
        size, _ = cv2.getTextSize(text, font, scale, thick)
        y += size[1] + 20
        cv2.putText(frame, text, (x, y), font, scale, black, glow)
        cv2.putText(frame, text, (x, y), font, scale, yellow, thick)
        y += 5
        cv2.line(frame, (x,y), (x+w,y), black, 1)
    size, _ = cv2.getTextSize('happiness', font, scale, thick)
    t = size[0] + 20
    w = 150
    h = size[1]
    for l, v in zip(labels, values):
        lab = '{}:'.format(l)
        val = '{:.2f}'.format(v)
        size, _ = cv2.getTextSize(l, font, scale, thick)
        # Set a red color for the emotion with bigger probability
        color = red if l == bigger else yellow
        y += size[1] + 15
        p1 = (x+t, y-size[1]-5)
        p2 = (x+t+w, y-size[1]+h+5)
        cv2.rectangle(frame, p1, p2, black, 1)
        # Draw the filled rectangle proportional to the probability
        p2 = (p1[0] + int((p2[0] - p1[0]) * v), p2[1])
        cv2.rectangle(frame, p1, p2, color, -1)
        cv2.rectangle(frame, p1, p2, black, 1)
        # Draw the emotion label
        cv2.putText(frame, lab, (x, y), font, scale, black, glow)
```

```
            cv2.putText(frame, lab, (x, y), font, scale, color, thick)
            # Draw the value of the emotion probability
            cv2.putText(frame, val, (x+t+5, y), font, scale, black, glow)
            cv2.putText(frame, val, (x+t+5, y), font, scale, white, thick)
        except Exception as e:
            print(e)
            pass
#----------------------------------------------
def main(argv):
    # Parse the command line
    args = parseCommandLine(argv)
    # Loads the video or starts the webcam
    if args.source == 'cam':
        video = cv2.VideoCapture(args.id)
        if not video.isOpened():
            print('Error opening webcam of id {}'.format(args.id))
            sys.exit(-1)
        fps = 0
        frameCount = 0
        sourceName = 'Webcam #{}'.format(args.id)
    else:
        video = cv2.VideoCapture(args.file)
        file_name=args.file
        if not video.isOpened():
            print('Error opening video file {}'.format(args.file))
            sys.exit(-1)
        fps = int(video.get(cv2.CAP_PROP_FPS))
        frameCount = int(video.get(cv2.CAP_PROP_FRAME_COUNT))
```

```
    sourceName = args.file
video.set(cv2.CAP_PROP_FRAME_WIDTH, 1280);
video.set(cv2.CAP_PROP_FRAME_HEIGHT, 720);
# Create the helper class
data = VideoData()
# Text settings
font = cv2.FONT_HERSHEY_SIMPLEX
scale = 1
thick = 1
glow = 3 * thick
# Color settings
color = (255, 255, 255)
paused = False
frameNum = 0

# Process the video input
while True:
    if not paused:
        start = datetime.now()
    ret, img = video.read()
    if ret:
        frame = img.copy()
    else:
        paused = True
    drawInfo(frame, frameNum, frameCount, paused, fps, args.source)
    data.detect(frame)
    data.draw(frame)
    cv2.imshow(sourceName, frame)
```

```
    if paused:
        key = cv2.waitKey(0)
    else:
        end = datetime.now()
        delta = (end - start)
        if fps != 0:
            delay = int(max(1, ((1 / fps) - delta.total_seconds()) * 1000))
        else:
            delay = 1
        key = cv2.waitKey(delay)
    if key == ord('q') or key == ord('Q') or key == 27:
        break
    elif key == ord('p') or key == ord('P'):
        paused = not paused
    elif args.source == 'video' and (key == ord('r') or key == ord('R')):
        frameNum = 0
        video.set(cv2.CAP_PROP_POS_FRAMES, frameNum)
    elif args.source == 'video' and paused and key == 2424832: # Left key
        frameNum -= 1
        if frameNum < 0:
            frameNum = 0
        video.set(cv2.CAP_PROP_POS_FRAMES, frameNum)
    elif args.source == 'video' and paused and key == 2555904: # Right key
        frameNum += 1
        if frameNum >= frameCount:
            frameNum = frameCount - 1
    elif args.source == 'video' and key == 2162688: # Pageup key
        frameNum -= (fps * 10)
```

```
            if frameNum < 0:
                frameNum = 0
            video.set(cv2.CAP_PROP_POS_FRAMES, frameNum)
        elif args.source == 'video' and key == 2228224: # Pagedown key
            frameNum += (fps * 10)
            if frameNum >= frameCount:
                frameNum = frameCount - 1
            video.set(cv2.CAP_PROP_POS_FRAMES, frameNum)
        elif key == 7340032: # F1
            showHelp(sourceName, frame.shape)
        if not paused:
            frameNum += 1
    video.release()
    cv2.destroyAllWindows()
    print("\n")
    print("Average_Neutral ",(val_0/count))
    print("Average_Happiness ",(val_1/count))
    print("Average_Sadness ",(val_2/count))
    print("Average_Anger ",(val_3/count))
    print("Average_Fear ",(val_4/count))
    print("Average_Surprise ",(val_5/count))
    print("Average_Disgust ",(val_6/count))

best_emotion=max(val_0/count,val_1/count,val_2/count,val_3/count,val_4/count,val_5/count,va
l_6/count)
    if best_emotion==(val_0/count):
        import pandas as pd
        df=pd.read_excel("../output.xlsx", header=None)
```

```
    numrows=len(df)
    df.loc[numrows+1]=[file_name,'Neutral',",",]
    df.to_excel("../output.xlsx",header=None,index=None)
    print("Best average emotion is Neutral with optimality of ",best_emotion)
 if best_emotion==(val_1/count):
    import pandas as pd
    df=pd.read_excel("../output.xlsx", header=None)
    numrows=len(df)
    df.loc[numrows+1]=[file_name,'Happiness',",",]
    df.to_excel("../output.xlsx",header=None,index=None)
    print("Best average emotion is Happiness with optimality of ",best_emotion)
 if best_emotion==(val_2/count):
    import pandas as pd
    df=pd.read_excel("../output.xlsx", header=None)
    numrows=len(df)
    df.loc[numrows+1]=[file_name,'Sadness',",",]
    df.to_excel("../output.xlsx",header=None,index=None)
    print("Best average emotion is Sadness with optimality of ",best_emotion)
 if best_emotion==(val_3/count):
    import pandas as pd
    df=pd.read_excel("../output.xlsx", header=None)
    numrows=len(df)
    df.loc[numrows+1]=[file_name,'Anger',",",]
    df.to_excel("../output.xlsx",header=None,index=None)
    print("Best average emotion is Anger with optimality of ",best_emotion)
 if best_emotion==(val_4/count):
    import pandas as pd
    df=pd.read_excel("../output.xlsx", header=None)
```

```
        numrows=len(df)
        df.loc[numrows+1]=[file_name,'Fear',",",]
        df.to_excel("../output.xlsx",header=None,index=None)
        print("Best average emotion is Fear with optimality of ",best_emotion)
    if best_emotion==(val_5/count):
        import pandas as pd
        df=pd.read_excel("../output.xlsx", header=None)
        numrows=len(df)
        df.loc[numrows+1]=[file_name,'Surprise',",",]
        df.to_excel("../output.xlsx",header=None,index=None)
        print("Best average emotion is Surprise with optimality of ",best_emotion)
    if best_emotion==(val_6/count):
        import pandas as pd
        df=pd.read_excel("../output.xlsx", header=None)
        numrows=len(df)
        df.loc[numrows+1]=[file_name,'Disgust',",",]
        df.to_excel("../output.xlsx",header=None,index=None)
        print("Best average emotion is Disgust with optimality of ",best_emotion)
#---------------------------------------------
def drawInfo(frame, frameNum, frameCount, paused, fps, source):
    # Font settings
    font = cv2.FONT_HERSHEY_SIMPLEX
    scale = 0.5
    thick = 1
    glow = 3 * thick
    # Color settings
    black = (0, 0, 0)
    yellow = (0, 255, 255)
```

```
# Print the current frame number and timestamp
if source == 'video':
    text = 'Frame: {:d}/{:d} {}'.format(frameNum, frameCount - 1,
                                        '(paused)' if paused else '')
else:
    text = 'Frame: {:d} {}'.format(frameNum, '(paused)' if paused else '')
size, _ = cv2.getTextSize(text, font, scale, thick)
x = 5
y = frame.shape[0] - 2 * size[1]
cv2.putText(frame, text, (x, y), font, scale, black, glow)
cv2.putText(frame, text, (x, y), font, scale, yellow, thick)
if source == 'video':
    timestamp = datetime.min + timedelta(seconds=(frameNum / fps))
    elapsedTime = datetime.strftime(timestamp, '%H:%M:%S')
    timestamp = datetime.min + timedelta(seconds=(frameCount / fps))
    totalTime = datetime.strftime(timestamp, '%H:%M:%S')
    text = 'Time: {}/{}'.format(elapsedTime, totalTime)
    size, _ = cv2.getTextSize(text, font, scale, thick)
    y = frame.shape[0] - 5
    cv2.putText(frame, text, (x, y), font, scale, black, glow)
    cv2.putText(frame, text, (x, y), font, scale, yellow, thick)
# Print the help message
text = 'Press F1 for help'
size, _ = cv2.getTextSize(text, font, scale, thick)
x = frame.shape[1] - size[0] - 5
y = frame.shape[0] - size[1] + 5
cv2.putText(frame, text, (x, y), font, scale, black, glow)
cv2.putText(frame, text, (x, y), font, scale, yellow, thick)
```

```
#-----------------------------------------

def showHelp(windowTitle, shape):

    # Font settings

    font = cv2.FONT_HERSHEY_SIMPLEX

    scale = 1.0

    thick = 1

    # Color settings

    black = (0, 0, 0)

    red = (0, 0, 255)

    # Create the background image

    image = np.ones((shape[0], shape[1], 3)) * 255

    # The help text is printed in one line per item in this list

    helpText = [

    'Controls:',

    '-----------------------------------------',

    '[q] or [ESC]: quits from the application.',

    '[p]: toggles paused/playing the video/webcam input.',

    '[r]: restarts the video playback (video input only).',

    '[left/right arrow]: displays the previous/next frame (video input only).',

    '[page-up/down]: rewinds/fast forwards by 10 seconds (video input only).',

    ' ',

    ' ',

    'Press any key to close this window...'

    ]



    # Print the controls help text

    xCenter = image.shape[1] // 2

    yCenter = image.shape[0] // 2
```

```
    margin = 20 # between-lines margin in pixels
    textWidth = 0
    textHeight = margin * (len(helpText) - 1)
    lineHeight = 0
    for line in helpText:
        size, _ = cv2.getTextSize(line, font, scale, thick)
        textHeight += size[1]
        textWidth = size[0] if size[0] > textWidth else textWidth
        lineHeight = size[1] if size[1] > lineHeight else lineHeight
    x = xCenter - textWidth // 2
    y = yCenter - textHeight // 2
    for line in helpText:
        cv2.putText(image, line, (x, y), font, scale, black, thick * 3)
        cv2.putText(image, line, (x, y), font, scale, red, thick)
        y += margin + lineHeight
    # Show the image and wait for a key press
    cv2.imshow(windowTitle, image)
    cv2.waitKey(0)
#---------------------------------------------
def parseCommandLine(argv):
    parser = argparse.ArgumentParser(description='Tests the face and emotion '
                                     'detector on a video file input.')
    parser.add_argument('source', nargs='?', const='Yes',
                        choices=['video', 'cam'], default='cam',
                        help='Indicate the source of the input images for '
                        'the detectors: "video" for a video file or '
                        '"cam" for a webcam. The default is "cam".')
    parser.add_argument('-f', '--file', metavar='<name>',
```

```
                help='Name of the video file to use, if the source is '
                "'video". The supported formats depend on the codecs '
                'installed in the operating system.')
    parser.add_argument('-i', '--id', metavar='<number>', default=0, type=int,
                help='Numerical id of the webcam to use, if the source '
                'is "cam". The default is 0.')
    args = parser.parse_args()
    if args.source == 'video' and args.file is None:
        parser.error('-f is required when source is "video"')
    return args
#-----------------------------------------------
# namespace verification for invoking main
#-----------------------------------------------
if__name__== '_main_':
    main(sys.argv[1:])
```

Text module source code

```
import speech_recognition as sr
AUDIO_FILE = ("happy.wav")
# use the audio file as the audio source
r = sr.Recognizer()
with sr.AudioFile(AUDIO_FILE) as source:
    #reads the audio file. Here we use record instead of
    #listen
    audio = r.record(source)
try:
    r.recognize_google(audio)
```

```
  #print("The audio file contains: " + r.recognize_google(audio))
except sr.UnknownValueError:
  print("Google Speech Recognition could not understand audio")
except sr.RequestError as e:
  print("Could not request results from Google Speech Recognition service; {0}".format(e))
import paralleldots
api_key = "PDoPGu3RJ9hHoKfxioElOzakPy7aRl1yAARJv7sR0mk"
#text      = "There are Kashmiris who love India & hate terrorism & Pakistan. We support them. We can't forget how Jawan Aurangzeb, Lieutenant Ummer & many Kashmiri policemen, armymen made sacrifices for India. Now, It's time for Kashmiris to look reality that How Separaists are using them."
text=r.recognize_google(audio)
paralleldots.set_api_key( api_key )
print( "\nEmotion" )
print( paralleldots.emotion( text ) )
import json
y=json.dumps(paralleldots.emotion( text ))
z=json.loads(y)
print(z['emotion'])
import pandas as pd
df=pd.read_excel("../output.xlsx", header=None)
numrows=len(df)
df.iloc[numrows-1,2]=z['emotion']
df.to_excel("../output.xlsx",header=None,index=None)
```

Audio module source code

```
data, sampling_rate = librosa.load('happy.wav')// audio input
%pylab inline
import os
```

```
import pandas as pd

import librosa

import glob

plt.figure(figsize=(15, 5))

librosa.display.waveplot(data, sr=sampling_rate)

#livedf= pd.DataFrame(columns=['feature'])

X, sample_rate = librosa.load('happy.wav',
res_type='kaiser_fast',duration=2.5,sr=22050*2,offset=0.5)

sample_rate = np.array(sample_rate)

mfccs = np.mean(librosa.feature.mfcc(y=X, sr=sample_rate, n_mfcc=13),axis=0)

featurelive = mfccs

livedf2 = featurelive

livedf2= pd.DataFrame(data=livedf2)

livedf2 = livedf2.stack().to_frame().T

livedf2

twodim= np.expand_dims(livedf2, axis=2)

livepreds = loaded_model.predict(twodim, batch_size=32, verbose=1)

livepreds

livepreds1=livepreds.argmax(axis=1)

liveabc = livepreds1.astype(int).flatten()

livepredictions = (lb.inverse_transform((liveabc)))

livepredictions

import pandas as pd

df=pd.read_excel("../output.xlsx", header=None)

numrows=len(df)

df.iloc[numrows-1,3]=livepredictions[0]

df.to_excel("../output.xlsx",header=None,index=None)
```

6.3 System Implementation Summary

This chapter shows the implementation of the structure created during architectural design and the results of system analysis to construct system elements that meet the stakeholder requirements and system requirements developed in the early life cycle phases. It shows the segment of programming code that is used in order to implement this project.

CHAPTER 7

TESTING

Software testing is defined as an activity to check the whether the actual results match the expected results and to ensure that the software system is Detect free. It involves execution of a software component or system component to evaluate one or more properties of interest.

Software testing also helps to identify errors, gaps or missing requirements in contrary to the actual requirements. It can be either done manually or using automated tools. Some prefer saying software testing as a White Box and Black Box Testing.

7.1 Validation and System Testing

Validation testing is a concern which overlaps with integration testing. Ensuring that the application fulfils its specification is a major criterion for the construction of an integration test. Validation testing also overlaps to a large extent with ***System Testing***, where the application is tested with respect to its typical working environment. Consequently for many processes no clear division between validation and system testing can be made. Specific tests which can be performed in either or both stages include the following.

7.1.1 Software Testing: Usually software testing is considered as one phase of the software development life cycle. In the software development life cycle (SDLC) the Testing is plays an importance role, which helps to improve the quality, reliability & performance of the system with all check what all functions software supposed to do & also check that Software is not doing what he not supposed to do. There are major Importance of testing in the part of SDLC and it is better to introduce testing in the early stage of SDLC phases so it help to identify the defects in the early stage & try to avoid the bugs finding & get resolve in the last critical stage.

7.1.2 Validation: Validation is the process of evaluating software during the development process or at the end of the development process to determine whether it satisfies specified business requirements. Validation Testing ensures that the product actually meets the client's needs. It can also be defined as to demonstrate that the product fulfils its intended use when deployed on appropriate environment.

7.1.3 Reasons for performing software validation: From ensuring the fulfilment of particular requirements of a specific intended use to investigating any deviations from the established parameters, there are various reasons for performing software validation. Other reasons for implementing software validation are:

- It ensures that the software has been built in conformance with the requirements of the customer.
- Helps catch errors that were missed or overlooked during the software verification.
- Enables the team to build a software that fulfils user's requirements.
- Execution of code comes under software validation.
- Helps improve the quality and value of the software product.

We have tested our modules for multiple input files (.mp4) for accessing the accuracy of our modules. During validation we found that our models are working fine and actual outcomes are as per expected outcomes.

	Expected Outcome			Actual Outcome		
Input File (.mp4)	Video	Text	Audio	Video	Text	Audio
video1	fear	fear	female_fearful	fear	fear	female_fearful
video2	anger	angry	male_angry	anger	angry	male_angry
video3	neutral	happy	female_calm	neutral	happy	female_calm
video4	anger	angry	male_angry	anger	angry	male_angry
video5	fear	fear	female_fearful	fear	fear	female_fearful
video6	sadness	sad	female_sad	sadness	sad	female_sad
video7	sadness	sad	male_sad	sadness	sad	male_sad
video8	fear	fear	male_fearful	fear	fear	male_fearful
video9	sadness	sad	male_sad	sadness	sad	male_sad
video10	fear	fear	male_fearful	fear	fear	male_fearful
video11	neutral	happy	male_calm	neutral	happy	male_calm
video12	anger	angry	male_angry	anger	angry	male_angry
video13	fear	fear	male_fearful	fear	fear	male_fearful
video14	happiness	happy	male_happy	happiness	happy	male_happy
video15	sadness	sad	male_sad	sadness	sad	male_sad
video16	happiness	happy	female_happy	happiness	happy	female_happy
video17	anger	angry	female_angry	anger	angry	female_calm
video18	sadness	sad	male_sad	sadness	sad	male_sad
video19	happiness	happy	female_happy	happiness	happy	female_happy
video20	anger	angry	female_angry	anger	angry	female_calm

Table 7.1: Comparison between actual and expected outcome

As shown in table 6.1 we have divided outcomes model wise. Then we have tested each video for each models video, text and audio and found no bugs in the program and working fine and giving correct output. We compared expected outcome for each model with observed output. We found that our models outputs are very close to expected model and fulfils the requirement.

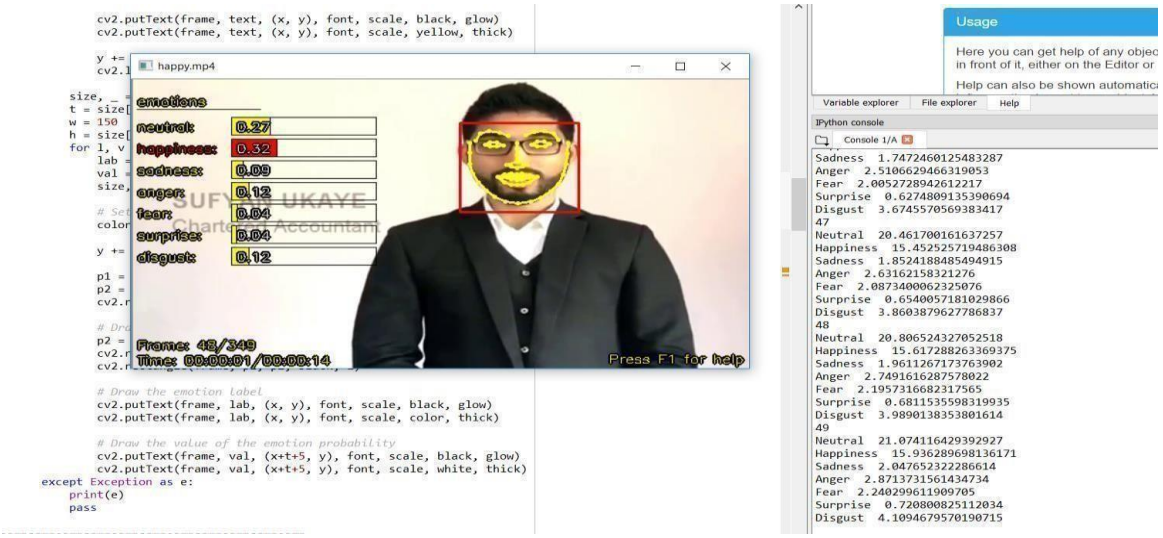
7.2 Testing Summary

This chapter shows the various test results produced by the system. Various kinds of test are performed for each part of the system and as well as the whole system. It shows various pre-defined test cases and result of running these test cases on the system. It provides the comparison of expected output and the actual output produced by system based on which bugs are identified and eliminated.

CHAPTER 8

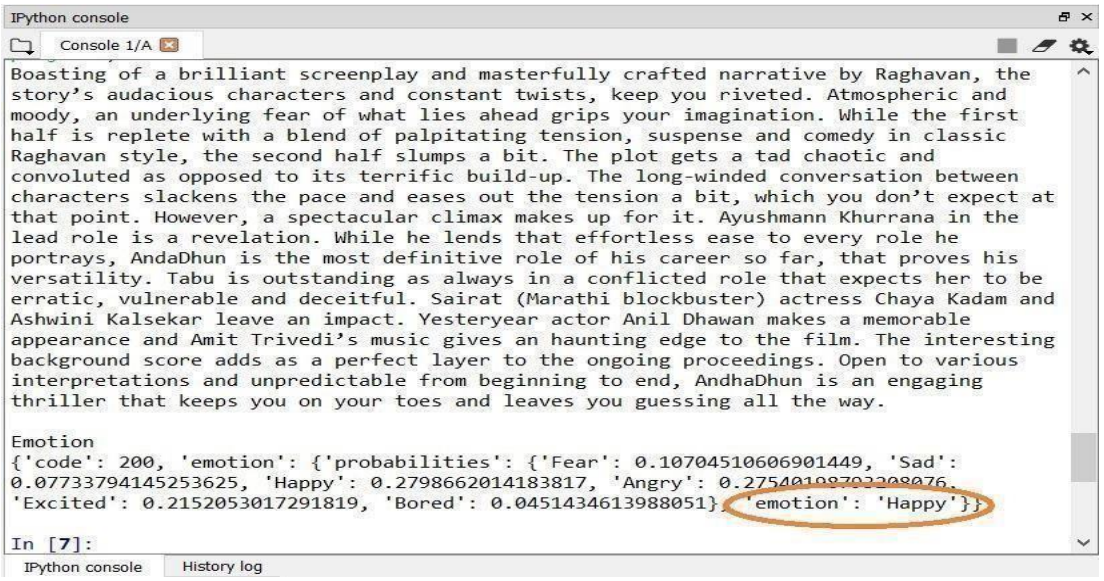
SAMPLE OUTPUT

8.1 Video module



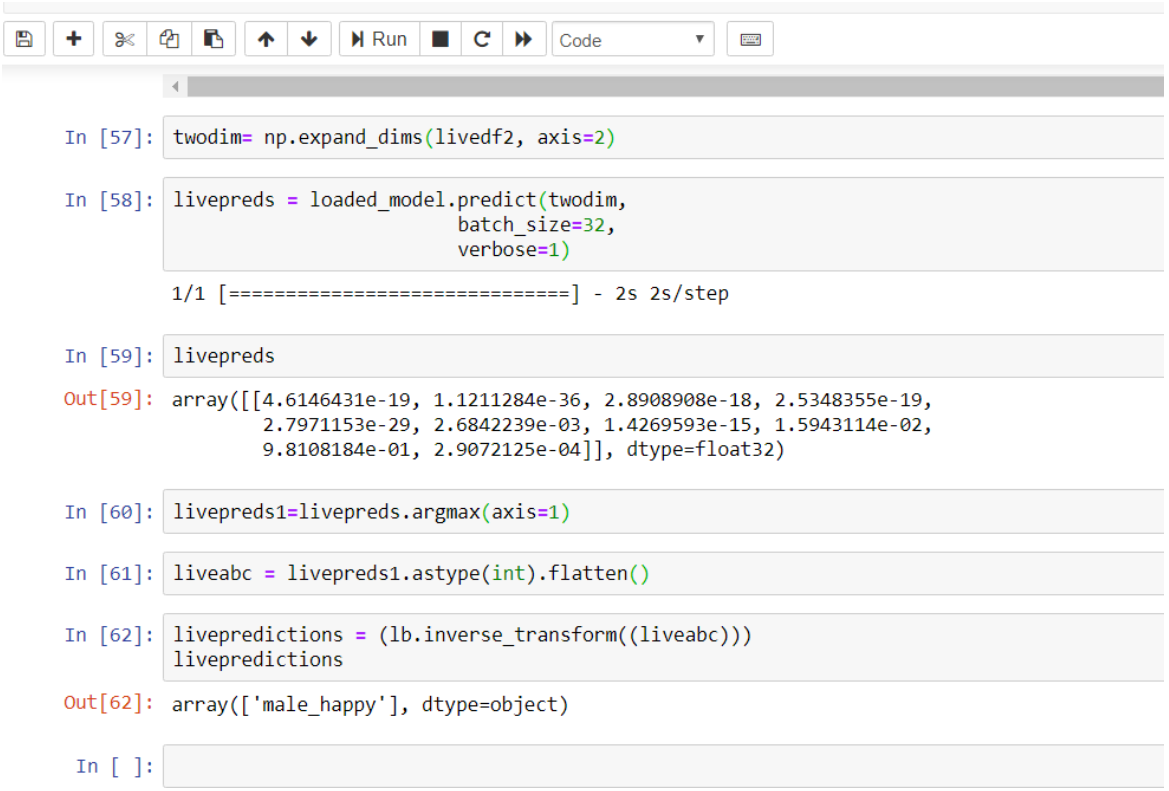
Snapshot 8.1 Emotions extracted from video frames

8.2 Text module



Snapshot 8.2 Emotion extracted from text

8.3 Audio module



Snapshot 8.3 Emotion extracted from Audio

8.4 Sample Output Summary

This chapter shows various screenshots of the system. It also shows how data processing happens at various stages of the system and the final output is also displayed. And it also shows the outer interface design of the system.

CONCLUSION

We were successfully able to extract emotions from different parts of a media file such as video, audio and text. We were also successful in getting the best optimal emotion for each modules of the media file under consideration. We were able to represent outcomes of input files module wise into csv file which can be further analyzed for more insights and decision makings. We have faced certain issues during this project like person should continuously maintain eye contact with webcam otherwise frames are skipped. In voice analysis, voice should be clear and loud and without noise.

Future Scope

As future work, we intend to perform experiments and expand sentiment analysis on different areas and datasets. This project can be further enhanced by integrating all the three sub-models to get optimal emotion depicted by the media file. This will increase the accuracy and will reduce the redundancy of the model. Video module can be further enhanced for overcome skipping issue and some work is required over voice part so that it can generate accurate outcomes even in noise full environment.

REFERENCES

[1] G. L. Kaushik, A. Sangwan and J. H. L. Hansen, "Automatic Sentiment Detection in Naturalistic Audio," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 8, pp. 1668-1679, Aug. 2017. Doi: 10.1109/TASLP.2017.2678164

[2] K. P. Seng, L. Ang and C. S. Ooi, "A Combined Rule-Based & Machine Learning Audio-Visual Emotion Recognition Approach," in IEEE Transactions on Affective Computing, vol. 9, no. 1, pp. 3-13, 1 Jan.-March 2018. Doi: 10.1109/TAFFC.2016.2588488 Eason

[3] Georgios Paltoglou and Mike Thelwall. 2012. Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media. ACM Trans. Intell. Syst. Technol. 3, 4, Article 66 (September 2012), 19 pages. Doi=http://dx.doi.org/10.1145/2337542.2337551

[4] Chung-Hsien Wu, Ze-Jing Chuang, and Yu-Chung Lin. 2006. Emotion recognition from text using semantic labels and separable mixture models. 5, 2 (June 2006), 165-183. Doi=http://dx.doi.org/10.1145/1165255.1165259

[5] Pérez Espinosa H, Escalante H, Villaseñor-Pineda L, Montes-y-Gómez M, Pinto-Avedaño D, Reyez-Meza V. Fusing Affective Dimensions and Audio-Visual Features from Segmented Video for Depression Recognition. Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge - AVEC '14. 2014. Doi:10.1145/2661806.2661815

[6] Wang R, Wang W, daSilva A et al. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2018;2(1):1-26. doi:10.1145/3191775

[7] A. Wongkoblap, M. A. Vadillo and V. Curcin, "Classifying Depressed Users With Multiple Instance Learning from Social Network Data," 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, 2018, pp. 436-436. doi: 10.1109/ICHI.2018.00088

Online References:

- Anaconda:
  https://docs.anaconda.com/anaconda/install/

  https://docs.anaconda.com/anaconda/user-guide/getting-started/
- Python
  http://tutorialspoint.com/python
  http://documentation.flytheme.net/how-to-use-anaconda-navigator-for coding/
- Sentiment analysis through python
  https://www.baeldung.com/getting_starte
  ed _with_twitter_datasets
- https://towardsdatascience.com/sentiment-
  analysis-with-python-part-1-5ce197074184
- YouTube:
  https://www.youtube.com/watch?v=KMEe2ni92rQ

BASE PAPER

Appreciation of Customer Satisfaction Through Analysis Facial Expressions and Emotions Recognition

Moulay Smail Bouzakraoui
Department of Informatics, Faculty of sciences Ibn Tofail University
Kenitra, Morocco
Bouzakraoui.smail14@gmail.com

Abdelalim Sadiq
Department of Informatics, Faculty of sciences Ibn Tofail University
Kenitra, Morocco
a.sadiq@uit.ac.ma

Abdessamad Youssfi Alaoui
National School of Computer Science and Systems Analysis Mohammed V University Rabat, Morocco
Abdessamad.youssfialaoui@gmail.com

***Abstract*— This article presents the modelling of a system that aims to predict the satisfaction of a customer through his emotions. This system must predict customer's behavior in the decision-making process. For this end, first we extract geometric features form customer's emotional faces, captured from local camera placed near the products. Then, to predict customer satisfaction, we have classified these features using adapted SVM classifier. The kinds of customer's satisfaction are satisfied, not satisfied and neutral. Our system shows a good performance, testing it on the JAFFE dataset.**

Keywords- Facial expression; customer emotion; face detection; customer behavir; Emotion recognition

I. INTRODUCTION

In today's world, the understanding and exploitation of human emotion are challenges in Marketing; it allows marketers to assimilate customer satisfaction and opinion about their products.

For a long time, companies have become increasingly interested in understanding customers' buying behavior and their relationship with products, including what attracts their attention and what remains unobserved [17]. By analyzing the nonverbal behavior of customers, we can recognize their emotions. Their perception of a product leads us to the extraction of many clues: body language and tone of speech, but the type of product interested in the client is often indicated by most facial expressions. Mehrabian said that the communication inter persons is represented by 55% of facial expression, 38% of the vocal part, and 7% of the semantic content [1].

The main objective of this work is to model an automatic system that aims to detect the nature of the facial expressions displayed by a client when he appreciates a product. This nature can be either positive or negative, or likes or dislikes, and so on. We can use this type of information to compute statistics on products and their exposure, we can also determine if the positive appreciation push the customers to buy the products. After that, we can have placed in the foreground products valued by positive appreciation, and present in other way products with negative appreciation. Another application to understand and recognize customers' product emotions supports the development of intelligent systems that can react to user preferences in a commercial context. This motivates us to propose a new system based on the analysis of the facial expressions able to deduce automatically the satisfactions of the customers towards a product.

This paper is organized as follows: Section II presents some work done in the literature. Section III presents the proposed approach for customer satisfaction based on facial expression. Section IV presents the Experimental results. In section V, we will present a summary of our work, some conclusions and perspectives.

II. RELATED WORK

The study of positive and negative consumer emotions through facial recognition has emerged in marketing and advertising.

Emotion representation models are classified according to two main criteria: discrete emotion states and dimensional persistent emotion space. He notable proceeding with emotional valence and arousal continuous model [18] appears to be increasingly suitable for assessing customer gratefulness, as it characterizes outward appearances as indicated by the charming (valence) improvements on a hub and initiation (arousing) on the other.

Ekman and Friesen defined six universal basic emotions: happiness, sadness, fear, disgust, surprise, and anger [2], they have produced FACS-(Facial Action Coding System for describing visually distinguishable Facial motion .Using the FACS, Action parameters designated to each of the expressions which classifies the Human Emotions [2].

For a shorter time and on a more marketing than commercial plan, the facial expressions can also be used to measure the advertising attention and the reactions of the individuals exposed to a message. The analysis or measurement of the facial expression can be used in particular to detect and analyze

978-1-7281-1232-9/19/$31.00 ©2019 IEEE

the emotions felt during the exposure to an advertisement. Eye tracking gives good attention when navigating through advertisements. Recently, there are a lot of works used The eye tracking in costumers' emotion recognition [4]. Shergill et al. [5] develop a computerized sales support system that relies heavily on facial expression recognition technology. Desmet [6] achieved the products' characteristics, which effect the state of emotion for users. Hi used 21 classes of emotions.

The analysis of facial expression is a very broad field that generates a multitude of methods, algorithms and techniques. All of these methods can be divided into two parts: The first groups methods based on the characteristics of the whole face (holistic methods). While the second group brings together methods that only focus on a few more important parts of the face (geometric methods), that works on Facial Action Units. Several algorithms have been proposed for the first group like Principle Component Analysis [7], Local Binary Pattern [8], Independent Component Analysis [7], and Linear Discriminate Analysis [9]. For the second group, we find Elastic Bunch Graph Matching [10], Active Shape Model [11] and Active Appearance Model [11] [12].

III. Our Approach

In this section, we present our algorithm to recognize customers' emotions of product. Customer satisfaction recognition through Analysis of facial expressions and Emotions consist in first step to acquire face image, second extracted expression features, especially geometric feature, to detect emotion. These features are obtained by transforming the input image into geometric primitives such as points and curves. By the way, it located distinctive patterns such as eyes, mouth, nose and chin, and measures their relative position, width, and other parameters like distance between noses, eyes, mouth… The challenge is to select the most relevant distances that can differentiate facial expression related to the emotion and customer's satisfaction. Finally, we use distances to classify emotions as a positive, negative or neutral appreciation towards a product. We analysis the results of classification to generate values to enterprise. Customers will use these values to extract how the product is perceived.

Our proposed model is illustrating in figure1:

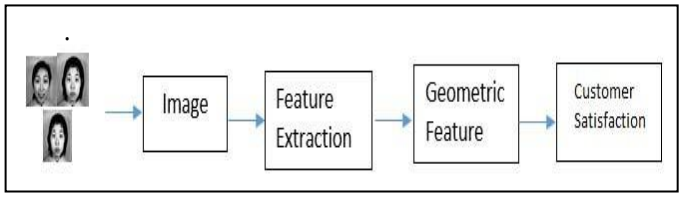


Figure1: Customer emotions analysis process

A. *Features extraction (Geometric features)*

This step consists to illustrate the shape and location of facial landmarks (e.g. mouth, eyes, eyebrows, nose, and chin) to represent the emotional face image. Although, the deformation of the facial landmarks changes the facial expressions. To locate the facial landmarks point we choose the model proposed by [9] shown in Figure (2).

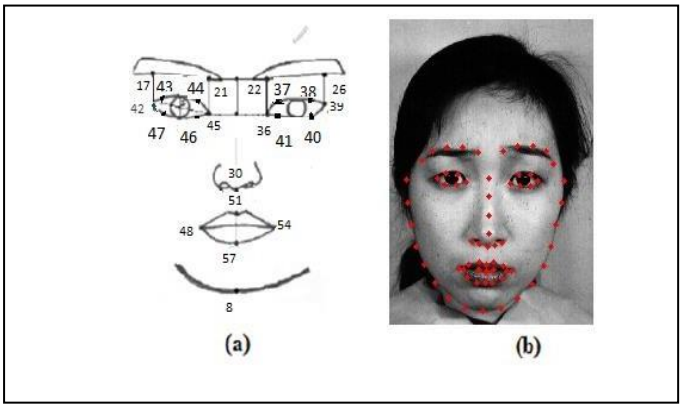


Figure2: (a) Facial model depicting the position of the chosen landmark point, (b) Landmark point on the test subject face

However, we used twenty-two facial landmarks points to calculate nineteen distances. Therefore, a vector of nineteen values will represent each face. The selected distance are:
For the eyes: six features
Left eye width, Right eye width, Right eye height, left eye height, Right eye height, left eye height
For the mouth: two features
Mouth Width, mouth height
For the chin: two features
Chin to nose distance, Chin to mouth distance
For the eyebrow: seven features
Left eyebrow width
Right eyebrow width
Outer left eyebrow distance to eye
Inner left eyebrow distance to eye
Outer right eyebrow distance to eye
Inner right eyebrow distance to eye
Inner left eyebrow distance to eye
For the nose: two features
Nose distance to left eye
Nose distance to right eye
In the current approach, we detected the landmarks points of the face for each image, we compute geometric features relative Euclidean distances between landmarks points which will be fed as feature input vectors supervised classifier Support Vector Machine (SVM)
Considered distance set is illustrated in Table I.
Finally, all distances are concatenated and form the geometric feature set: fG = (D1; D2;….; D19), formed by nineteen values.

Distance	Description	Formula	Distance	Description	Formula
D1	Mouth Width	d(X48,X54)	**D11**	Right eye height	d(X37,X4)
D2	Mouth height	d(X51,X57)	**D12**	Right eye height	d(X38,X40)
D3	Chin to mouth distance	d(X8,X57)	**D13**	Outer left eyebrow distance to eye	d(X17,X42)
D4	Chin to nose distance	d(X8,X30)	**D14**	Inner left eyebrow distance to eye	d(X21,X45)
D5	Left eyebrow width	d(X17,X21)	**D15**	Outer right eyebrow distance to eye	d(X26,X39)
D6	Right eyebrow width	d(X22,X26)	**D16**	Inner right eyebrow distance to eye	d(X22,X36)
D7	Left eye width	d(X42,X45)	**D17**	Inner left eyebrow distance to eye	d(x21,x45)
D8	Right eye width	d(X36,X39)	**D18**	Nose distance to left eye	d(X45,X30)
D9	Left eye height	d(X43,X47)	**D19**	Nose distance to right eye	d(X36,X30)
D10	Left eye height	d(X44,X46)			
D11	Right eye height	d(X37,X4)			

Table1: Definitions of Geometric features

B. Training and Classification

In this step, we classify geometric features, described in the last section, in three classes according to customer's satisfaction using SVM classifier.

To separate between two classes of Data, Support Vector Machine create a hyper plan with a maximum margin of hyper plan. This hyper plan maximizes the distance between two hyper plan [24]. The optimum hyper plan is which maximize the minimum margin between two sets. Therefore, the border training patterns are only taken to compute the hyper plan. These border are called support vectors. Furthermore, closest is lied to decision surface by using Support vector [14]. Recently studies have shown that support vector Machine illustrate higher accuracy in terms of data classification compared with other classifier [15].

The following equation define the equation of Hyper plane:

$$\sum_{i=1}^{N} \alpha_i d_i k(x, x_i) = 0 \quad (1)$$

Or

$$K(x, x_i) = M^T(x) M(x_i)$$

Demotes that inner-product kernel kernel [16], [17] is the inner product between by the input vector x and input pattern xi relating to the ith example.

Where

$$W = \sum_{i=1}^{N} \alpha_i d_i M(x_i)$$
$$M(x) = [M_0(x), M_1(x), \ldots\ldots, M_{m1}(x)]^T \quad (2)$$

$$6xM_0(x) = 1$$

We had used RBF (radian basis function kernel) function as a SVM Kernel.

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (3)$$

The following equation define the function Q(α), the Lagrange multipliers {αi} for i = 1 to N maximize this function.

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j K(x, x_i) \quad (4)$$

Subject to the following constraints:

$$\sum_{i=1}^{N} \alpha_i d_i = 0$$
$$0 \le \alpha_i \le C \text{ for } i = 1, 2, \ldots\ldots, N \quad (5)$$

The following equation is used to compute the ideal of the Lagrange multipliers comparing to the linear weight vector W0.

We used the following equation to find the ideal values of the Lagrange multipliers corresponding to the linear weight vector w0

$$W_0 = \sum_{i=1}^{N} \alpha_{0,i} d_i M(x_i) \quad (6)$$

$M(x_i)$ is the image initiated in the feature space due to xi

In our proposed approach, we used distances illustrated in table 1. By the way, we had used a mask to detect point. After that, we computed distances. Therefore, the input of our SVM classifier will be vectors, which contain nineteen distances. Alternatively, a vector of distances will represent each emotional face.

In this paper, we classify our features by using a multi class Support Vector Machine (SVM). Thus, we used the decision function to each class trained. After that, we took the largest values of the outputs of decision functions. The following function show how we compute output of multi class SVM classifier.

$$output = \arg\max_i (f_i) \qquad i = 1,2,3. \quad (7)$$

C. Experimental Results

We used Jaffe dataset to test our proposed method. By the way, Jaffe dataset planned and assembled by Michael Lyons, Miyuski Kamachi and Tiro Cyoba. The photos were taken at the psychology department in Kplyushu University and contains 213 images posed by 10 female Japanese. It is used for facial expression recognition. So, we adapted this dataset to our algorithm. We classified these images to three classes.

The first one contains images for satisfied persons, the second is for non-satisfied persons and the third class is for neural persons.

We used RBF kernel methods for SVM classifier.Then, we computed The ROC Curve in figure3 and 4 shows the performance of our proposed method

(8)

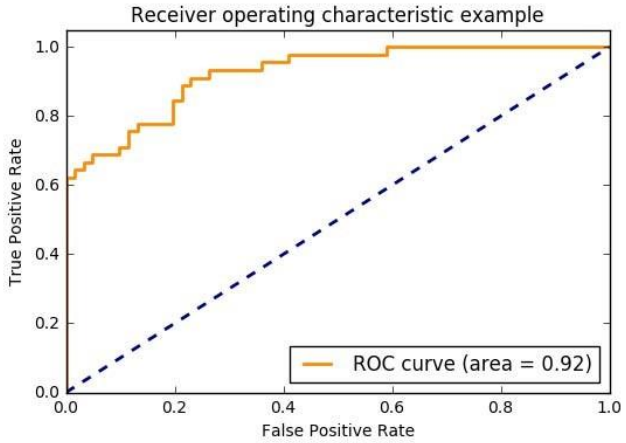


Figure 3 : Global ROC curves for classification method

(9)

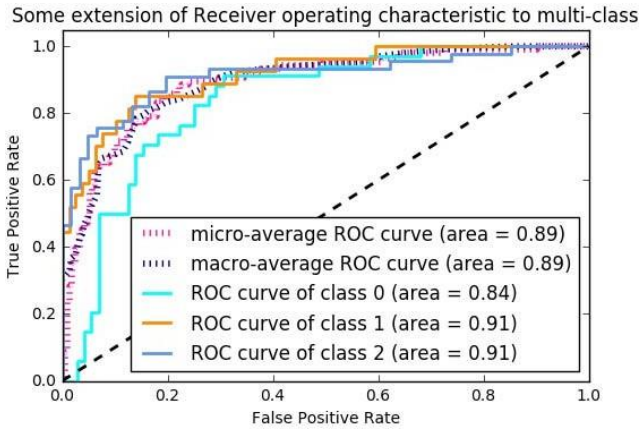


Figure 4 : ROC curves for classification satisfaied, not satisfaied ans neutral

IV. CONCLUSION AND PERSPECTIVES

In this article, we have present an algorithm to recognize customer's satisfaction. The main idea is to classify various facial expressions using Support Vector Machine (SVM). For this end, we extract geometrics features that represent distances between facial landmarks points. Next, we generate a vector of nineteen values that represent customer's face expression. To improve the performance of our approach, we applied our algorithm using JAFFE database. Our algorithm achieves a performant results. The future work, will focused on the use of adequate customers' shopping behaviour database, and enrich our approach by others features like eyes gaze, eyes tracking Authors and Affiliations.

REFERENCES

[1] "Emotion Classification Using Facial Expression", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 7, 2011.

[2] Hyisung Hwang, David Matsumoto, "Evidence for the Universality of Facial Expressions of Emotion",2014.

[3] Tobias Otterbring, Erik Wästlund, Anders Gustafsson, "Eye-tracking customers' visual attention in the wild: Dynamic gaze behavior moderates the effect of store familiarity on navigational fluency", 2008

[4] Shergill et al Computerized sales assistants: the application of computer technology to measure consumer interest – a conceptual framework, Journal of Electronic Commerce Research, VOL 9, NO 2, 2008

[5] Desmet P. M. A. and P. Hekkert, "The basis of product emotions", In W. Green and P. Jordan (Eds.), Pleasure with Products, Beyond usability, pp. 60-68, London.

[6] Ian T. Jolliffe, Jorge Cadima, "Principal component analysis: a review and recent developments", 2012 .

[7] Sarah Adel Bargal, Emad Barsoum, Cha Zhang, Cristian Canton Ferrer, "Emotion recognition in the wild from videos using images", 2016.

[8] Aloa shrama,Kudlip k.paliwal "Linear discriminant analysis for the small sample size problem: an overview", 2015.

[9] Pensyl William Russel, Xiaoping Min, Song Shuli Lily "Facial Recognition and Emotion Detection in Environmental Installation and Social Media Applications", 2016.

[10] Reda Shbib, Shikun Zhou, "Facial Expression Analysis using Active Shape Model", Vol. 8 No. pp. 9-22 ,2015.

[11] Maja Pantic, Leon J.m. Rothkrantz, "Automatics Analysis of Facial Expressions: The State of the Art", IEEE transactions on pattern analysis and machine intelligence, Vol. 22, No. 12, December2000

[12] Kim, Sangwook, et al. "Deep learning of support vector machines with class probability output networks." Neural Networks (2014)F.

[13] bdat, C.Maaoui et A.Pruski : Gradient based method for static facial features localization. International Conference Visualization,Imaging and Image Processig , VIIP07, 2007.

[14] Deepak Ghimire, Sunghwan Jeong, Joonwhoan Lee,"Facial expression recognition based on local region specific features and support vector machines",2017

[15] Foody, Giles M., and Ajay Mathur."The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM.",2006

[16] Lima, Ana Carolina ES, and Leandro Nunes De Castro. "A multi-label, semi-supervised classification approach applied to personality prediction in social media." Neural Networks 58,2014.

[17] Sandra Maria Correia Loureiro ISCTE-IUL, Portugal,"Consumer-brand relationship brand relationship brand relationship": foundation and state- foundation and state-of-art, 2012

[18] Hatice Gunes, Björn Schuller, Maja Pantic, Roddy Cowie, "Emotion representation, analysis and synthesis in continuous space: A survey", 2012

,

.