

Improving User Experience with Multi-Dimensional Factors for Effective Interactions
in a Dynamic Environment

By

Wasura Dananjith Watterachchi

15001415

This dissertation is submitted to the University of Colombo School of Computing
in partial fulfilment of the requirements for the
Degree of Bachelor of Science Honours in Computer Science

University of Colombo School of Computing

35, Reid Avenue, Colombo 07,

Sri Lanka

July 2020

Declaration

I, W.D. Wattearachchi (15001415) hereby certify that this dissertation entitled “Improving User Experience with Multi-Dimensional Factors for Effective Interactions in a Dynamic Environment” is entirely my own work and it has never been submitted nor is currently been submitted for any other degree.

.....
Date Signature of the Student

I, Prof. K.P. Hewagamage, certify that I supervised this dissertation entitled “Improving User Experience with Multi-Dimensional Factors for Effective Interactions in a Dynamic Environment” conducted by W.D. Wattearachchi in partial fulfilment of the requirements for the degree of Bachelor of Science Honours in Computer Science.

.....
Date Signature of the Supervisor

I, Dr. K.H.E.L.W. Hettiarachchi, certify that I supervised this dissertation entitled “Improving User Experience with Multi-Dimensional Factors for Effective Interactions in a Dynamic Environment” conducted by W.D. Wattearachchi in partial fulfilment of the requirements for the degree of Bachelor of Science Honours in Computer Science.

.....
Date Signature of the Co-Supervisor

Abstract

In the field of Human-Computer Interaction (HCI), improving User Experience (UX) of mobile devices has become a necessity due to the emergence of smart technologies and with the popularity of using mobile devices in day to day life rather than traditional desktop systems. The main aim of this research is to develop a model for a mobile device that can suggest adaptive functionalities, based on the current user emotion and the context. To the best of our knowledge, there are not any systems which provide not only adaptive interfaces but also provides adaptive functionalities within a mobile device which will enhance the acceptability and usability of that particular system.

As a proof of concept, a keyboard named “Emotional Keyboard” was developed through five prototypes using Evolutionary Prototyping, iteratively. As the methodology, Action Research together with User-Centered Design (UCD) was followed which also included two user surveys. Initial decisions were taken after conducting the first survey and Prototypes 1, 2 and 3 have been developed and those have been evaluated with the participation of 40 users. Prototype 3 has incorporated an Artificial Neural Network (ANN) that was trained using the data collected during the evaluations of the Prototypes 1 and 2, which can decide the most optimal emotion by combining the emotions from facial expressions and text. Consequently, Prototype 4 and 5 were developed which can suggest the most affective function based on the emotion and the context (location, time and user activity), by incorporating the data that was collected by conducting the second survey and built a “Preference Tree” which consists the probabilities of choosing functions and also considering the frequently used functions. The evaluation of Prototype 4 and 5 have been carried out with the participation of 18 users, where an individual and general analysis has been performed and proved that with time the model was able to correctly suggest adaptive functions. This evaluation yields the conclusion of the research, thus paving the way to an “Adaptive System with User Control” to improve the acceptability and the usability of a mobile device which aligns with the research aim.

Preface

A novel approach and a model were introduced in this research in order to improve the User Experience (UX) of a mobile device to assist a user by providing more affective functionalities with the aid of affective computing and multimodal information such as emotions and context. According to the best of our knowledge, no research has been conducted by considering both emotions and context to provide not only adaptive interfaces but also to provide adaptive functionalities which contribute to the novelty of this research.

The design of the research incorporates Action Research with User-Centered Design (UCD) where Evolutionary Prototyping has been performed which introduces an innovative approach to conduct a Human-Computer Interaction (HCI) research. The first three prototypes have been developed by incorporating two state-of-the-art SDK/APIs such as Affectiva SDK (to detect emotions from facial expressions) and IBM Watson Tone Analyzer API (to detect emotions from the text), but the integration and evaluation of these three prototypes introduced an innovative way to combine emotions using these two tools in a practically applicable manner. To improve this further, an Artificial Neural Network (ANN) has been developed with my own investigation.

Before developing the next prototypes, a user survey was conducted and finally, Prototype 4 and 5 were developed to introduce a model named “Preference Tree” that was designed using the data from the survey. It is my own invention of storing probabilities about how generally the users would like to perform a particular task, in a tree. The probabilities were updated in a manner where for each user his/her own set of probabilities will be stored in his/her Preference Tree by also incorporating the concept named “Frequently used Functions” that has been introduced by the literature. Together, this research introduces a new model that suggests the users the most adaptive function thus improving the acceptability and usability of a mobile device.

Acknowledgement

I would like to express my sincere gratitude to my research supervisor, Prof. K.P. Hewagamage, Director of the University of Colombo School of Computing and my research co-supervisor, Dr. K.H.E.L.W. Hettiarachchi, Senior Lecturer of the University of Colombo School of Computing for providing me with continuous guidance and supervision throughout the research.

I would also like to extend my sincere gratitude to Dr. T.A. Weerasinghe, Senior Lecturer of the University of Colombo School of Computing and Dr. K.T. Karunanayaka, Senior Lecturer of University of Colombo School of Computing for providing feedback on my research proposal and interim evaluation to improve my study. I also take the opportunity to acknowledge the assistance provided by Dr. H.E.M.H.B. Ekanayake as the final year computer science project coordinator.

I appreciate the feedback and motivation provided by my friends to achieve my research goals. This thesis is also dedicated to my loving mother and sister who have been an immense support to me throughout this journey of life. It is a great pleasure for me to acknowledge the assistance and contribution of all the people who helped me to complete my research successfully.

Table of Contents

Declaration.....	i
Abstract.....	ii
Preface.....	iii
Acknowledgement	iv
Table of Contents	v
List of Figures.....	xi
List of Tables	xiii
List of Acronyms	xiv
Chapter 1 - Introduction	1
1.1 Background to the Research	1
1.2 Research Problem and Research Questions	3
1.2.1 Research Problem	3
1.2.2 Research Questions	4
1.3 Research Aim and Objectives	4
1.3.1 Research Aim.....	4
1.3.2 Research Objectives.....	5
1.4 Justification for the Research.....	5
1.5 Methodology	6
1.5.1 Diagnosing	7
1.5.2 Action Planning	7
1.5.3 Action Taking	8
1.5.4 Evaluating	8
1.5.5 Learning	8
1.6 Outline of the Dissertation	9

1.7 Delimitations of Scope.....	10
1.7.1 In Scope	10
1.7.2 Out Scope.....	11
1.8 Summary	12
Chapter 2 - Literature Review	13
2.1 Emotions	13
2.2 Emotion Detection	14
2.2.1 Approaches to Detect Emotions Using Facial Expressions.....	14
2.2.2 Approaches to Detect Emotions Using Text Analytics	15
2.2.3 Approaches to Detect Emotions Using Typing Pattern	17
2.2.4 Appropriateness of Different Methods to Detect Emotions When Using Mobile Devices	19
2.2.5 Best Methods to Detect Emotions When Using Mobile Devices	20
2.3 Context.....	21
2.4 Proposed Management Techniques When Designing an Adaptive System	21
2.5 Applications	22
2.5.1 EmotionPush	22
2.5.2 Situation Dependent Browser	23
2.5.3 AUBUE (Adaptive User-Interface Based on User’s Emotion)	24
2.5.4 Spatiotemporal Structure Learning with Samsung Galaxy S to Develop an Adaptive User Interface	25
2.5.5 Automatic Mobile Menu Customization System.....	26
2.6 Summary	27
Chapter 3 - Design	28
3.1 Justification for the Methodology.....	28
3.2 Design Overview	29
3.2.1 Conducting User Survey 1	29

3.2.2 Strengthen the Initial Decisions	30
3.2.3 Prototype Evaluation.....	31
3.2.4 Conducting User Survey 2	31
3.2.5 Final Analysis	32
3.2.6 Final Results.....	32
3.3 High-Level Architecture (Abstract Design).....	32
3.3.1 Context Model	33
3.3.2 User Emotion Model.....	34
3.3.3 Aggregator Model.....	35
3.4 Ethical Issues and Steps Taken	36
3.4.1 Ethical Issues	36
3.4.2 Steps Taken.....	36
3.5 Summary	37
Chapter 4 - Implementation.....	38
4.1 Prototype 1	38
4.1.1 Functionalities of Prototype 1	39
4.1.2 Emotion Detection Procedure Using Text	40
4.1.3 Emotion Detection Procedure Using Facial Expressions	42
4.1.4 Feedback Process	43
4.2 Prototype 2	44
4.3 Prototype 3	46
4.3.1 Developing the Artificial Neural Network (ANN)	46
4.3.2 Improving Android App	47
4.3.3 Feedback Process	49
4.4 Prototype 4	50
4.4.1 Preference Tree and the Registration Process.....	50
4.4.2 Retrieving Location of the User.....	51

4.4.3 Detecting the Current User Activity	53
4.4.4 Retrieving the Time	54
4.4.5 Emotion Detection	54
4.4.6 Demonstration of Suggesting the Adaptive Functions	55
4.4.7 Data for Evaluation	57
4.5 Prototype 5	59
4.5.1 “frequentFunctions” Field.....	59
4.5.2 Improved Method to Suggest the Adaptive Functions	60
4.6 Summary	62
Chapter 5 - Results and Evaluation	63
5.1 Prototype 1	63
5.1.1 Participant Selection and Experiment Setting.....	63
5.1.2 Evaluation Procedure	63
5.1.3 Analysis.....	64
5.1.4 Observations	65
5.1.5 The Need for Prototype 2.....	66
5.2 Prototype 2	66
5.2.1 Analysis.....	67
5.2.2 Comparing Prototype 1 and Prototype 2 Using Precision	68
5.2.3 The Need for Prototype 3.....	71
5.3 Prototype 3	71
5.3.1 Training and Validating the ANN.....	71
5.3.2 Confusion Matrix and Interpretation	72
5.3.3 Performance Measures of the Emotions “joy”, “sadness” and “anger”.....	76
5.4 Prototype 4	77
5.4.1 Participant Selection and Evaluation Procedure	77
5.4.2 Analysis.....	77

5.4.3 The Need for Prototype 5.....	78
5.5 Prototype 5.....	79
5.5.1 Individual Analysis.....	79
5.5.2 Overall Analysis.....	82
5.6 Summary.....	83
Chapter 6 - Conclusions	85
6.1 Conclusions about the Research Questions (Aims and Objectives).....	85
6.2 Conclusions about the Research Problem.....	87
6.3 Research Contribution	87
6.3.1 Contribution to CS.....	87
6.3.2 Contribution to HCI.....	88
6.4 Limitations.....	89
6.5 Implications for Further Research	89
References.....	90
Appendix A: User Survey 1.....	97
A.1 Questionnaire (Google Form).....	97
A.2 Results of User Survey 1.....	99
Appendix B: User Survey 2.....	102
B.1 Questionnaire (Google Form).....	102
B.2 Results of User Survey 2.....	109
Appendix C: Informed Consent Form.....	116
Appendix D: Code Listings	118
D.1 Code Listings of Prototype 1	118
D.2 Code Listings of Prototype 2	120
D.3 Code Listings of Prototype 3	121
D.4 Code Listings of Prototype 4	123
D.5 Code Listings of Prototype 5	132

Appendix E: Deciding the Thresholds	136
E.1 Deciding Threshold 1 (T1)	136
E.2 Deciding Threshold 2 (T2)	137
Appendix F: Prototype Evaluations	138
F.1 Prototype 1 Evaluation	138
F.2 Prototype 2 Evaluation	139
F.3 Prototype 3 Evaluation	140
F.4 Prototype 4 Evaluation	141
G.5 Prototype 5 Evaluation.....	142
Appendix G: Publications	145
G.1 Publication 1	145
G.2 Publication 2	151

List of Figures

Figure 1.1: Action research.....	7
Figure 2.1: Text analysis methods [4].....	17
Figure 2.2: Dividing the context into subcategories [4]	21
Figure 2.3: Architecture of AUBUE [6]	24
Figure 2.4: Architecture of spatiotemporal structure learning application [18]	25
Figure 3.1: Design steps of the research	29
Figure 3.2: High-level architecture	33
Figure 3.3: Context Model.....	33
Figure 3.4: User Emotion Model	35
Figure 3.5: Aggregator Model	35
Figure 4.1: Prototype 1- Demonstration and stacked notifications.....	39
Figure 4.2: Prototype 1 - Asking feedback	40
Figure 4.3: Prototype 1 - If user pressed NO	40
Figure 4.4: Prototype 1 - YES/NO count in Firebase database	43
Figure 4.5: Prototype 1 - A record with respect to a particular notification	44
Figure 4.6: Prototype 3 - ANN architecture.....	47
Figure 4.7: Prototype 3 - Feedback recorded format	49
Figure 4.8: Prototype 4 - Register screen.....	51
Figure 4.9: Prototype 4 - User database record.....	51
Figure 4.10: Prototype 4 - Edit location details	51
Figure 4.11: Prototype 4 - After setting the location details.....	52
Figure 4.12: Prototype 4 - Updated database after the location setting	52
Figure 4.13: Prototype 4 - Suggesting functions	55
Figure 4.14: Prototype 4 - Traversed branch	56
Figure 4.15: Prototype 4 - When user pressed NO	57
Figure 4.16: Prototype 4 - “adaptiveChoices” when pressed YES	58
Figure 4.17: Prototype 4 - “adaptiveChoices” when pressed NO.....	58
Figure 4.18: Prototype 4 - “adaptiveFunctionsEvaluation” to track predictions	59
Figure 4.19: Prototype 5 - “frequentFunctions” field	59

Figure 4.20: Prototype 5 - Example values for “frequentFunctions”	61
Figure 5.1: Prototype 1 - Face emotions analysis	64
Figure 5.2: Prototype 1 - Text emotions analysis	65
Figure 5.3: Prototype 2 - Face emotions analysis	67
Figure 5.4: Prototype 2 - Text emotions analysis	68
Figure 5.5: Prototype 3 - Training accuracy vs. Validation accuracy	72
Figure 5.6: Prototype 3 - Training loss vs. Validation loss	72
Figure 5.7: Prototype 3 - Confusion matrix for evaluation.....	72
Figure 5.8: Prototype 4 - User 1 evaluation.....	78
Figure 5.9: Prototype 5 - Overall analysis	82

List of Tables

Table 1.1: Objectives of Research	5
Table 2.1: Datasets for Text Analytics [4].....	16
Table 2.2: Justify the Suitable Methods to Detect Emotions in Mobile Devices	19
Table 3.1: Methods to Perceive Context Parameters and Descriptions	34
Table 4.1: Prototype 1 - Text Emotions Array	41
Table 4.2: Prototype 1 - Face Emotions Array	42
Table 5.1: The Need for Prototype 2	66
Table 5.2: Prototype 1 - Precision Values	69
Table 5.3: Prototype 2 - Precision Values	69
Table 5.4: Prototype 3 - Precision for Each Emotion	74
Table 5.5: Prototype 3 - Recall for Each Emotion.....	74
Table 5.6: Prototype 3 - Specificity for Each Emotion.....	75
Table 5.7: Prototype 3 - F-Score for Each Emotion	76
Table 5.8: Prototype 3 - Performance Measures for “joy”, “sadness” and “anger”	76

List of Acronyms

ANN	Artificial Neural Network
AOSP	Android Open Source Project
API	Application Programming Interface
AUBUE	Adaptive User-Interface Based on User's Emotion
CAFÉ	Child Affective Facial Expression
FACS	Facial Action Coding System
HCI	Human-Computer Interaction
JAFFE	Japanese Female Facial Expressions
KDEF	Directed Emotional Faces
OCC	Ortony, Clore, and Collins
PANA	Positive Activation – Negative Activation
SDK	Software Development Kit
SVM	Support Vector Machine
UCD	User-Centered Design
UI	User Interface
UX	User Experience

Chapter 1 - Introduction

This chapter provides an introduction to the research by justifying the motivation. Section 1.1 will discuss the background to the research, followed by Section 1.2 which will introduce the research problem and two research questions. Research aim and the objectives will be listed in Section 1.3 by relating to the research questions identified and the justification for the research will be discussed in Section 1.4. An overview of the research methodology will be described in Section 1.5 followed by the outline of the dissertation in Section 1.6. The chapter concludes by introducing the delimitations of the scope in Section 1.7.

1.1 Background to the Research

The term User Experience (UX) has a broad meaning which describes the feeling or the satisfaction which a user acquires when performing their day to day activities using a particular system or a product [1]. Improving UX of computer systems has become important and popular in the field of Human-Computer Interaction (HCI) because it will provide users with a positive and comfortable experience [2]. Improving UX leads to the concept of “Affective Computing” which provides a system that can adapt to the users’ needs by considering user emotions and the emotional communication between the user and the system [3].

As explained in “Affective Computing” [3], emotions can be considered as a factor to be used when designing a usable system that can provide a more positive and comfortable experience to the user [4]. The intelligence from emotions has been used such as when designing intelligent robots [5], user interfaces [6] and appliances for households [7]. Also, the emotions of a person depend on several factors such as relationships, mental health, personality, and context [4].

Several emotion models were introduced such as Ekman’s model [8], Plutchik’s wheel of emotions [9] which are known as emotion categories and models such as circumplex model which are known as emotion dimensions. Emotion detection can be

done using several methods such as using facial expressions [10], gestures [11], gaze direction [12], voice [13], text analytics, typing pattern or keystroke dynamics [14] and brain signals [15]. One can apply these methods individually or combine them thus leading to multimodal emotion detection [16]. Among them, emotion detection through facial expressions recognition and text analytics are most suitable when using mobile devices which will be justified in [Section 2.2.5](#).

Apart from the emotions, by identifying the context and the environment of the user, more dynamic and affective system functionalities such as changing the colour, font size, widgets of the system has been performed in previous research work [17], [18]. Context can be subdivided into categories based on the user, platform and environmental aspects [19]. User aspects such as user profiles, cognition, user activities and environmental aspects like location, time and weather are commonly considered when developing adaptive systems [17]–[19]. It has proven that multimodal information can be perceived using multiple sensors to accomplish human-like decision making based on data and these inputs should be processed and analyzed by consolidating them which leads to a context-dependent model [20].

There are several management techniques [21] to do the adaptation of user interfaces such as a system where the user can change the settings manually, adaptable with system support, adaptive with user control, and fully adaptive. In the fully adaptive technique, the entire management process is done by the system which does predictions without disturbing the user by considering the user profile and other variables but this research focusses on developing a system that is “Adaptive with User Control”.

Due to the convenience of smart technologies, the incorporation of functionalities to enhance affective user experience is more popular these days by embedding them into mobile devices [22], [23] rather than embodying them in desktop systems [24] and many applications were developed by taking user emotions and the context into account to produce effective and affective usable systems through personalization and customization [4]. The literature review will discuss some of these applications (Refer [Section 2.5](#)) such as EmotionPush [22], Situation-dependent browser [24], AUBUE

(Adaptive User-Interface Based on User's Emotion) [6] and an adaptive user interface based on spatiotemporal structure learning for Samsung Galaxy S [18]. Some of these applications were not able to give high performance for some emotions such as neutral and fear [22] while some have asked the users to manually enter the emotions [18].

To facilitate further improvement, models could be created to gather user data and emotions which could then be stored in a profile. It can later be used to predict the most affective action to be performed by the system based on the context data and also to provide affective feedback [4]. Considering the above, there is a need to introduce a system that can provide more affective functionalities to improve the UX of a mobile device while considering user emotions based on the context in real-time. (Refer [Section 1.4](#) for further justifications)

1.2 Research Problem and Research Questions

1.2.1 Research Problem

There is a need (which will be explained in [Section 1.4](#)) to improve the User Experience (UX) of a mobile device in order to assist a user by providing more affective functionalities to a user with the assistance of affective computing and multimodal information.

To the best of our knowledge (Refer [Section 2](#) for Literature Review), there are not any systems which provide not only adaptive interfaces but also provides adaptive functionalities within a mobile device, by considering the user emotion and the context parameters such as location, time and user activity. Also, detecting emotions through both facial expressions and text analytics has not been carried out for mobile devices and most of the researches have used user emotions or context alone to provide adaptive user interfaces.

1.2.2 Research Questions

1. How to aggregate different emotion detection techniques to decide the most meaningful emotion?

Most of the researches were carried out to detect emotions using facial expressions and text analytics but the combination of these two was not done. When these two techniques are combined, the most meaningful decision about the user emotions can be taken when using mobile devices which will be justified in [Section 2.2.5](#).

2. How a new model can be developed to do optimal decision making in order to perform affective interactions by considering user emotions based on the context?

Aggregating the emotions of the user based on the context was not done simultaneously before. According to the literature (Refer [Section 2.5](#)), there are several models where the context has played a main role when deciding the adaptive action to be taken in the system interface. By combining the emotions and context, a new architectural model can be introduced where more affective and usable decisions can be taken and executed in a system, which will not only be limited to the user interface.

1.3 Research Aim and Objectives

1.3.1 Research Aim

Developing a model for a mobile device which will have the ability to take more optimal decisions about the functionalities, by considering user emotions based on the current context in real-time, while also enhancing the acceptability and usability of that particular mobile device.

1.3.2 Research Objectives

The following objectives will be covered under the identified research questions are shown in Table 1.1.

Table 1.1: Objectives of Research

Research Questions	Objectives
1. How to aggregate different emotion detection techniques to decide the most meaningful emotion?	<ul style="list-style-type: none">• Detect user emotions through facial expressions.• Detect user emotions through text analytics.• Combine 1st and 2nd steps and get the most optimal emotion.
2. How a new model can be developed to do optimal decision making in order to perform affective interactions by considering user emotions based on the context?	<ul style="list-style-type: none">• Detect the context of the user by considering the physical location, time and user activities.• Develop a model that can make a decision to provide affective functionalities using the emotion based on the context.

1.4 Justification for the Research

When justifying the first research question, [Section 2.2](#) will discuss the methods of detecting emotions through facial expressions [10] and text analytics [14]. Further, [Section 2.2.4](#) justifies the inconvenience of using other methods such as typing pattern analysis [14], voice [13], body gestures [11], gaze direction [12] and brain signals [15] to detect emotions when using mobile devices, thus confirming facial expressions and text analytics might be the only ways to detect emotions in this scenario. The limitations when practically detecting emotions through these two techniques should be tested using the state of the art APIs and SDKs [25], [26] and a method to get the most meaningful emotion by combining these two techniques must be investigated which will lead to multimodal emotion detection.

The second research question investigates a way to incorporate the context parameters such as the time, location and user activities along with the emotion when deciding the

most optimal decision to suggest an adaptive function to the user. The literature shows different types of applications such as Situation Dependent Browser [24], Adaptive Interface using Spatiotemporal Structure Learning [18] and Automatic Mobile Menu Customization [27]. The concepts from these applications will justify the way to develop an adaptive system with user control [21] by paving the way to find answers to the second research question.

Due to the emergence of smart technologies, mobile devices have become more popular than traditional desktop systems [28], and it has proven that improving UX in mobile devices has vital importance [23]. The motivation for the main research aim to develop a system for a mobile device which can take more optimal decisions about the functionalities, by considering user emotions based on the current context in real-time has come from this necessity to improve the UX in mobile devices.

1.5 Methodology

The problem covers under this research is, not having a proper way to combine emotions from different modalities in a suitable manner to integrate with the context to increase the User Experience (UX) of mobile devices by providing adaptive system functionalities. The research methodology used here is **Action Research** [29] with **User-Centered Design (UCD)** [30] where every step from requirement gathering to evaluation was performed with user participation. User participation will happen through prototype evaluation with Evolutionary Prototyping [31] and the usage of two questionnaire surveys.

In the field of Human-Computer Interaction (HCI), User-Centered Design (UCD) helps to understand users more within a context by iteratively using their participation in each iteration. By integrating it with Action Research (as shown in Fig. 1.1), more descriptive and meaningful interpretation of decisions can be done while carrying out this research as an iterative process as follows. [Section 3.1](#) will explain more justifications to use the above techniques in this research.



Figure 1.1: Action research

Each step of the research methodology depicted in Figure 1.1 can describe as follows.

1.5.1 Diagnosing

In the first iteration, since the prior knowledge is not available, the Diagnosing has been done based on the literature by identifying the problems in the existing approaches. After the first iteration, in each iteration, Diagnosing was done based on the decisions taken by analyzing the feedback from the Learning stage.

1.5.2 Action Planning

The current iteration was planned from the problems which were identified through Diagnosing. Planning the design and participants selection (selecting the focus group) were done in this phase.

1.5.3 Action Taking

User-Centered Design (UCD) Process consists of four steps [30] as follows.

1. User Surveys.
2. Analysis of Requirements.
3. Prototyping.
4. Assessment of Usability

In this stage, the development of User Surveys (Questionnaires and Interviews) was done to develop a platform to identify user requirements while strengthening the decisions to be taken in this research. Prototypes were developed to eliminate the mismatch between the user's mental model and the designer's mental model which will pave the way to develop a more usable system. This is the first phase where the User-Centered Design (UCD) was incorporated into the Action Research Methodology.

1.5.4 Evaluating

The developed user survey or the prototype was evaluated with the participation of users. The speciality here is, from the first iteration, in each phase, the evaluation should be carried out rather than doing it at the end of the research.

1.5.5 Learning

After the evaluation, the analysis was carried out by considering both quantitative and qualitative data. The feedback and the results from the observation were fed to the Diagnosing phase to start the next iteration.

Section 3 will discuss this more with the Research Design.

1.6 Outline of the Dissertation

This dissertation consists of six chapters as described below.

Chapter 1 - Introduction

This chapter provides an introduction to the research by discussing the background to the research, followed the research problem and two research questions, research aim and objectives, the justification for the research, an overview of the research methodology and concludes by introducing the delimitations of the scope.

Chapter 2 - Literature Review

This chapter discusses a complete review of the literature related to this research by five sections. The main aim of this chapter is to discuss the basic concepts, the advantages and disadvantages of existing work which corresponds to the two research questions mentioned in Section 1.2.2.

Chapter 3 - Design

This chapter consists of four sections which are, the justification for the methodology, design overview, high-level architecture and ethical issues and steps taken.

Chapter 4 - Implementation

This chapter mainly discusses the aspects of the implementation of the proof of concept, which was done in each iteration of the research and will be presented in five sub-sections as five prototypes.

Chapter 5 - Results and Evaluation

This chapter mainly describes the details related to the results and the evaluation process of the research which was done iteratively. In each iteration, the results from the previous prototype evaluation were compared and the decisions were taken to improve the next iteration.

Chapter 6 - Conclusions

This chapter discusses the conclusions of the research that have obtained from the results of the analysis. It will provide the answers to the research questions by reviewing the aims and objectives, discusses the conclusions obtained for the research problem followed by the research contribution to the fields of Computer Science and Human-Computer Interaction. Consequently, the limitations will be explained and the chapter ends with giving details about the further implications of this research.

1.7 Delimitations of Scope

1.7.1 In Scope

The following will be covered under the scope of the project.

- **Use facial expression recognition and text analytics to detect user emotions.**

Section 2.2.5 will explain that although detecting user emotions through facial expressions and text analytics separately given good results, there are certain limitations when practically applying them [32], [33]. This motivates to find a method to practically check the appropriateness of these two techniques, while also combining them in a meaningful manner to eliminate the limitations and to obtain better results.

- **Consider analyzing English text when identifying emotions from text analytics.**

Only text in English was considered in this research when detecting emotions since there are well-known APIs that can provide this functionality which will be discussed in Section 2.2.2.

- **Consider physical location, time and user activities as the context.**

Many researches have defined the context in many ways [4] but the above-mentioned factors were considered only in here. The physical location was taken from GPS, and time was taken from the internet or system time and the

user activity was perceived from the accelerometer data. The raw data which were measured from these sensors should represent meaningful information and should be analyzed thoroughly. For example, based on the longitude and latitude, how can we say that the user is in office or home? How can we interpret data from the accelerometer to decide whether the user is running, walking or standing?

- **The project considered only the user emotions and context to decide the affective system functionalities.**

After retrieving the most optimal emotion based on the current context, how can we decide the best functionality to provide is questionable.

- **Only the adaptive system functionalities such as opening the music player, display popup messages, open chat applications were done in the system.**

When considering the functionalities in a system it is not practical to address all of them in a short period. To narrow down the scope, only a few frequently used functionalities were considered. These frequently used functionalities were identified based on User Survey 1 (Refer [Section 3.2.1](#)) that have done using questionnaires.

The main deliverable of this research is a model to make decisions based on user emotions and context. As a supplement, the evaluation was done through an Android app that was easily distributed among the target set of participants to test and obtained their feedback.

1.7.2 Out Scope

The following will not be covered under the scope of the project.

- **Improve the accuracy of facial expression detection or text analytics.**

Since there are many accurate APIs and research work related to those (which will be discussed in [Section 2.2.1](#), [Section 2.2.2](#) and [Section 2.2.5](#)),

reimplementation or modification was not done to detect emotions from facial expressions or text individually.

- **Use gesture recognition, gaze detection, voice, typing pattern, and keystroke analysis to detect user emotions.**

Section 2.2.4 and Section 2.2.5 provide the justifications to choose only facial expressions and text data when detecting emotions. Also, the main contribution of this research is developing a model that will give affective feedback. Therefore, the emphasis will be placed on that.

- **Factors such as physical characteristics of the user, sensory abilities, cognition will not be considered when deciding the context.**

The scope was narrowed since the time is a limiting factor.

- **Manipulation of the operating system will not be done when deciding the functionalities.**

Operating systems such as Windows are not open source. Therefore, it is hard to manipulate them. As a proof of concept an Android app was implemented which was distributed among users and evaluation was done.

1.8 Summary

This chapter started by introducing the background concepts related to the research by giving an overview of this research domain. Then, the research problem was precisely defined and the research questions were introduced. Next, the research aim was introduced followed by research objectives explained by mapping to the two research questions. The justification section discussed the knowledge gap in this research and how to fulfil those gaps in this research. Methodological details of this research process illustrated in the methodology section. Hereafter, this chapter stated brief outline details about each chapter of this dissertation. This chapter concluded by discussing the delimitations of the scope by defining the boundary between the in scope and out scope.

Chapter 2 - Literature Review

This chapter discusses a complete review of the literature related to this research through five sections. The main aim of this chapter is to discuss the basic concepts, the advantages, and disadvantages of existing work related to the two research questions mentioned in Section 1.2.2. The basic concepts behind “Emotions” will be discussed in Section 2.1, followed by Section 2.2 which will discuss “Emotion Detection” techniques by justifying the need for the first research question. The facts to justify the second research question will be covered next by introducing the term “Context” in Section 2.3, followed by Section 2.4 which will discuss the “Different management techniques when designing an adaptive system”. The literature review concludes with Section 2.5 by describing the “Applications” which discusses the related work to this research and the basic concepts behind them.

2.1 Emotions

User emotions can be defined either using Emotional Categories or Emotional Dimensions [34].

In Emotional Categories, an emotion will only have one discrete label which will represent only one value. Ekman’s model [8] which constitutes six discrete emotions such as Anger, Disgust, Fear, Happiness, Sadness, and Surprise is widely used among many types of research. Also, Plutchik’s wheel of emotions is another example that considers the emotions as four couples – Joy-Sadness, Anger-Fear, Trust-Disgust, and Anticipation-Surprise [9].

In Emotional Dimensions, emotion will be represented as a collection of many emotions in a 2D or a 3D form [34]. Examples for 2D models are Circumplex Model [35] which states an emotion can be described as a linear combination of many using the two dimensions: Valence - Arousal and PANA (Positive Activation – Negative Activation) [36] which divides emotions into positive and negative.

Most of the researchers have used Emotional Categories rather than Emotional Dimensions, and among the categories, Ekman's model is the most popular classification [22], [32], [34], [37].

2.2 Emotion Detection

Several methods can be used to detect user emotions such as using facial expressions [10], text analytics and typing pattern analysis [14], voice [13], body gestures [11], gaze direction [12] and brain signals [15]. One can detect emotions using these methods individually which is the traditional way but more accurate and improved results can be obtained when two or more of the above methods are combined thus leads to multi-modal emotion detection [16].

2.2.1 Approaches to Detect Emotions Using Facial Expressions

Communication can be done verbally or non-verbally. 55% of the message is expressed through non-verbal communication using facial expressions [38]. Researchers have stated, "*Your Face speaks louder than your voice*" thus confirming the fact that the best way to analyze user emotions is through facial expressions [39]. Identifying visible factors in muscle actions can be performed using the Facial Action Coding System (FACS) which is a popular method when identifying emotions through facial expressions [40].

Detecting emotions through facial expressions can be done in four steps as image acquisition, preprocessing of images, extracting and selecting features, and predict the emotion [10], [37], [41].

Both static identification and active identification of emotions can be done by acquiring data as static images and videos accordingly. Data sets like KDEP (Karolinska Directed Emotional Faces) [10] and CAFÉ (Child Affective Facial

Expression) [37] and databases like Bosphorus and JAFFE (Japanese Female Facial Expressions) [41] are popular when acquiring a large number of images.

Preprocessing of images was done to convert images with different resolutions to the same size (with the same dimensions) and to optimize the brightness and grey levels to increase the efficiency and accuracy of the further processing. JAFFE data set consists of images with the same size while the Bosphorus data set consists of images with different sizes which has to be preprocessed [41].

When extracting features, most of the researches have focused on extracting regions of interest such as eye region, mouth region and auxiliary region [41]. Mostly Viola-Jones detector was used when identifying these regions of interest [10], [37]. To measure the distances between identified facial points Mahalanobis distance [37] or Euclidean distance [41] has been used.

To do the predictions and identify the emotions, models like Support Vector Machines (SVM), k-NN, Bagged Trees, RUSBoost Trees were used [10], [12], [41].

Since detecting emotions through facial expressions has given high accuracies, several APIs and SDKs such as Microsoft Cognitive Services, Affectiva, Emotient, Kairos, Sky biometric application, Eyris EmoVu, OpenFace, FaceReader 6.1, Nviso and Vision API [26], [25] have been introduced.

2.2.2 Approaches to Detect Emotions Using Text Analytics

Statement of emotions can be done either using vocabulary-based methods or affective terms-based methods when considering text data. The semantic meanings of words such as sad, love, hate, happy are considered in vocabulary-based methods, while words such as ugh, eww, yuhuu are considered in affective terms-based methods [34].

As shown in Table 2.1, there are datasets that can be used to train a model to analyze emotions from text data. Most datasets have categorized emotions according to Ekman's model [8] and some of them are free and open source.

Table 2.1: Datasets for Text Analytics [4]

Dataset	Description	Emotion Categories	Availability
AffectiveText	News headlines	Ekman's Model	Free
Blogs	Sentences from blogs	Ekman's Model	Contact Authors
DailyDialogs	Conversation data	Ekman's Model	Free
Tales	Sentences from fairy tales	Ekman's Model	Free
TEC	Twitter hashtags	Ekman's Model	Free
Emotion-Stimulus	Sentences annotated with emotion and cause	Ekman's Model + shame	Free
CrowdFlower	Tweets	Non-standard	Free
ISEAR	Questionnaires answered by people from different cultures	Non-standard	Free
Electoral-Tweets	Tweets on elections	Non-standard	Free
EmoInt	Tweets	Non-standard	Free
fb-valence-arousal	From Facebook posts	Valence, Arousal	Free
Neviarouskaya et al.'s dataset 1	Stores in 13 different categories	Non-standard	Not sure
Neviarouskaya et al.'s dataset 2	Sentences from diary-like posts	Non-standard	Not sure
Alm's dataset	Sentences from fairy tales	5 from Ekman's Model	Not sure
Aman's Dataset	Sentences from blogs	Ekman's Model + neutral	Not sure

There are two main approaches to carry out experiments on text analytics which are rule-based approaches such as corpus-based and dictionary-based, and machine learning-based techniques such as using classifier algorithms, deep learning-based techniques and supervised feature-based learning techniques [34], [42], [43]. Fig 2.1 shows some well-known approaches for text analytics [4].

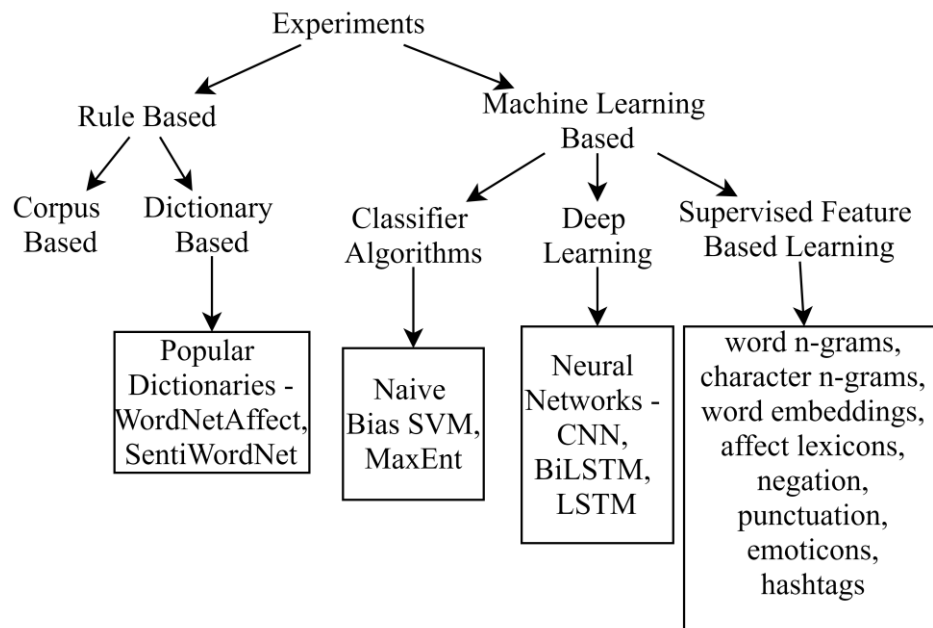


Figure 2.1: Text analysis methods [4]

There are some popular APIs and SDKs to detect emotions from text such as Tone Analyzer, Receptiviti, BiText and Synesketch [25] to detect emotions such as joy, fear, sadness, anger, analytical, confident and tentative tones.

2.2.3 Approaches to Detect Emotions Using Typing Pattern

Behavioural factors for emotion recognition can be identified using a typing pattern which can be mainly analyzed using keystroke dynamics and movement of the mouse [14], [44], [45].

Experiments were performed utilizing two techniques. The first technique was, participants were asked to read a paragraph, and with the emotion which they have perceived, they were asked to type something using a keyboard [46]. The second approach was, the participants were shown a video and again they were asked to type something with the emotion that generated while watching the video [14]. Assume, somebody who does not feel joy or sadness after viewing a video would imagine by communicating a fake feeling. This is a fundamental downside of utilizing these techniques where the participants can control their emotions as to suit the normal consequences of the experiment [4].

Few examinations were carried out by consolidating the above two techniques to analyze the mouse movements and the scrolling pattern. After watching the video, participants were asked to scroll down a paragraph while reading it and to type something to conclude the experiment. Assume the participant felt some emotion after watching the video. By the time when scrolling has come to an end, that emotion might have faded away. This leads the researcher to detect only neutral emotions which would have given bias results [14].

During these experiments, participants behaviour were mapped into the following six activities [14].

1. Participant types in the usual way with the usual speed.
2. Participant types at a speed higher than his/her usual speed.
3. Participant types at a speed lower than his/her usual speed.
4. Participant uses backspace key time to time.
5. Participant uses irrelevant keys from time to time.
6. Participant does not utilize the keyboard.

By examining the above behaviours, researches can answer questions like, “*What is the typing speed of a particular person?*”, “*How many numbers of characters were typed in a five seconds interval?*”, “*How much time it took to type a given paragraph or text content?*”, “*How often and how many times the backspace was used?*”, “*What is the time gap between the user is staying idle and typing?*” [46]. Additionally, average times to perform left, right and double mouse clicks, scrolling using the mouse, traverse between cursor x-axis and y-axis, key down to up, up to down and down to down, regular keypress, enter key, arrow key, backspace and function key press were able to analyze [14].

Basically, researchers were able to identify the emotion level of the user as positive, negative or neutral by the above examinations [46]. As the conclusions from the above experiments, slips, mistakes or errors happen due to the clicking of irrelevant keys when the user has negative emotions. To correct them backspace key was used often, which results in decreasing the usual typing speed of that particular person. When the

user is in a positive situation his/her typing speed was increased often. If he/she is in neutral mode, the usual typing speed can be seen.

2.2.4 Appropriateness of Different Methods to Detect Emotions When Using Mobile Devices

At the beginning of [Section 2.2](#), several methods were highlighted which can be used to detect user emotions such as facial expressions [10], text analytics and typing pattern analysis [14], voice [13], body gestures [11], gaze direction [12] and brain signals [15], but all of them are not suitable when detecting emotions using a mobile device. The reasons can be justified as follows.

Table 2.2: Justify the Suitable Methods to Detect Emotions in Mobile Devices

Method	Justification
Typing pattern	The typing pattern does not differ only because of the emotions. Gender, text familiarity, age of the user are also a fact which overrules the typing pattern of a particular person [47], [48]. It is clear that typing pattern will not be a reasonable method to be considered when detecting user emotions.
Voice	Voice is rarely used when interacting with a mobile device and it has proven that when expressing a message only 38% is covered from vocal part and 55% is covered by facial expressions [38].
Gaze Direction	When detecting emotions using gaze direction it is needed to consider action units that represent specific facial muscles and the location of eyes can be considered as the main fact to be focused when recognizing gaze direction [12]. Since facial expression recognition mainly uses FACS to locate different features of the face, eyes are also being covered under facial expression recognition [38]. It is reasonable to use facial expressions because it is the superset.
Gestures	When capturing gestures, head gestures (head position), hand gestures (shape and motion), body motion (spinal column, DOF body, body center mass, joints) should be considered to get better results [16]. It is hard to capture all of these through a mobile device.
Brain Signals	It is not practical to measure brain signals using a mobile device and it requires to take scalp EEG signals to be taken using special equipment [15].

The only remaining methods are, detecting emotions through facial expressions and text analytics. The next section discusses the limitations of the above two methods and the appropriateness of usage of those two techniques when using mobile devices.

2.2.5 Best Methods to Detect Emotions When Using Mobile Devices

According to the best of our knowledge, detecting emotions through facial expressions is more accurate [38] than other methods. Even though it is accurate there are limitations when practically applying this method, such as endures from face oval slanted cover, problems occurred due to the location of the camera, problems of lighting and uneven illumination [32].

The remaining method, text analytics plays a major role alongside facial expressions when discussing emotion detection [34], [42], [49]. Text analytics is practical when considering a mobile phone because a user usually interacts with a mobile device by typing, touching and swiping. As mentioned in Section 2.2.4, it is pointless to measure behavioural factors when typing, touching and swiping, rather it is meaningful to consider the content which the user is typing, which is text. Even though it is meaningful, text analytics suffers from limitations such as ambiguity of words, lack of ability to recognize sentences without keywords, the need of phonetic data and troubles in deciding emotion markers [33].

Although the above two methods have such limitations, both these methods have accurate APIs and SDKs which can be easily used when detecting emotions [25], [26]. The **first research question** suggests (Refer Section 1.2.2) to find a method to **combine these two techniques and to check the practical applicability and appropriateness of these methods when detecting emotions when using mobile devices.**

2.3 Context

The term “Context” has a broad meaning. Thus, can be sub-divided into the following sections as shown in Fig. 2.2. In the popular concepts like Spatiotemporal Structure Learning [18] and Situated Computing [24] context is defined using physical location, time, weather and user activities. Also, some recommender systems have kept track of time, to mark down the recently used and frequently used menu items, to recommend menus automatically when using mobile devices [27]. These applications will be further discussed in [Section 2.5](#).

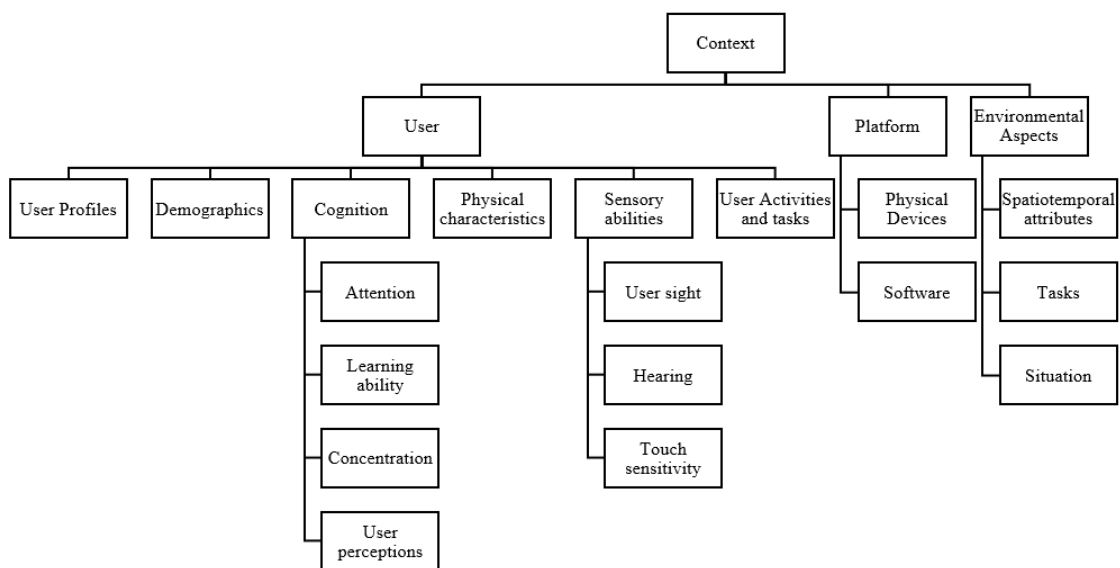


Figure 2.2: Dividing the context into subcategories [4]

2.4 Proposed Management Techniques When Designing an Adaptive System

A system which can adjust and blend its actions or configurations to assist different user requirements according to the status of the user, status of the system and the current context in real-time can be defined as an adaptive system [50], [51].

There are several management techniques [21] to the adaptation of user interfaces as follows.

- Adaptable or Manual System - A system where the user can change the settings manually.
- Adaptable with system support or User selection - A system which supports the user to do the adaptation process.
- Adaptive with user control or User approval – The system controls the adaptation process with the direction of the user and gives freedom to the user to choose the alternative actions if necessary.
- Fully adaptive - The entire management process is done by the system which does predictions without disturbing the user by considering the user profile and other variables.

This research will introduce a model which is “Adaptive with User Control”.

2.5 Applications

This section will discuss five applications such as EmotionPush, Situation Dependent Browser, AUBUE, Spatiotemporal Structure Learning with Samsung Galaxy S and Automatic Mobile Menu Customization System which utilize emotion detection, context detection and presenting adaptive user interfaces.

2.5.1 EmotionPush

EmotionPush is an Android app that can detect and display the emotion of text from Facebook Messenger as push notifications [22].

The researcher has used emotional categories instead of emotional dimensions and six colours were used such as red, yellow, navy blue, green, orange and purple to represent six discrete emotions such as Anger, Joy, Sadness, Fear, Anticipation and Tired accordingly. Messages with neutral tones were displayed without any colour seven binary classifiers were used to classify the text messages into one of the six emotions or Neutral, based on the highest probability. Machine learning techniques such as

LibSVM was used to enhance the performance of the application while incorporating multi-label classification in the future.

The app was deployed on the Google Play Store. Eight native English speakers were selected who use Facebook Messenger as the participants and the following facts were discovered from the studies.

- Wrong predictions did not meddle with the participant's chatting activities.
- After seeing the push notifications, the participant tends to change his/her behaviour.
- Messages with emotions such as Sadness or Anger were read immediately by participants.
- Messages with emotions such as Joy were not read immediately by participants but kept away to read later.

For emotions such as Joy, Sadness, and Anger, high performance could be gained while low performance for emotions such as Neutral and Fear which needs improvements in future research.

2.5.2 Situation Dependent Browser

This application can do searching depending on the active context which the authors have defined as the situation. The situation was represented by $S(L, t, a, /)$ where L denotes location, t denotes time, a denotes the set of user activities, and $/$ stands for a set of other critical circumstances.

The situation-dependent browser can do searching with simple discovery and efficient filtering of resources and data. Suppose that you have accessed a file last week at university during a meeting. This browser can find that file based on the situation (location, time, activity, other factors) and present it to the user. This was done by decomposing metadata into three levels as below.

1. **Physical level** - Retrieve primitive information such as location parameters from GPS, time from the system clock.

2. **Semantic level** – Identify the semantic meaning of the primitive information such as whether the user is at home or outside based on the GPS coordinates, whether today is a holiday or not.
3. **Pragmatic level** - Assess the assembled information with the current situation of the user.

They have mainly considered the desktop environment when implementing the situation-dependent browser and but they have stated this would be more useful if this can be implemented in a mobile computing environment too.

2.5.3 AUBUE (Adaptive User-Interface Based on User’s Emotion)

It has proven that different decisions can be made about a user’s emotion by analyzing the user’s behaviour when interacting with the system interface. Out of 22 emotions from OCC (Ortony, Clore, and Collins) [52], three were considered here such as Joy, Distress, and Anger. AUBUE (Adaptive User-Interface Based on User’s Emotion) can change the colour of the user interface based on the current emotion of the user [6]. This was implemented as a game that asks the participant to type the name of the shape displayed on the screen. By analyzing the typing pattern, the system will change the colour of the interface based on user emotion.

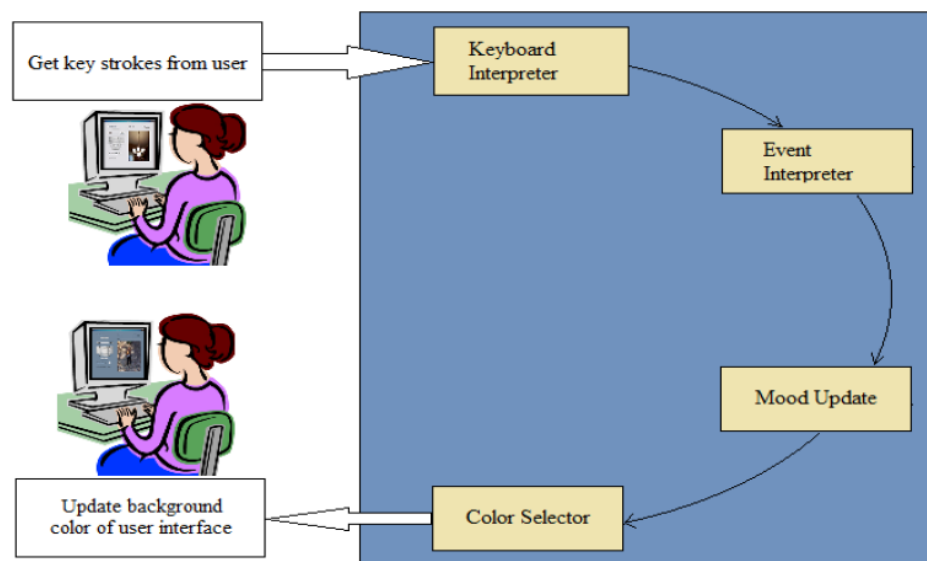


Figure 2.3: Architecture of AUBUE [6]

According to Fig 2.3, AUBUE consists of four components as follows.

1. Keyboard interpreter – Retrieves keyboard parameters such as typing speed, a number of times backspace was used and pressed irrelevant keys.
2. Event interpreter – Convert the keyboard parameters to events specified by Stathopoulou and et al. [53]
3. Mood Update – Update the current mood by mapping active emotions.
4. Color Selector – Selects the colour based on the identified mood.

2.5.4 Spatiotemporal Structure Learning with Samsung Galaxy S to Develop an Adaptive User Interface

This application can recommend apps to be used by considering the variables such as time, location, weather, emotion and the activities of the user using a supervised learning technique [18]. By the usage of Naïve Bayesian Classifier, they have created a machine learning model that can map relationships between context variables and predict the most suitable apps to be used based on the situation thus leading to an algorithm named “Spatiotemporal Structure Learning”.

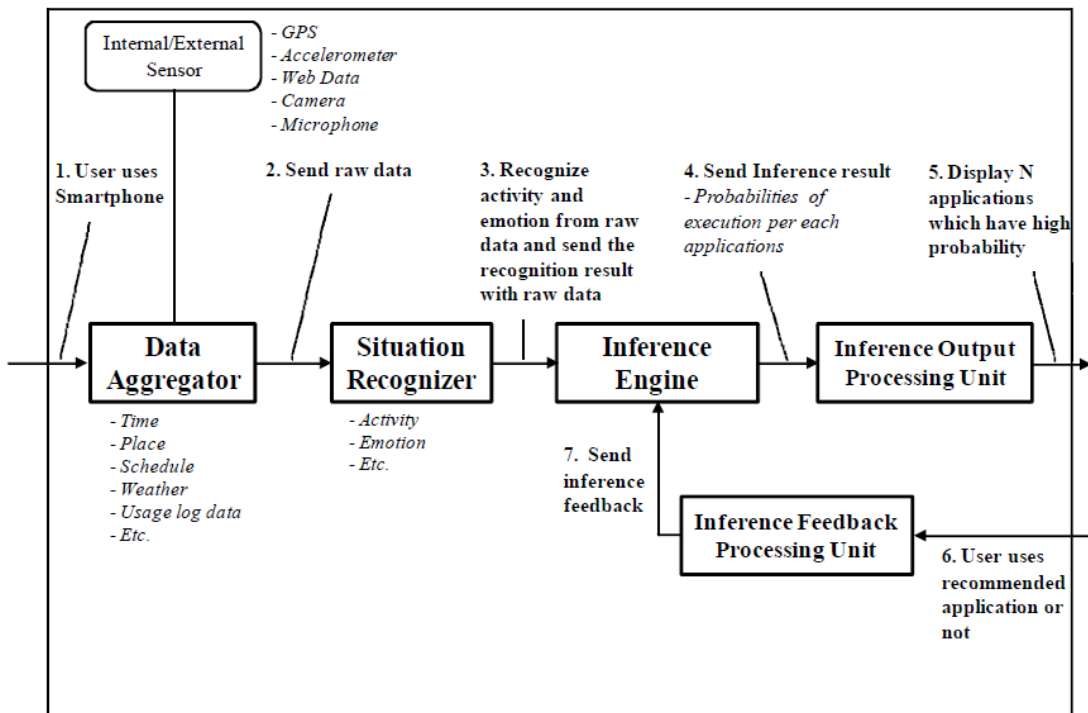


Figure 2.4: Architecture of spatiotemporal structure learning application [18]

Fig 2.4 describes the system architecture of this application. 3-axis accelerometer data was used to analyze the user activity as walk, stop, running or riding in a car. The module named “Situation Recognizer” has the responsibility to perceive user emotions, but in this system, they have entered the user emotion manually which they expect to improve in the future using facial expressions and voice analysis. “Inference Output Processing Unit” recommends the suitable applications to the user, and checks whether the user accepts it or not. If not, it resends the feedback to the “Inference Engine” through the “Inference Feedback Processing Unit” to re-learn the feedback.

The data collection for the machine learning model was done using a questionnaire which has distributed to several users and asking them to select the most preferred mobile apps to be used based on the following factors.

- Emotion - positive, negative or neutral
- Tasks or activities – walking, running or riding in a vehicle
- Time – morning, afternoon or evening
- Location – At home or workplace
- Weather condition – sunny, cloudy or rainy

2.5.5 Automatic Mobile Menu Customization System

This system can suggest and change the menus of a mobile device by keep track of the frequently and recently used activities when using mobile devices [27]. This system has eliminated famous collaborative filtering approach which has been used by most recommender systems [54], due to the privacy issue when uploading user histories to third party servers through mobile phone to get the recommendations. Instead, they have proposed a ranking mechanism by introducing a feature vector and with the help of a ranking SVM. The significance of this system in this literature is the following equation [27] as shown in eq. (1) which was used when calculating the rank of the menu items to be recommended.

$$H_i = \frac{1}{2} (RankByRecency_i + RankByFrequency_i) \quad (1)$$

The rank of the function/menu item by considering whether it has been used recently was denoted by *RankByRecency_i*, and the rank of the function/menu item by considering whether it has been frequently used was denoted by *RankByFrequency_i* and a total rank of H_i was achieved.

By considering all the above applications as explained in [Section 2.5](#) and the concepts of Emotions and Context, the conclusion paves the path to the **second research question** to find a way to **build a new model to do optimal decision making in order to perform affective interactions by considering user emotions based on the context.**

2.6 Summary

This chapter mainly focused on providing an extensive review of literature related to this research. The literature review began discussing the basic concepts of “Emotions”, followed by “Emotion Detection” techniques and then specially discussed the “Appropriateness and the best methods to detect emotions when using mobile devices”. The path to the need for the first research question was thoroughly elaborated by highlighting the advantages and disadvantages of existing techniques. Then the term “Context” was introduced by introducing the popular context parameters, followed by the discussion of “Different management techniques when designing an adaptive system”. The literature review concluded by discussing the “Applications” which are related to this research and the basic concepts behind them, associated with the second research question.

Chapter 3 - Design

This chapter consists of four sections that cover a detailed description of the research design. Section 3.1 will justify the usage of the methodology which was chosen in Section 1.5. Section 3.2 describes the detailed description of the research design step by step and the abstract view of the final system will be explained in Section 3.3. Finally, the ethical issues which may arise and the steps taken will be discussed in Section 3.4.

3.1 Justification for the Methodology

As explained in Section 1.5, Action Research was used as the methodology with User-Centered Design (UCD) [30] which was mainly done using Evolutionary Prototyping [31].

UCD was mainly used to identify the mental model of the user to facilitate the real expectations of the user. Otherwise, a system that was designed by only considering the designer's mental model may not be usable because the user's mental model and the designer's mental model may not always be matched [30].

To identify the user's mental model, a focus group of 40 participants was selected for the prototype evaluations of the first three prototypes (Prototypes 1, 2 and 3) which will be explained in Section 5.1.1. For the evaluations of the last two prototypes (Prototypes 4 and 5), a user group with nine males and nine females were selected from the previous 40 participants due to the time limitation as explained in Section 5.4.1. An iterative approach was carried out as mentioned in Section 1.5, where each phase of the development was evaluated with the user participation as soon as the prototyping was done. The feedback and observation from the prototype evaluations were the main incentives used to plan the next iteration.

3.2 Design Overview

Since this research uses an iterative approach as discussed in [Section 1.5](#), the design also consists of an iterative set of tasks as shown in Fig. 3.1.

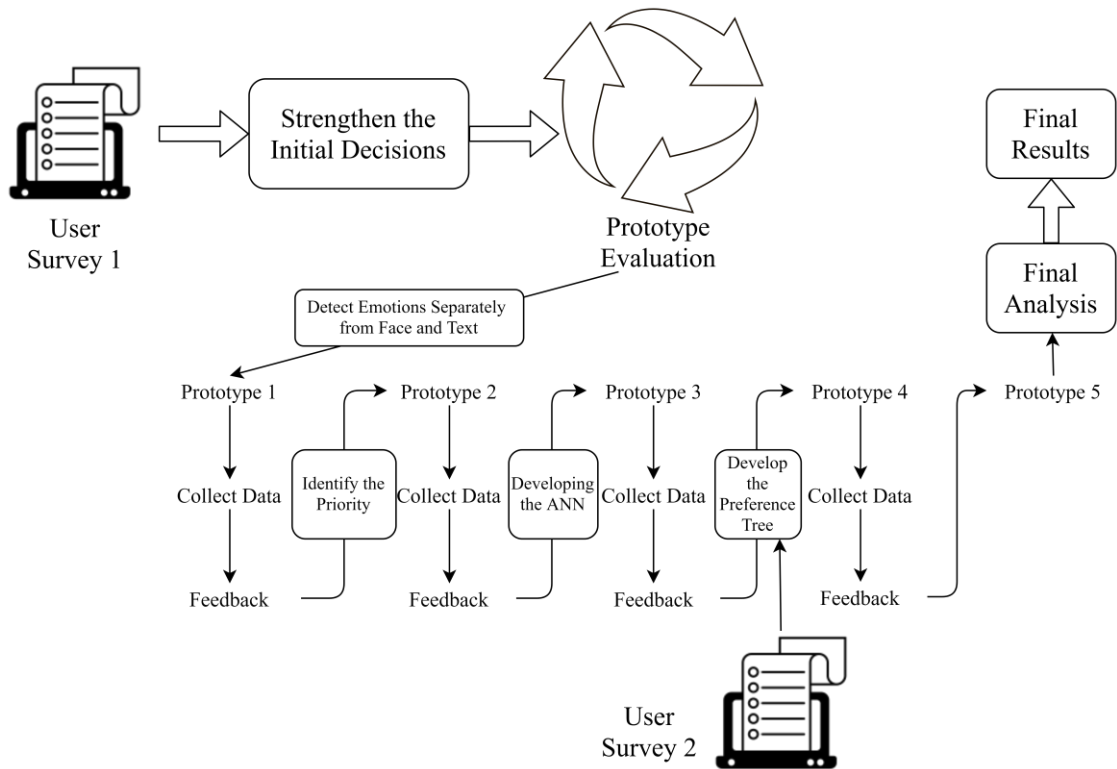


Figure 3.1: Design steps of the research

The research design consists of six main tasks as follows.

1. Conducting User Survey 1
2. Strengthen the Initial Decisions
3. Prototype Evaluation (an iterative process)
4. Conducting User Survey 2
5. Final Analysis
6. Final Results

3.2.1 Conducting User Survey 1

As shown in Fig. 3.1, the first step of the research design is to conduct User Survey 1. A questionnaire entitled “**Improving User Experience based on User Emotions when using Mobile Devices**” was created. The main aim of this study is to strengthen

the initial decisions to be taken before starting the research. This questionnaire was distributed as a Google Form, where the main expectation was to get more responses to find the facts which have to be improved through this research. There were seven questions, and the 4th question to the 7th question were the prominent ones which are listed below.

- **Question 4** - When you are using a mobile device, you can feel many emotions due to various reasons. Assume there is a system that can provide you with feedback and guidance based on your emotions. Choose the emotions which you would like the system to track and provide feedback and guidance?
- **Question 5** - In order to change your mood (emotion), what functionalities of a mobile device would you like to use?
- **Question 6** - Do those functionalities have an impact based on the time of the day? (Eg: In the morning you like to listen to music, but at night you like to chat with your friends)
- **Question 7** - Do those functionalities have an impact on the location where you are at that moment? (Eg: Work/University or Home)

A pilot study has been carried out with 10 participants, and then, the questionnaire was improved and distributed to the general public (Snowball sampling [29]). 310 responses were obtained. (Please refer [Appendix A.1](#) for full questionnaire/Google Form for this survey and [Appendix A.2](#) for results)

3.2.2 Strengthen the Initial Decisions

The initial decisions were taken as below to start the implementation of prototypes from the results obtained through User Survey 1.

- **The decision from Question 1** - Happy, Sad and Anger should be the important emotions to be tracked to provide adaptive functionalities. Since Happy is a positive emotion there is not a need to change that emotional state of the user, but Sad and Anger should be focused when suggesting adaptive functionalities.

- **The decision from Question 2** - Most people have chosen Listen to Music, Watch a Video, and Chat with Friends as the most affective functions to be considered to change their mood.
- **The decision from Question 3** - Respondents decided the time of the day should be considered when choosing the adaptive function.
- **The decision from Question 4** - Respondents decided the location of the particular person should be considered when choosing the adaptive function.

By considering the above decisions the development of prototypes was started and User Survey 2 was designed (Refer [Section 3.2.4](#) for User Survey 2).

3.2.3 Prototype Evaluation

Prototype evaluation is an iterative process where five prototypes have been developed and evaluated through user participation. Prototypes 1, 2 and 3 mainly covered the requirement which was emerged from the first research question, while the Prototypes 4 and 5 mainly served the purpose of the second research question (Refer [Section 1.2.2](#) for research questions). The implementation details of each prototype will be discussed in [Section 4](#) and the detailed evaluation procedure and results will be discussed in [Section 5](#).

By the end of each prototype evaluation the data was recorded and the decisions were taken to improve the prototype using the Evolutionary Prototyping approach [31].

3.2.4 Conducting User Survey 2

After analyzing the results of the previous User Survey 1, another questionnaire was designed, titled “**Collecting Data to Improve User Experience based on User Emotions when using Mobile Devices**”, to capture data in order to be used to develop Prototype 4 (Refer [Section 4.4](#) for implementation details). Similarly, as in User Survey 1, a pilot study has been carried out with 10 participants and after that, the questionnaire was improved and distributed to the general public (Snowball sampling

[29]) where the target audience was the students or employed people who are 18 years or above, and 209 responses were able to obtain. (Please refer [Appendix B.1](#) for full questionnaire/Google Form for this survey and [Appendix B.2](#) for results)

3.2.5 Final Analysis

This is the last evaluation step where after this step another prototype will not be developed. The final evaluation will be done for the final prototype (Prototype 5) with user participation and by giving the application to use while performing their day to day activities.

3.2.6 Final Results

Final results will be obtained using the inference from the final evaluation. Two steps will be considered as an **individual analysis** which depicts how the model performs for each individual user and as an **overall analysis** which discusses how the individual insights can be used to do an inference of the model generally, which provides the ability to map the performance to any user who may use this in future. Simply, it will show how the model will generally behave for any user.

3.3 High-Level Architecture (Abstract Design)

The high-level architecture of the proposed system is shown in Fig. 3.2. This diagram corresponds to the final prototype which will be used as the proof of concept which is a keyboard, where a user can install this and use it as an ordinary keyboard but can suggest activities (functions) based on the emotions (The users were asked to sign a consent form which can be found in [Appendix C](#)) captured by facial expressions and text while also considering the current context. More details about the prototypes can be found in [Section 4](#).

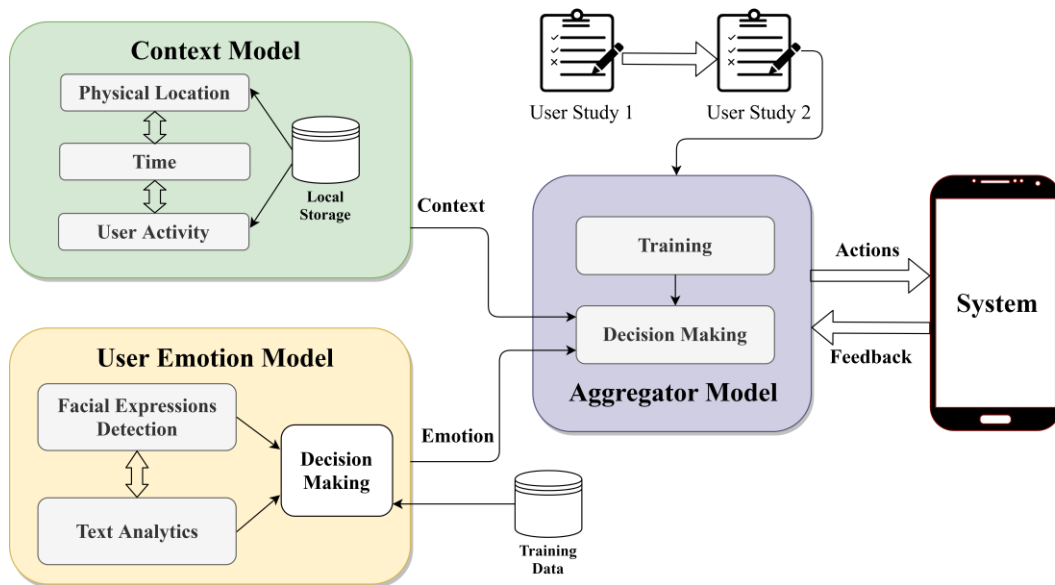


Figure 3.2: High-level architecture

3.3.1 Context Model

As shown in Fig. 3.3, the Context Model will be used to perceive the context parameters such as physical location, time and user activity. The descriptions of each method can be found in Table 3.1.

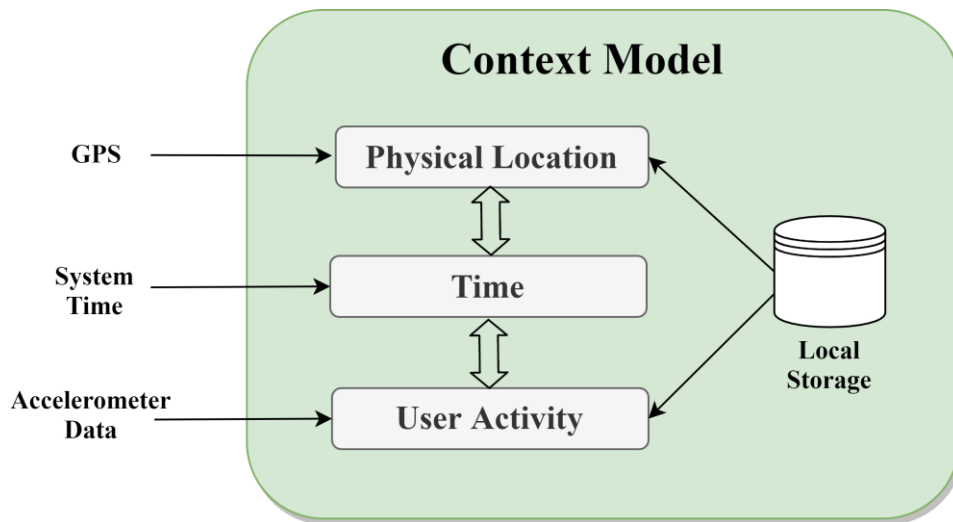


Figure 3.3: Context Model

Table 3.1: Methods to Perceive Context Parameters and Descriptions

Parameter	Method to perceive the real-time input	Description
Physical Location	GPS coordinates (Google Play Services [55])	In the beginning, the user will be asked to choose his/her location of home and work, and it will be stored on Shared Preferences (Local Storage) in the app. When the detection is taking place, latitude and longitude will be measured and must cross-reference it with stored data in Shared Preferences in Android.
Time	System Time or Internet Time	Local time should be taken.
User Activity	ActivityRecognitionClient [56] in Android based on Accelerometer Data	Several activities can be recognized from sensor data and can be divided into 3 groups as follows. <ol style="list-style-type: none"> 1. In a vehicle - IN_VEHICLE, ON_BICYCLE 2. Walking – ON_FOOT, WALKING 3. Still (Standing) – STILL, TILT <p>These groups will be stored on local storage and when the input comes, those will be cross-referenced.</p>

3.3.2 User Emotion Model

Emotion Model will be used to perceive the user emotions from facial expressions and text analytics. As shown in Fig. 3.4, **Affectiva SDK** [26] was used to detect emotions from facial expressions and **IBM Watson Tone Analyzer API** [25] was used to detect emotions from text analytics.

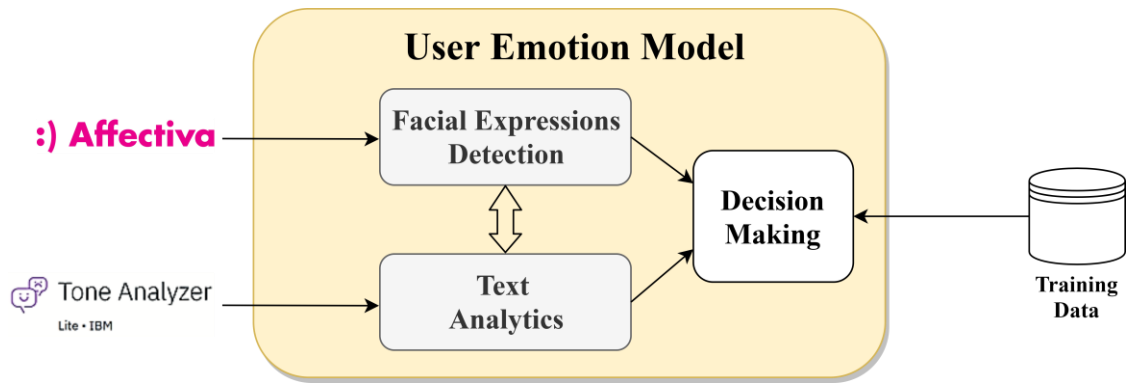


Figure 3.4: User Emotion Model

The Decision-Making component can provide optimal emotion from the perceived data. It is an Artificial Neural Network (ANN) [57] implemented using TensorFlow with Keras (Refer [Section 4.3](#) for implementation details).

3.3.3 Aggregator Model

As shown in Fig. 3.5, the Aggregator Model combines the emotion and the context and will decide the appropriate actions to be taken in the system. The basic data for this model was collected from User Survey 2 and a Preference Tree was built with probabilities of using functions based on a particular situation. The implementation details in [Section 4](#) will discuss this thoroughly.

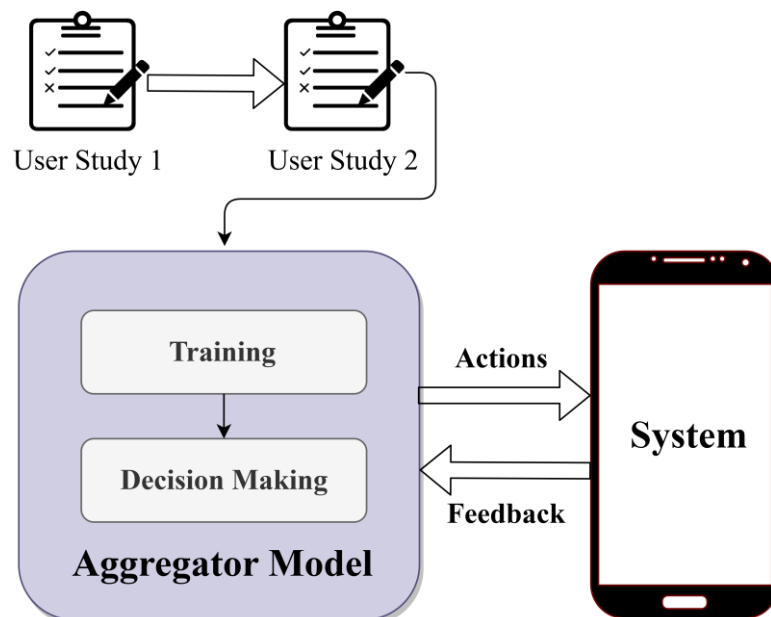


Figure 3.5: Aggregator Model

3.4 Ethical Issues and Steps Taken

3.4.1 Ethical Issues

- Users may not know that the face tracking is happening all the time when they have started using the prototype since a background service was used here (Refer [Section 4.1.3](#)).
- The keyboard can track what the user is typing which may arise a privacy issue, since the users may think that the text content will be saved in the system.
- The location of the user will be perceived using GPS (through Google Play Services [55]) by running a background service which may arise a privacy issue.
- The user activity will be captured using ActivityRecognitionClient [56] which may also a privacy issue.

3.4.2 Steps Taken

- To eliminate the above ethical issues and to inform the participants about the studies an Informed Consent Form (Refer [Appendix C](#)) was signed with the participants and explained that their confidentiality will be guaranteed throughout the studies.
- From Android 6.0 (and higher) the user who is using the app must explicitly give permission to the app to use the camera and to track the location and to track the user activity [58]. If the user does not want to give permission, he/she can revoke the permission by himself/herself.

3.5 Summary

This chapter mainly explained the design aspects of the research. First, the justification to use the research methodology and how it is suitable to an HCI research was explained, followed by the design overview which elaborated the steps of the research. Then the high-level architecture of the proof of concepts was explained and the chapter concluded by discussing the ethical concerns and the steps taken to eliminate those.

Chapter 4 - Implementation

This chapter mainly focuses on the aspects of the implementation of the proof of concept, which was done iteratively and will be discussed through five sub-sections. [Section 4.1](#) will discuss the implementation details of Prototype 1 which has the ability to detect emotions from facial expressions and text data separately, along with the basic details of the keyboard and the functionality of it. [Section 4.2](#) will discuss the development of Prototype 2, followed by [Section 4.3](#) which will discuss the implementation of Prototype 3, where an ANN was incorporated into the application which can decide the most optimal emotion based on facial expressions and text data. Those three implementations mainly cover the first research question which was mentioned in [Section 1.2.2](#). Next, the development of Prototype 4 will be discussed in [Section 4.4](#), which can suggest functions based on the current user emotion and the context parameters. An improvement to Prototype 4 was done by tracking frequently used functionalities and implemented Prototype 5 which will be discussed in [Section 4.5](#) which concludes the second research question which was mentioned in [Section 1.2.2](#). For each prototype, the evaluation details will be discussed in the next chapter.

4.1 Prototype 1

The main reason for developing Prototype 1 was to get an insight on how emotions can be detected **separately** from facial expressions and text data. By identifying the practical limitations through evaluations with user participation, a comparison will be done after combining the emotions from both these methods in future prototypes while observing whether an improvement has been made. (Refer [Section 5.2.2](#) for a comparison of Prototype 1 and Prototype 2)

A keyboard named “**Emotional Keyboard**” (compatible with Android devices) was developed that can track emotions from facial expressions and text **separately** and send notifications stating the emotion. This keyboard was designed with QWERTY layout, which has lowercase and uppercase letters (with the caps lock) with numbers

from 0 to 9 and basic symbols such as . , / * etc. Its full functionality is discussed in Section 4.1.1.

4.1.1 Functionalities of Prototype 1

When a user was chatting using a chat application, the keyboard can detect the emotion from the text and will send push notifications. Also, a background service will be running to capture facial expressions through the front camera of the device, and the notifications will be sent when an emotion is detected through that. (The notifications from both text and facial expressions will be stacked on the mobile device, allowing the user to easily distinguish the two types of notifications as shown in Fig. 4.1.)

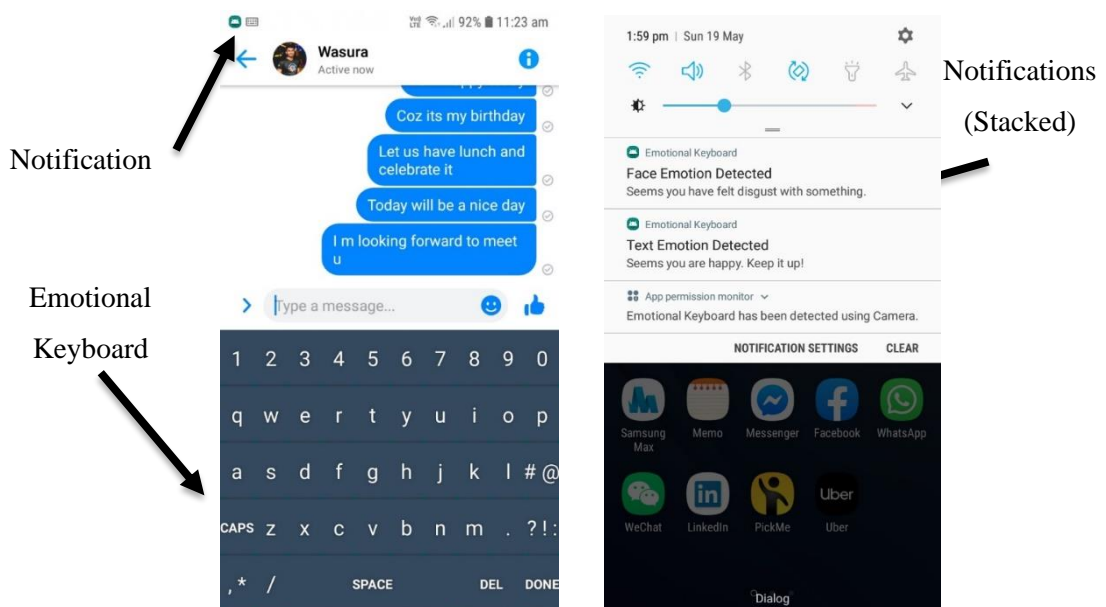


Figure 4.1: Prototype 1- Demonstration and stacked notifications

When the user taps on a notification, an Alert Dialog will appear as shown in Fig. 4.2 and will ask the user to enter his/her feedback/satisfaction with the current detected emotion. It was recorded in a database which was used later to analyze the user satisfaction concerning this prototype. The satisfaction for emotions from text and facial expressions will be recorded separately. If the user press YES, the data will be recorded instantly. If he/she press NO, another screen will appear asking the user to select the actual emotion that he/she is feeling right now as shown in Fig. 4.3.

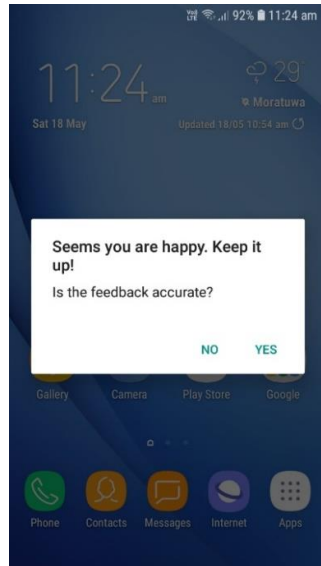


Figure 4.2: Prototype 1 - Asking feedback

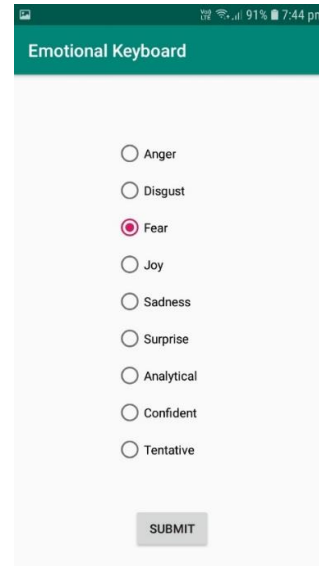


Figure 4.3: Prototype 1 - If user pressed NO

4.1.2 Emotion Detection Procedure Using Text

This keyboard tracked each keystroke even it is a backspace. **IBM Watson Tone Analyzer API** [25] was used to detect emotions from the text. The emotion will be detected in four situations as stated below.

- Once the user has finished typing one line and enter the DONE button. (DONE will be moving the cursor to a newline or sending the message)
- Once the user has finished typing one message and send that message.
- Once the user has started to type again after sending the previous message.
- Once the user closes the current chat application.

In these situations, the keyboard will send requests to the API with the captured text and will retrieve the response. This response will consist of emotions such as “**joy**”, “**fear**”, “**sadness**”, “**anger**”, “**analytical**”, “**confident**” and “**tentative**”.

Eg: -

```
{
  "document_tone": {
    "tones": [
      {
        "score": 0.931034,
        "tone_id": "fear",
        "tone_name": "Fear"
      },
      {
        "score": 1,
        "tone_id": "anger",
        "tone_name": "Anger"
      },
      {
        "score": 0.916667,
        "tone_id": "sadness",
        "tone_name": "Sadness"
      }
    ]
  }
}
```

Based on the response, an array will be updated as percentages as shown in Table 4.1. (After receiving the previous response, the array will look like below)

Table 4.1: Prototype 1 - Text Emotions Array

joy	fear	sadness	anger	analytical	confident	tentative
<i>0.0</i>	<i>93.1034</i>	<i>91.6667</i>	<i>100.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>

When more responses are retrieved, this array will keep updating by adding that particular values to that particular emotion and by calculating the average. This will happen for 60 seconds (one minute). After that, the emotion which has the maximum

percentage will be taken from the array and a notification will be sent to the user stating that emotion. (When the notification is sent, the values of the array will be reinitialized with zeros)

4.1.3 Emotion Detection Procedure Using Facial Expressions

This application captures the image frames from the camera. Those captured frames were processed using **Affectiva SDK** [26] which outputs emotion percentages for emotions such as “**joy**”, “**fear**”, “**sadness**”, “**anger**”, “**surprise**” and “**disgust**” according to Ekman’s Model [8]. For every 60 seconds (one minute), a captured frame will be processed and will be analyzed by the SDK and an output will be perceived as shown below with six emotions.

Eg: -

joy: 98.1125%
fear: 0.0042%
sadness: 1.0232%
anger: 10.2314%
surprise: 64.0142%
disgust: 10.1015%

Based on the output, another array will be updated as percentages as shown in Table 4.2 (similarly as mentioned in [Section 4.1.2](#)). The emotion which has the maximum percentage will be taken from the array and a notification will be sent to the user based on that emotion. (When the notification is sent, the values of the array will be again reinitialized to zeros)

Table 4.2: Prototype 1 - Face Emotions Array

joy	fear	sadness	anger	surprise	disgust
<i>98.1125</i>	<i>0.0042</i>	<i>1.0232</i>	<i>10.2314</i>	<i>64.0142</i>	<i>10.1015</i>

4.1.4 Feedback Process

When the user taps on the notification, he/she will be asked to enter the satisfaction by pressing either “YES”, or “NO”. Then, as previously shown in Fig. 4.3, another screen will appear to capture the actual emotion that he/she was feeling at that moment.

The feedback was recorded in a **Firestore Real-time Database** for emotions from text and facial expressions separately which will look as shown in Fig. 4.4. The YES/NO count of the particular emotion was increased according to what the user has chosen. Also, the respective emotion object with percentages was recorded with respect to each notification as shown in Fig. 4.5, and that data was used to train the ANN when developing Prototype 3 (Refer [Section 4.3](#)).

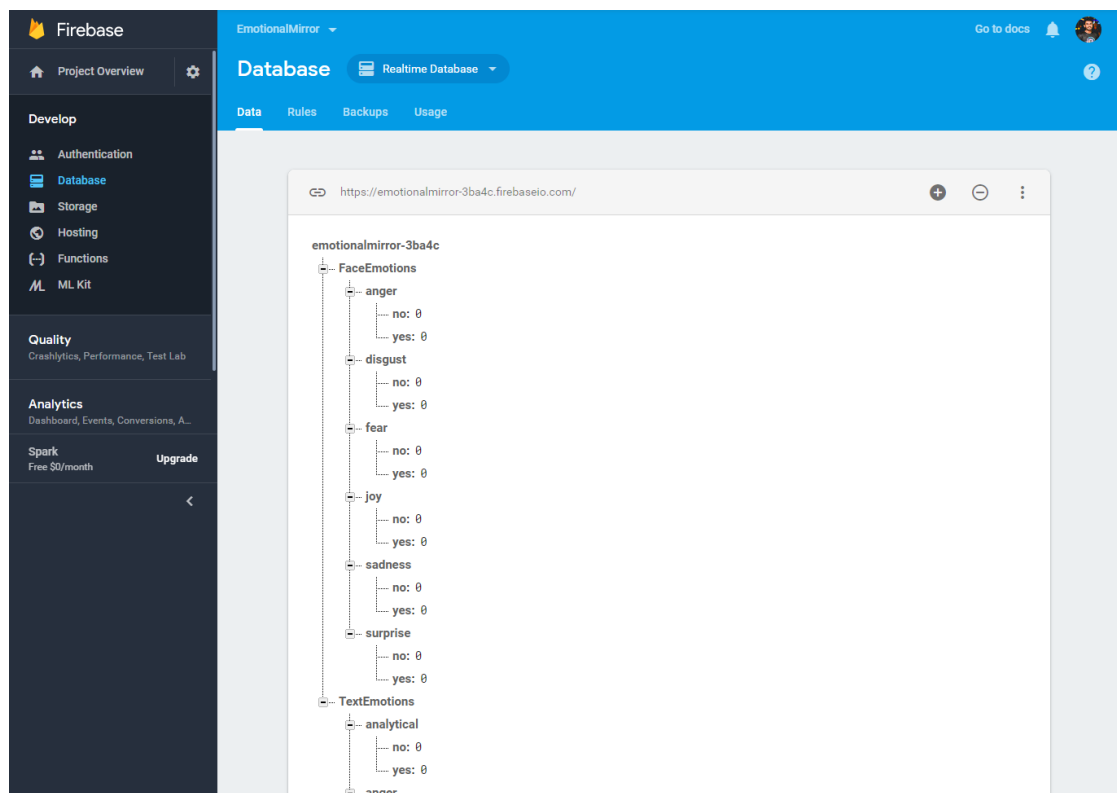


Figure 4.4: Prototype 1 - YES/NO count in Firebase database

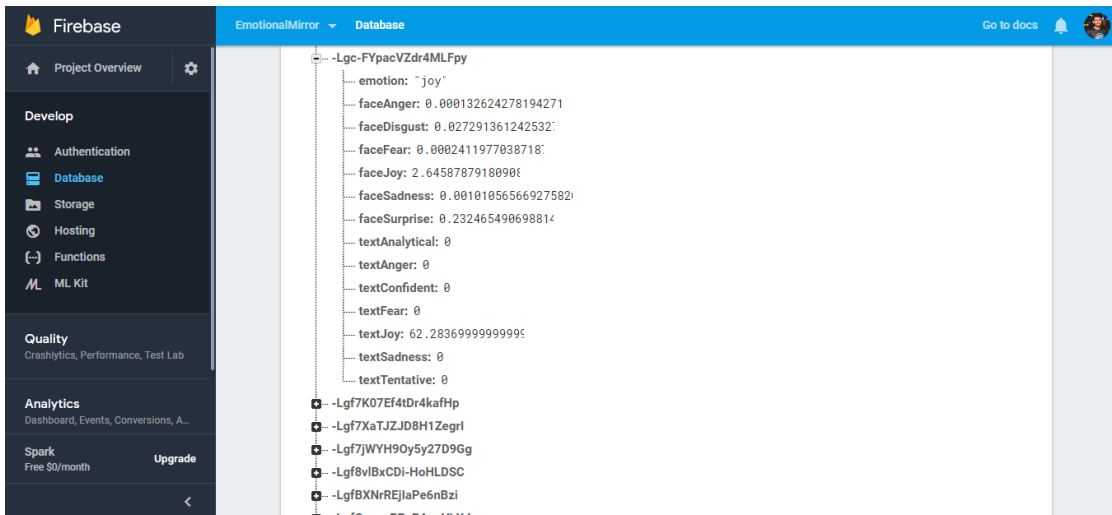


Figure 4.5: Prototype 1 - A record with respect to a particular notification

The implementation details are attached in [Appendix D.1](#). The evaluation and results of Prototype 1 will be discussed in [Section 5.1](#). From the observations carried out during the prototype evaluation, a need was arisen (Refer [Section 5.1.5](#)) to build another prototype. The next section discusses the implementation details of that.

4.2 Prototype 2

After the evaluation of Prototype 1, the same prototype was changed to improve the identified limitations which will be discussed in [Section 5.1.5](#). Slight advancements have been performed without doing major changes.

Since the results from facial expressions were not accurate, it has proven that the practical usage of detecting emotions through face has limitations as discussed in [Section 2.2.5](#). If **emotions from both face and text** were available, a decision was taken to consider the emotion only from text, since it has been proven the emotion from facial expressions might be deceiving (which will be discussed in [Section 5.1.4](#)). If the **emotions were available only from the face**, not text, then consider a confidence threshold which is 50% which was decided through trial and error during the evaluations. If the confidence of the emotion from the face was greater than the threshold, the notification was sent stating that emotion, otherwise discarded it. To reduce the number of mistakes that have been done when detecting emotions from the text, another threshold which is also 50% was considered when deciding the emotion.

(Refer [Appendix E](#) for deciding the threshold values). The complete changes suggested here are summarized in the algorithm given below. No other changes have been done to the prototype.

Algorithm: Prototype 2 Emotion Detection

Input: EmotionFromFace, EmotionFromText

```
If ((EmotionFromFace != null AND EmotionFromText != null) OR
      (EmotionFromFace == null AND EmotionFromText !=null) ){
    If (Confidence(EmotionFromText) >= 50){
        Output Emotion;
    }
    Else {
        Discard Emotion;
    }
}
Else If (EmotionFromFace != null AND EmotionFromText == null ){
    If (Confidence(EmotionFromFace) >= 50){
        Output Emotion;
    }
    Else {
        Discard Emotion;
    }
}
```

The implementation details are attached in [Appendix D.2](#). The evaluation and the results of the Prototype 2 will be discussed in [Section 5.2](#). It was proven that the need for Prototype 2 was satisfied (which was to give a high priority to emotions from the text and to detect emotions from the face only if text emotions cannot be detected) but a more practical system can be developed if a model can be created using the collected data during the previous prototype evaluations because that data has automatically incorporated the behavioural factors and the compatibility of the context. As a result, a new prototype was implemented as Prototype 3, which incorporated an ANN, which

can find the most meaningful emotion by considering data from both facial expressions and text.

4.3 Prototype 3

To develop a more practical system using the collected data during the previous prototype evaluations another prototype was implemented. The keyboard User Interface (UI) of this prototype was similar to the previous prototypes but a pre-trained TensorFlow Lite model (an ANN) was used to decide on the most optimal combined emotion to produce notifications to the user. The feedback was recorded to develop the confusion matrix to do the final analysis of this prototype.

4.3.1 Developing the Artificial Neural Network (ANN)

A simple neural network has been developed using Keras (with TensorFlow).

Preprocessing

After loading the dataset, all the values were normalized by dividing them by 100, therefore the values will only be in the range of 0 to 100. Converted the categorical string labels such as “analytical”, “anger”, “confident”, “disgust”, “fear”, “joy”, “sadness”, “surprise” and “tentative” to categorical integer labels and encoded them again as using one-hot encoding [59] because many machine learning algorithms cannot operate on label data directly and require all input variables and output variables to be numeric.

Instantiation and Configuration of the model

The ANN consists of three layers as shown in Fig. 4.6. The first layer is a Dense layer which consists of 13 input nodes to perceive 13 values corresponding to emotions from face and text. 30 nodes were used for output to the next layer. Rectified Linear Unit (relu) [60] was used as the activation function here. The second layer is a Dropout layer, which consists of randomly setting a fraction rate of input units to zero at each update during training time, which helped to prevent overfitting. 0.5 was selected as

the fraction rate here. The third layer is the output layer which is also a Dense layer with nine output nodes corresponding to nine emotion classes (combined optimal emotions) to be produced as output. Softmax function [60] was used as the activation function which will produce nine probabilities corresponding to nine classes of emotions which are “analytical”, “anger”, “confident”, “disgust”, “fear”, “joy”, “sadness”, “surprise” and “tentative”. (The combined optimal emotion would be the emotion class which has the maximum value)

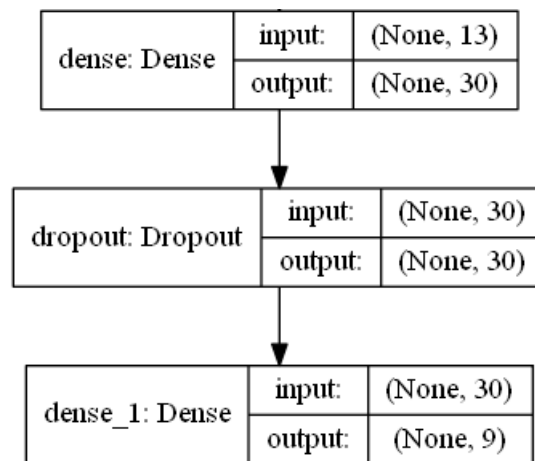


Figure 4.6: Prototype 3 - ANN architecture

The model was trained with a batch size of 10 and 200 epochs. Adam optimizer [61] with a learning rate of 0.001 was used. Categorical cross-entropy was used as the loss function since it represents the true class as a one-hot encoded vector, and closer the outputs of the model are to that vector, lower the loss. 70% data was used to train the model while 30% was used to do the validation (Refer [Section 5.3.1](#) for results of training and validation and refer [Appendix D.3](#) for the Python implementation of the model).

4.3.2 Improving Android App

The ANN which was created in [Section 4.3.1](#) was saved as a .h5 model in Keras and converted it to a **TensorFlow Lite model** to be incorporated in the Android app. As mentioned in Prototype 1 ([Section 4.1](#)), the emotions from the text was tracked for 60 seconds and updated the array by averaging each value of it. The emotions from facial expressions was detected by capturing and processing an image frame every 60

seconds. Both these techniques will update their corresponding arrays and those were merged as shown below, to be fed as inputs to the ANN (the TensorFlow Lite model).

```
Input [] = {  
    {  
        face("anger"),  
        face("disgust"),  
        face("fear"),  
        face("joy"),  
        face("sadness"),  
        face("surprise"),  
        face("analytical"),  
        text("anger"),  
        text("confident"),  
        text("fear"),  
        text("joy"),  
        text("sadness"),  
        text("tentative")  
    }  
};
```

Above Input array can be fed to the ANN and output with nine values corresponding to nine combined emotions such as “analytical”, “anger”, “confident”, “disgust”, “fear”, “joy”, “sadness”, “surprise” and “tentative” can be retrieved. One value from the Output array corresponds to the composite emotion probability which was decided by the Input emotion array that has been fed to the model. The maximum of the nine values was selected and the corresponding composite emotion was chosen as the optimally combined emotion and was sent to the user via a notification and the user feedback was recorded. Refer [Appendix D.3](#) for the Java code which was used to incorporate the trained ANN into the Android app and to pass inputs to get the most meaningful emotion as the output.

4.3.3 Feedback Process

As mentioned in previous prototypes (Refer [Section 4.1.4](#)), when the user taps on the notification, he/she was asked to enter the satisfaction which can be either “YES” or “NO”. The feedback was recorded in the **Firestore Real-time Database** in a different format than earlier as shown in Fig. 4.7.

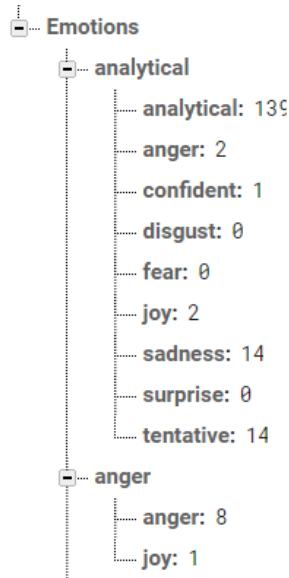


Figure 4.7: Prototype 3 - Feedback recorded format

Consider the emotion “analytical” above. It has sub-branches with nine emotions (similar to the Output array) including itself. Sub-branch “analytical”, keeps the number of records (corresponds to notifications) where the notification has stated (predicted) that it is “analytical”, and the user has also agreed that the actual emotion is “analytical”. Sub-branch “anger” records the number of records (corresponds to notifications) where the notification has stated (predicted) that it is “analytical”, but the user has disagreed and has chosen “anger” as the actual emotion. This is similar to all the other sub-branches “**confident**”, “**disgust**”, “**fear**”, “**joy**”, “**sadness**”, “**surprise**” and “**tentative**”.

Similarly, for all the nine emotions there were records as shown in Fig. 4.7, which were used to develop the **confusion matrix** after the prototype evaluations have been done (Refer [Section 5.3.2](#)). Also, as mentioned in previous prototypes (Refer Fig. 4.5) the respective emotion object with percentages was recorded with respect to each

notification, therefore that data can be used to train the machine learning model further if needed. According to [Section 5.3.3](#), this prototype is adequate to satisfy the first research question and the next prototypes will be mainly focusing on the second research question.

4.4 Prototype 4

Since the previous three prototypes mainly covered the first research question which was mentioned in [Section 1.2.2](#) the development of Prototype 4 was carried out to find answers for the second research question.

The UI of the keyboard and feedback procedure is similar to the previous prototypes but this can suggest adaptive functions by considering the user emotion based on the context such as time, location, and user activity based on a **Preference Tree** which will be discussed in [Section 4.4.1](#). The feedback was recorded to evaluate this prototype.

4.4.1 Preference Tree and the Registration Process

The Preference Tree consists of the probabilities that a particular user performs a function based on his/her current emotion and the context. The development of the Preference Tree was done using the data from User Survey 2 which was discussed in [Section 3.2.4](#) (Refer [Appendix D.4](#) for more details).

When a particular user installs the app, the user must register by creating an account as shown in Fig. 4.8. The main reason for this is to inform the users what is happening and to take the consent by registering through the app. The purpose of registration is to maintain a profile for that particular user by recording his/her Preference Tree in the Firebase database as shown in Fig. 4.9. The probabilities will be updated in the user profile based on his/her personal choices using the feedback which will be discussed in [Section 4.4.7](#).



Figure 4.8: Prototype 4 - Register screen

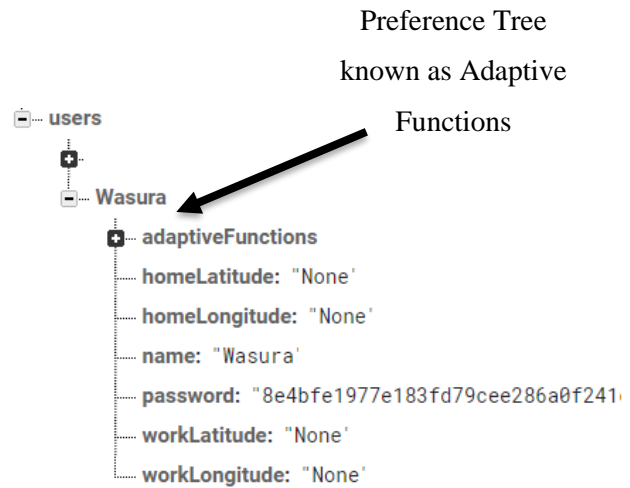


Figure 4.9: Prototype 4 - User database record

4.4.2 Retrieving Location of the User

After the registration, the user can log in to his/her account and the following Home Screen will be displayed as shown in Fig. 4.10. The coordinates of the current location were retrieved using Google Play Services [55], but the locations of Home and Work were initially “None”. User needs to set them manually (the longitude and latitude of the locations) by clicking “EDIT LOCATION DETAILS”.

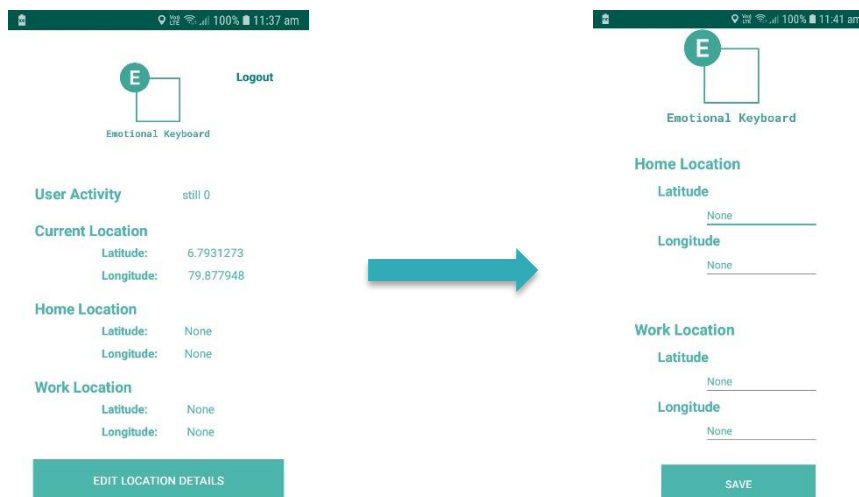


Figure 4.10: Prototype 4 - Edit location details

The appearance of the Home Screen after a user SAVES his/her Home and Work locations is shown in Fig. 4.11 and the data will be stored in Firebase Realtime Database as shown in Fig. 4.12.

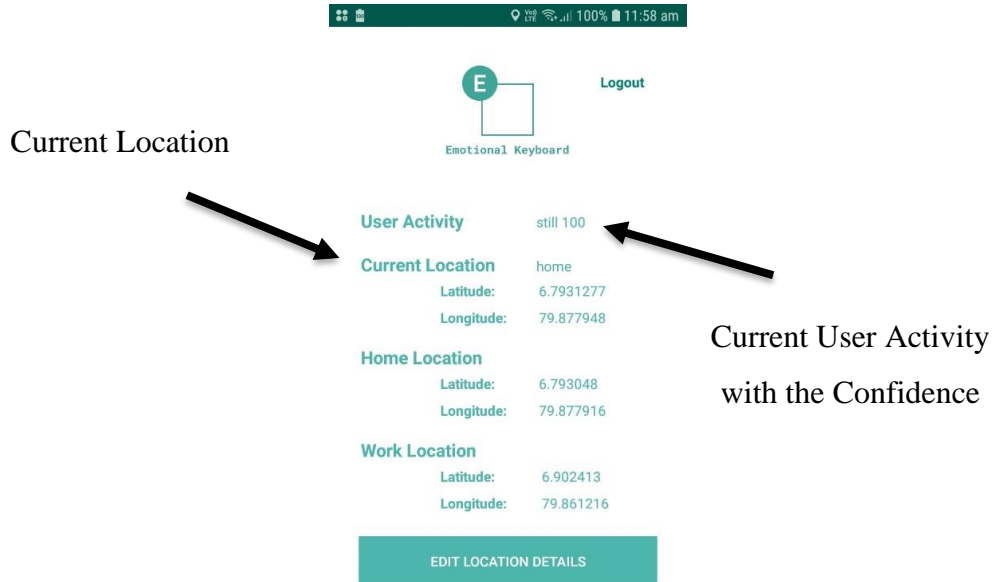


Figure 4.11: Prototype 4 - After setting the location details



Figure 4.12: Prototype 4 - Updated database after the location setting

When deciding the current location of the user, first the latitude and longitude were perceived through GPS in real-time and the perceived coordinates were compared with stored Home and Work location coordinates as shown below.

Algorithm: Prototype 4 Location Detection

```
If (Distance between home and current location is less than the distance between  
the workplace and the current  
location) {  
    If (Distance between home and current location is less than or equal to 20  
                                                meters)  
    {  
        location = "home";  
    }  
    Else {  
        location = "outside";  
    }  
}  
Else {  
    If (Distance between workplace and current location is less than or equal to  
                                                20 meters)  
    {  
        location = "work";  
    }  
    Else {  
        location = "outside";  
    }  
}
```

The code for the above implementation is presented in [Appendix D.4](#).

4.4.3 Detecting the Current User Activity

Current user activity was detected using the Activity Recognition Client in Android based on Accelerometer data [56]. Activities such as IN_VEHICLE, ON_FOOT, RUNNING, WALKING, ON_BICYCLE, STILL, TILT, UNKNOWN can be recognized from sensor data. They were grouped as follows.

1. **In a vehicle** - IN_VEHICLE, ON_BICYCLE
2. **Walking** – ON_FOOT, WALKING
3. **Still (Standing)** – STILL, TILT

The activities (along with the confidence value) were stored in an array and the activity with the maximum confidence was taken as the current user activity. (Confidence means the ability to predict what the user is currently doing. For example, the sensors of the android device might detect that the user is “Walking” with a 70% confidence and the user is “Still” with a 10% confidence which says that there is a high probability that the user was “Walking” rather than standing “Still” because $70\% > 10\%$) The code for this implementation can be found in [Appendix D.4](#).

4.4.4 Retrieving the Time

Time will be taken from the system. If the current time is between 5.00 am and 1.00 pm, then decided that it was morning, or else decided as evening. The time range was decided by considering the practical usage of this app during the prototype evaluations.

4.4.5 Emotion Detection

Emotion detection was done in the same manner as explained in [Section 4.3](#) using the ANN which was developed as Prototype 3. Nine emotions can be identified such as “analytical”, “anger”, “confident”, “disgust”, “fear”, “joy”, “sadness”, “surprise” and “tentative”. From User Survey 1 (Refer [Section 3.2.2](#)), most of the people have stated that “joy”, “sadness”, and “anger” should be tracked by this type of a system and produce feedback and guidance. In here “joy” is a positive emotion, but “sadness” and “anger” are negative emotions. Since our ultimate goal is to provide the user with a positive experience, “sadness” and “anger” should be tracked and must provide guidance to eliminate negative feelings. Therefore, in this prototype, the adaptive response will be issued only for emotions “**sadness**” and “**anger**”.

4.4.6 Demonstration of Suggesting the Adaptive Functions

The system decides the suggestion by first detecting the emotion. Then the time, then location, then the user activity was considered accordingly. This means a branch from the Preference Tree will be traversed from top to bottom. Consider Fig. 4.13 below.

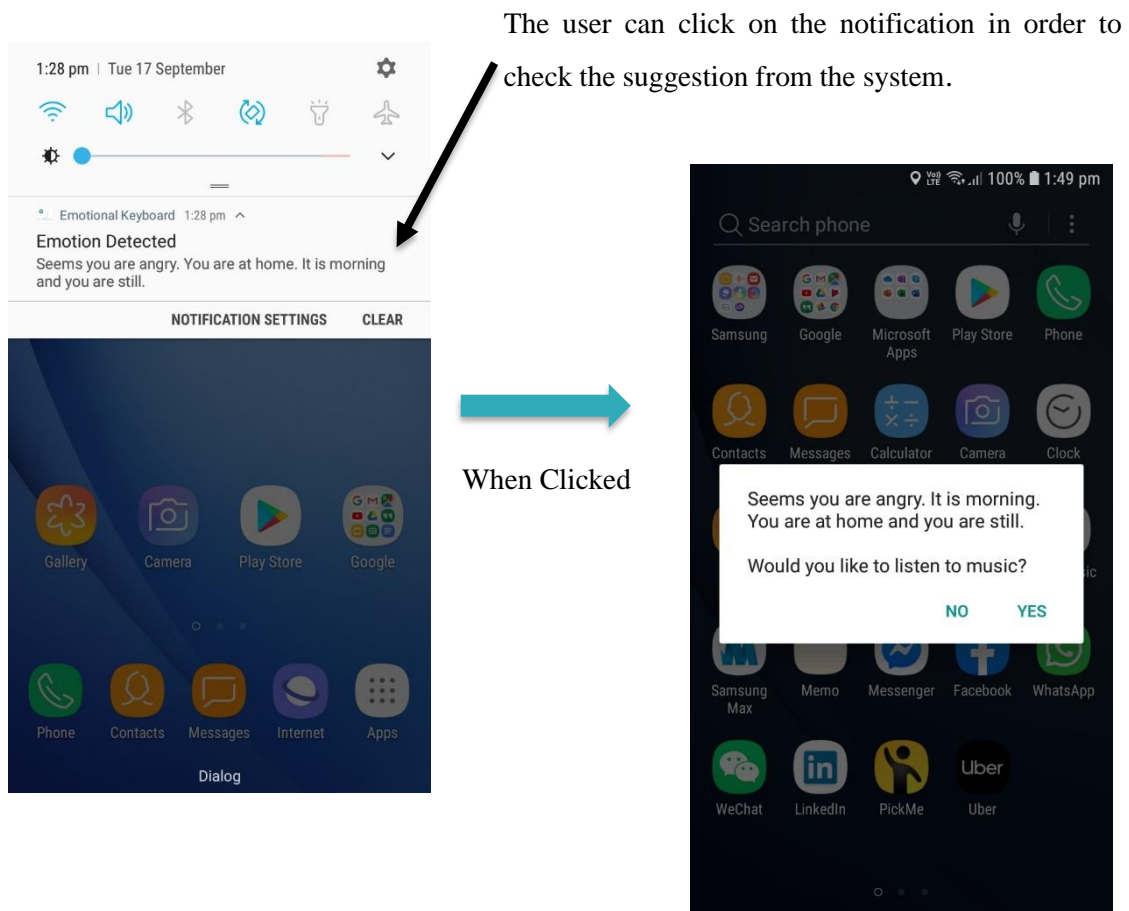


Figure 4.13: Prototype 4 - Suggesting functions

In the above example, the branch which should be traversed in the Preference Tree is “**anger -> morning -> home -> still**” as shown in Fig. 4.14 below. The leaf node of this particular branch consists of the functionalities to be suggested with the probabilities based on User Survey 2. The system will suggest the function which has the highest probability, which is in here “Listen to Music” as shown in Fig. 4.13.



Figure 4.14: Prototype 4 - Traversed branch

If the user pressed “YES”, the probability of the **suggested function** was **increased** in that particular branch using eq. (2) since the user has **chosen** the suggested function to use.

Increased probability of the chosen function

$$= \left(\frac{\text{Current probability of the chosen function} + 1}{101} \right) \times 100\% \quad (2)$$

According to Fig. 4.14, the current probability of “Listen to Music” is 32.25%. Therefore, the increased probability of “Listen to Music” should be $\left(\frac{32.25 + 1}{101} \right) \times 100 = 32.92\%$. Also, the probabilities of **other functions** will be **decreased** in that particular branch using eq. (3).

Decreased probability of other function

$$= \left(\frac{\text{Current probability of other function}}{101} \right) \times 100\% \quad (3)$$

For example, the current probability of “Browse Internet” is 7.95%. Therefore, the decreased probability of “Browse Internet” should be $\left(\frac{7.95}{101} \right) \times 100 = 7.87\%$. The new probabilities of the functions of the traversed branch shown in Fig. 4.13 should be as below.

- Browse internet: 7.87%
- Browse the gallery: 3.07%
- Chat with friends: 21.23%
- Listen to music: 32.92% (Increased)**
- Play a game: 7.67%
- Use social media: 27.22%

If the user pressed “NO”, then he/she will be asked for the function which would like to be performed at the current situation as shown in Fig 4.15. The probability of the **chosen function** will be **increased** and the probabilities of **other functions** will be **decreased** using the previous equations.

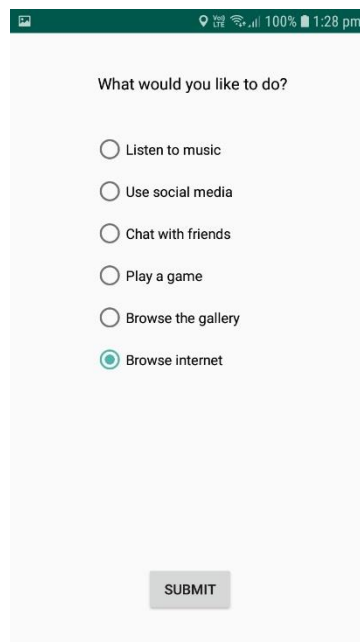


Figure 4.15: Prototype 4 - When user pressed NO

Finally, the branch in the database will be updated for the particular user.

4.4.7 Data for Evaluation

For evaluation purposes, another set of data for each user will be recorded in the user profile on a branch named “adaptiveChoices”. Each record of that branch will consist of the below fields.

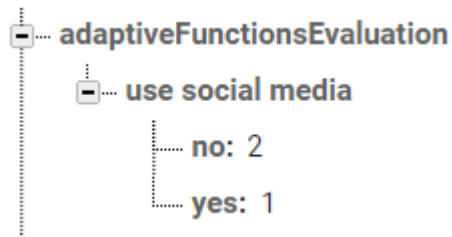


Figure 4.18: Prototype 4 - “adaptiveFunctionsEvaluation” to track predictions

Based on the results and evaluation of this prototype (Refer [Section 5.4](#)) there is a need for a another prototype that can suggest adaptive functions correctly and quickly without taking much time to adapt to the particular user. Prototype 5 was implemented as discussed in the next section to achieve this expectation.

4.5 Prototype 5

Since the analysis of Prototype 4 has shown that it takes more time to produce the correct suggestions which are not practical to evaluate, a new prototype was developed as Prototype 5. The functionality of this prototype is similar to Prototype 4 but the implementation differs a bit by keep tracking frequently used functions as explained in [Section 2.5.5](#) in the Literature Review.

4.5.1 “frequentFunctions” Field

When the user is registering for the first time as discussed in [Section 4.4.1](#), a new field named “frequentFunctions” will be added to the database as shown in Fig. 4.19.

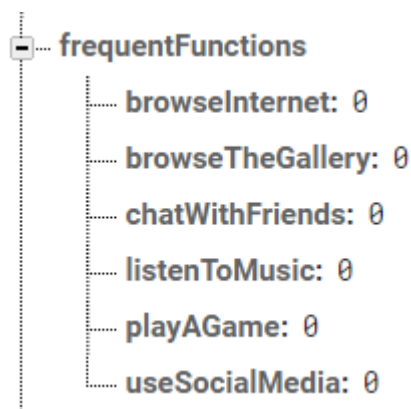


Figure 4.19: Prototype 5 - “frequentFunctions” field

When the user is providing feedback as mentioned in [Section 4.4.6](#), each time, the count of the particular function in the field “frequentFunctions” was incremented. For example, if the system suggested “Listen to Music” and user hit YES, then the count of “listenToMusic” will be incremented in the “frequentFunctions” field. If the user pressed NO, then again, the system will ask the user to choose the function which the user would like to do, and when that function was chosen the count of that function will be incremented in the “frequentFunctions” field.

4.5.2 Improved Method to Suggest the Adaptive Functions

The main difference of this prototype and the previous prototype occurs here, which is the usage of “frequentFunctions” field when suggesting adaptive functions to the user. As in the previous prototype, the probabilities will be taken from the Preference Tree in the database as discussed in [Section 4.4.1](#) by traversing through the branch, based on the current emotion and the context parameters which will output an array of probabilities for each adaptive function as discussed previously. Previously the function which has the highest probability from the Preference Tree was suggested to the user. In here, a new probability value named “Frequent Probability” will be calculated for each adaptive function using eq. (4).

Frequent Probability of function A

$$= \left(\frac{\text{Count of "function A" has in frequentFunctions}}{\text{Total count in frequentFunctions}} \right) \times 100\% \quad (4)$$

Assume that the “frequentFunctions” field consists of below values in the database for a particular user as shown in Fig. 4.20.

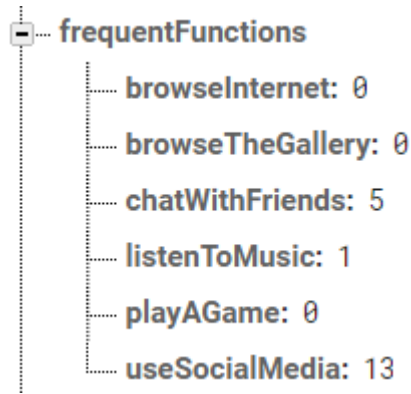


Figure 4.20: Prototype 5 - Example values for “frequentFunctions”

Assume that function A = “useSocialMedia”, then

$$\begin{aligned} \text{Frequent Probability of useSocialMedia} &= \left(\frac{13}{0 + 0 + 5 + 1 + 0 + 13} \right) \times 100\% \\ &= \frac{13}{19} \times 100\% = 68.4\% \end{aligned}$$

Likewise, after calculating the Frequent Probabilities for each function, those will be added to the corresponding Preference Tree probabilities and will be averaged using the eq. (5).

New Probabilty of function A

$$= \left(\frac{\text{Probabilty of "function A" from Preference Tree} + \text{Frequent Probability of "function A"}}{2} \right) \times 100\% \quad (5)$$

These new probabilities will be compared with each other and the adaptive function which has the maximum new probability will be suggested to the user. When the user has given feedback, the same procedure will be used to update the Preference Tree as discussed in [Section 4.4.6](#) and the count of the chosen function will be increased in the “frequentFunctions” field in the database as discussed in [Section 4.5.1](#). This Frequent Probability has the characteristic of “Which function the user has frequently used?” and this helps to keep track of the functions that the user has used with time, which eliminates the time constraint which will be discussed in [Section 5.4.3](#).

Code segments for the above implementations are attached in [Appendix D.5](#) and the evaluation and results of this prototype will be discussed in [Section 5.5](#).

4.6 Summary

This chapter mainly explained the implementation details of the proof of concept through five prototypes. The first section discussed the development of Prototype 1 which can detect emotions from facial expressions and text separately, followed by the next section which discussed the implementation of Prototype 2 which eliminated the limitations of Prototype 1. Then, another prototype was developed which has embedded an ANN in the application that was trained using the data collected during the evaluations of the previous prototypes. It paved the way to the conclusion of the first research question. Then Prototype 4 was developed which has the functionality to suggest functions based on the current user emotion and the context parameters. To eliminate the limitation of the time constraint of user evaluations, an improvement has been made and developed as Prototype 5 by considering the frequently used functions. It paves the path to the conclusion of the second research question. For each prototype, the evaluation details and the results will be discussed in the next chapter.

Chapter 5 - Results and Evaluation

This chapter mainly describes the details related to the results and the evaluation process of the research. According to the methodology described in [Section 1.5](#), the evaluation was done in each iteration. Since this methodology has incorporated User-Centered Design (UCD) [30], the evaluation was mainly done using prototype evaluations. In each iteration, the results from the previous prototype evaluation were compared and the decisions were taken to improve the next iteration. In this chapter, the evaluation and results of each prototype will be discussed from [Section 5.1](#) to [Section 5.5](#), which finally paves the path to the conclusion of the research.

5.1 Prototype 1

The 1st iteration of the research design started by implementing Prototype 1 which was discussed in [Section 4.1](#). After the implementation, the evaluation was performed with user participation which will be explained throughout this section.

5.1.1 Participant Selection and Experiment Setting

A focus group of **40 participants** was selected using convenient sampling and expert sampling [29] with informed consent. (The informed consent form can be found in [Appendix C](#)) Three rooms with adequate lighting were mainly used to perform studies. A mobile phone with a 5-megapixel front camera was used for the studies.

5.1.2 Evaluation Procedure

The user evaluation has been mainly conducted by using two methods as mentioned below.

1. Showing videos to participants
2. Questioning (chatting with) the participants through Google Hangouts

Three videos were used where each of them mainly generated “joy”, “sadness”, and “anger” and the videos were shown one by one to the participant. If a notification comes while watching a video, the participant was asked to check it and give his/her feedback as mentioned in [Section 4.1.4](#). After watching one video, the participants were asked to chat using Google Hangouts and some unstructured questions about the videos were asked. Here also if a notification popped up, the same steps were asked to perform. After showing the three videos, another set of dynamic/unstructured questions (which changes with users’ answers) were asked using Google Hangouts again. If a notification comes while chatting the same steps were asked to perform. Evidence for prototype evaluations can be found in [Appendix F.1](#).

5.1.3 Analysis

492 emotions (records) were recorded, where 100 were from facial expressions and 392 were from text analytics.

As shown in Fig. 5.1, when considering the emotions from facial expressions, almost all the emotions were predicted in a wrong manner, because the orange bars, which depicts the “NO percentages” have gone high. Only joy was detected correctly.

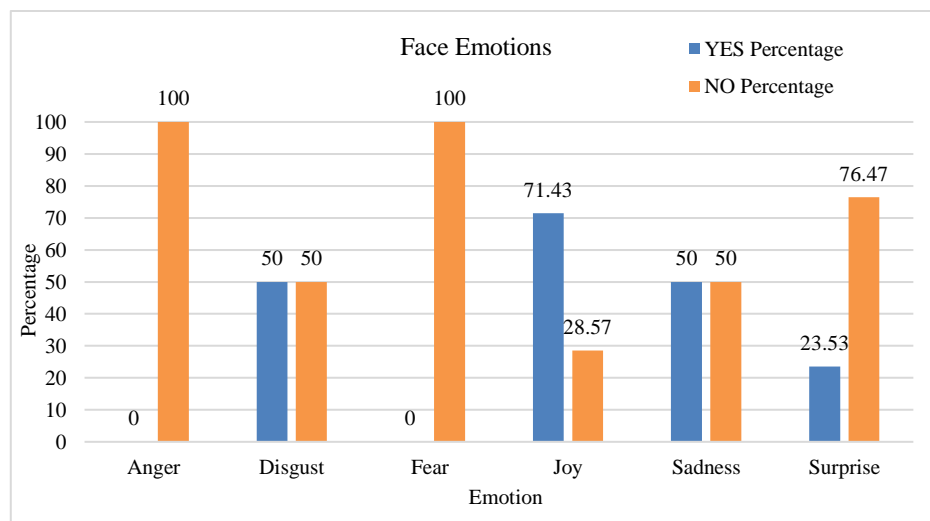


Figure 5.1: Prototype 1 - Face emotions analysis

As shown in Fig. 5.2, when considering the emotions which were detected by text, almost all the emotions were predicted correctly because the blue bars have gone high.

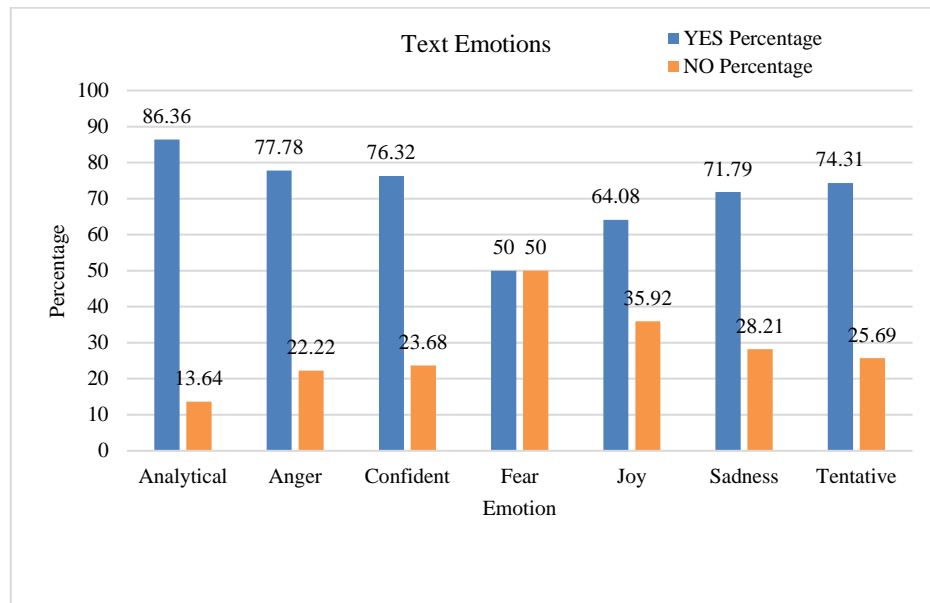


Figure 5.2: Prototype 1 - Text emotions analysis

5.1.4 Observations

Emotions from facial expressions were not detected accurately when using the mobile phone. The face was not captured and the emotions were not detected when the phone was rotated to the landscape mode while watching the videos. The participants were asked to hold the phone in the portrait mode. Most of the emotions were detected as incorrect for the participants who were wearing spectacles. Sometimes, the emotion was detected as “angry” because the frame of the spectacles was detected as brow lowering which resulted in detecting the facial muscles such as Depressor Glabellae, Depressor Supercilii, Corrugator [40].

“sadness”, “anger”, and “fear” were hard to capture from the face because generally, the participants did not express that kind of emotion using facial expressions when interacting with the mobile device. Even when a participant does not express any emotion, the app decided that the participant has “disgust”. “joy” was easily captured from the face when the participants were smiling. When the participants were raising

eyebrows or opening mouth, the emotion was detected as “surprise”, which is 76.47% wrong as shown in Fig. 5.1.

5.1.5 The Need for Prototype 2

Analyzed results of Prototype 1 pave the way for three needs to be covered from the next prototype. Table 5.1 discusses the need for the 2nd prototype and the solution that has been incorporated to improve the practical applicability of the prototype.

Table 5.1: The Need for Prototype 2

Need	Solution
1. A method was needed to give high priority to emotions from the text.	When the emotions from both text and face are available, consider emotions only from the text.
2. Detect emotions from the face, only if text emotions cannot be detected.	If a particular emotion is detected from the face, and the confidence is greater than or equal to a predefined threshold (50%), then express the particular emotion. The threshold value was decided through trial and error during the evaluations.
3. Need to reduce the number of mistakes when deciding the emotions from the text.	Define a predefined threshold (50%) for text too, and express the emotion only if the confidence is greater than that value.

5.2 Prototype 2

The 2nd iteration of the research design started with the implementation of Prototype 2, which was discussed in [Section 4.2](#) and [Section 5.1.5](#). After the implementation, evaluation was carried out with user participation where participant selection, experiment setting, and the evaluation procedure were similar to Prototype 1 which were explained in [Section 5.1](#). The analysis of the results and observations of Prototype 2 and the road to Prototype 3 will be explained throughout this section.

5.2.1 Analysis

502 emotions (records) were retrieved where 59 were from facial expressions and 443 were from text analytics.

As shown in Fig. 5.3, the number of emotions detected from facial expressions has been reduced to 59, which proves that the priority that has been given for the emotions from the facial expressions was reduced.

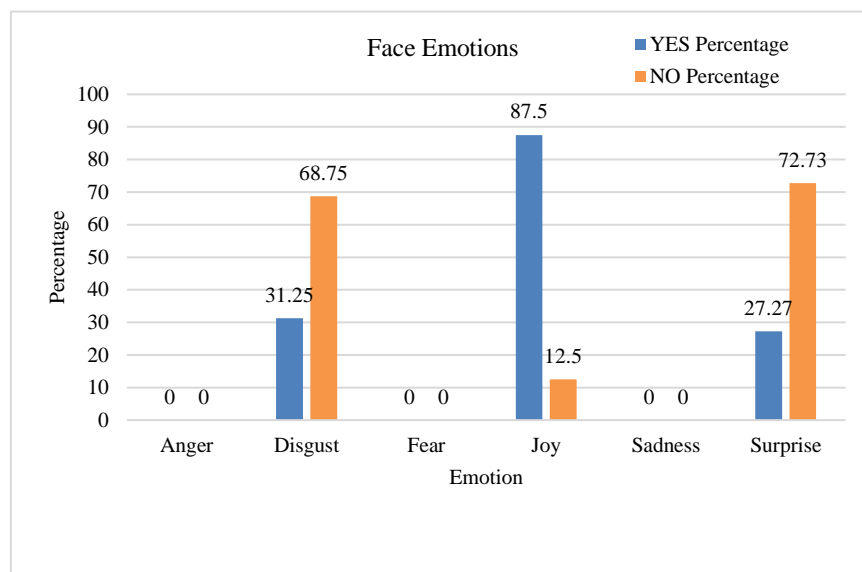


Figure 5.3: Prototype 2 - Face emotions analysis

As shown in Fig. 5.4, when considering the emotions from the text, the blue bar has gone higher than in the analysis of Prototype 1, thus proving that all the emotions have been almost detected correctly.

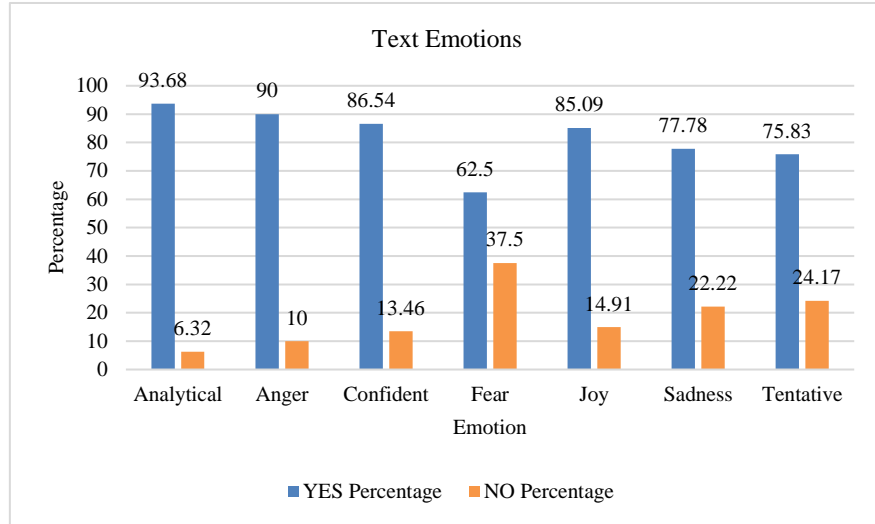


Figure 5.4: Prototype 2 - Text emotions analysis

5.2.2 Comparing Prototype 1 and Prototype 2 Using Precision

Precision expresses the proportion of the observed data that the model says was relevant, actually were relevant [62] which can be calculated using the eq. (6).

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

During the Prototypes 1 and 2 evaluations, the precision for each emotion was calculated with the YES and NO counts as shown in Tables 5.2 and 5.3 using the eq. (7).

$$Precision = \frac{YES\ Count}{YES\ Count + NO\ Count} \times 100\% \quad (7)$$

Table 5.2: Prototype 1 - Precision Values

Emotion Type	Emotion	YES count	NO count	Precision (%)
Face	anger	0	1	0
	disgust	21	21	50
	fear	0	3	0
	joy	25	10	71.43
	sadness	1	1	50
	surprise	4	13	23.53
Text	analytical	76	12	86.36
	anger	7	2	77.78
	confident	29	9	76.32
	fear	3	3	50
	joy	66	37	64.08
	sadness	28	11	71.79
	tentative	81	28	74.31

Table 5.3: Prototype 2 - Precision Values

Emotion Type	Emotion	YES count	NO count	Precision (%)
Face	anger	0	0	0
	disgust	10	22	31.25
	fear	0	0	0
	joy	14	2	87.5
	sadness	0	0	0
	surprise	3	8	27.27
Text	analytical	89	6	93.68
	anger	9	1	90
	confident	45	7	86.54
	fear	10	6	62.5
	joy	97	17	85.09
	sadness	28	8	77.78
	tentative	91	29	75.83

Justification to use Precision

In here the user was the person who decided whether a particular emotion is correctly detected or not. Since the user was given the high priority, the decision-maker was the user. The user could choose whether an emotion was correct or not, and if the emotion was not correct, he/she was asked to enter the actual emotion he/she was feeling at that moment. When considering a person, he/she would express that emotion regardless of the method which he/she uses to express it. If someone is happy (“joy”), he/she is “happy” with regardless of which way he/she expresses it. It can be through facial expressions or text data or anything.

For further explanation, assume that the system says, “You are happy. Keep it up!”. If the user was not satisfied, he/she would select NO, then the actual emotion should be selected. (Check [Section 4.1.4](#) for more information) But in here the user cannot say in which way he/she has felt that emotion because, for a human, the emotion is the same whether he/she has expressed it from face or text. This fact implies that it is not practical to get the way/method that the user has felt the actual emotion, instead only the NO count was incremented. In other words, the user cannot say whether a particular emotion was expressed from face or text and there is no method to record the misclassified data with the incorrectly predicted label. Even though the user cannot decide the way the emotion was generated, when a correct classification was done, after the user has pressed YES, **the system knows** from which way **it** decided the emotion.

Assuming the detected emotion belongs to the positive class, the YES count corresponds to the **True Positives** (where a positive sample is predicted as positive) and NO count corresponds to the **False Positives** (where a negative sample is predicted as positive).

The results in Tables 5.2 and 5.3 show that the Precision value for all the emotions in Prototype 2 is higher than in Prototype 1, except for “disgust”. Thus, can be proven the Prototype 2 has performed better than Prototype 1.

5.2.3 The Need for Prototype 3

The evaluation results of Prototype 2 which were discussed in [Section 5.2.1](#) proved that the need for Prototype 2 was satisfied up to some extent. Using the data which was collected during the previous prototype evaluations, an ANN was trained and embedded in the next prototype. This prototype aimed proposing a more practical system because the data that have been used consists of the behavioural factors and the compatibility of the context. As a result, a new prototype was implemented as Prototype 3 which was discussed in [Section 4.3](#) and the results of the Prototype 3 evaluation will be discussed in the next section.

5.3 Prototype 3

The 3rd iteration of the research design started with the implementation of the ANN to be used in Prototype 3, which was discussed in [Section 4.3](#). After the configuration of the ANN, the training and validation process will be explained in [Section 5.3.1](#). After that, the model was embedded into the Android app and the evaluation (testing) was carried out with user participation where participant selection, experiment setting and the evaluation procedure were similar to the Prototype 1 which were explained in [Section 5.1](#). The results of training and validation of the ANN will be discussed in this section along with the design of a confusion matrix which was done after the user evaluation of Prototype 3.

5.3.1 Training and Validating the ANN

As mentioned in [Section 4.3.1](#), the model was trained with a batch size of 10 and 200 epochs and Adam optimizer [61] with a learning rate of 0.001 was used. Categorical cross-entropy was used as the loss function and 70% data was used to train the model while 30% was used to do the validation. As shown in Fig. 5.5, the training accuracy has converged to 75.04% while validation accuracy has converged to 78.17%. As shown in Fig. 5.6, the training loss has converged to 0.89 while the validation loss has converged to 0.7658 which clearly states that the model has not overfitted.

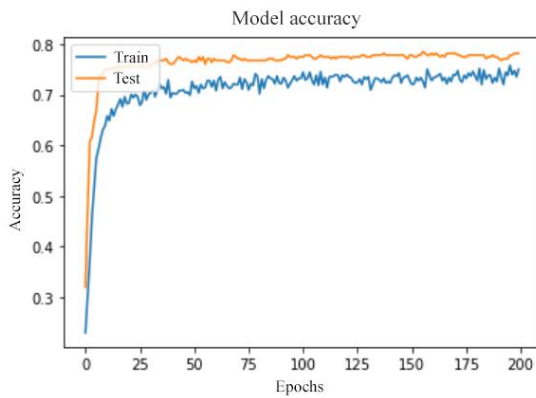


Figure 5.5: Prototype 3 - Training accuracy vs. Validation accuracy

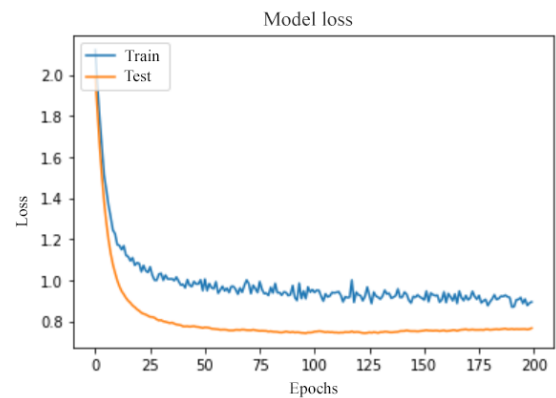


Figure 5.6: Prototype 3 - Training loss vs. Validation loss

5.3.2 Confusion Matrix and Interpretation

The user evaluation was done using the same procedure and with the same set of participants as mentioned in [Section 5.1](#), but the data was recorded in a slightly different manner when compared with previous prototype evaluations as mentioned in [Section 4.3.3](#), in a way that is easy to create a multiclass confusion matrix as shown in Fig. 5.7. Performance measures such as Accuracy, Precision, Recall (Sensitivity/TP Rate), Specificity (TN Rate) and F-Score (F-Measure) for each emotion was calculated [63].

		PREDICTED CLASS								
		analytical	anger	confident	disgust	fear	joy	sadness	surprise	tentative
KNOWN CLASS	analytical	139	0	0	3	0	0	0	0	0
	anger	2	8	0	0	0	0	0	0	1
	confident	1	0	27	0	0	0	0	0	0
	disgust	0	0	0	9	0	1	0	0	0
	fear	0	0	0	0	2	0	0	0	0
	joy	2	1	0	1	0	61	0	0	2
	sadness	14	0	1	7	0	1	18	0	3
	surprise	0	0	0	0	0	0	0	0	0
	tentative	14	0	0	1	0	1	3	0	36

Figure 5.7: Prototype 3 - Confusion matrix for evaluation

General Observations

Most of the time the model has predicted the emotion as “analytical”, which was correct but the model has incorrectly predicted “disgust”. “surprise” was not detected at all from the model. Also, the users have not selected “surprise” when a wrong prediction was done by the model. This means, that “surprise” is not an emotion to be considered in this kind of scenario (when using mobile phones). The reason for this might be, “surprise” is mostly part of “joy”, which can be proven by a large number of predictions of “joy” which was correctly predicted.

Accuracy

As discussed in [Section 5.3.1](#), the training accuracy was converged to 75.04% and the validation accuracy was converged to 78.17%. According to the confusion matrix which was prepared after the prototype evaluations, **83.57%** accuracy was shown by applying the model practically, which is an improvement. Below is the **testing accuracy** that was able to get from this model calculated using eq. (8).

$$\begin{aligned} Accuracy &= \frac{TP + TN}{TP + FP + FN + TN} \times 100\% & (8) \\ &= \frac{300}{359} \times 100\% \\ &= 83.57\% \end{aligned}$$

Precision

Precision indicates the proportion of the instances that the model says was relevant, actually were relevant [62]. For example, “How many observations that were predicted as sadness is **actually** sadness?”. Precision values for each emotion were calculated as shown in Table 5.4 using the eq. (6) which was mentioned in [Section 5.2.2](#) and only “disgust” had a low value.

Table 5.4: Prototype 3 - Precision for Each Emotion

Emotion	Precision
analytical	80.81%
anger	88.89%
confident	96.43%
disgust	42.86%
fear	100%
joy	95.31%
sadness	85.71%
surprise	-
tentative	85.71%

Recall (Sensitivity/TP Rate)

Recall indicates the proportion of actually positive instances that are correctly predicted [62]. For example, “Of all observations which are **truly** sadness, how many we correctly predicted?”. Recall values for each emotion were calculated as shown in Table 5.5 using the eq. (9) and both “sadness” and “tentative” had low values.

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

Table 5.5: Prototype 3 - Recall for Each Emotion

Emotion	Recall
analytical	97.89%
anger	72.73%
confident	96.43%
disgust	90%
fear	100%
joy	91.04%
sadness	40.91%
surprise	-
tentative	65.45%

Specificity (TN Rate)

Specificity indicates the proportion of real negative instances that are correctly predicted negative [62]. For example, “Of all observations which are **not** sadness, how many we correctly predicted?”. Specificity values for each emotion were calculated as shown in Table 5.6 using the eq. (10) and there is high specificity for every emotion.

$$Specificity = \frac{TN}{TN + FP} \times 100\% \quad (10)$$

Table 5.6: Prototype 3 - Specificity for Each Emotion

Emotion	Specificity
analytical	84.79%
anger	99.71%
confident	99.7%
disgust	96.56%
fear	100%
joy	98.97%
sadness	99.05%
surprise	100%
tentative	98.03%

F-Score (F-Measure)

F-Score considers both precision and recall. It is the harmonic mean (average) of the Precision and Recall. F-Score is best if there is some sort of balance between Precision and Recall in the system [62]. F-Score values for each emotion was calculated as shown in Table 5.7 using the eq. (11) and got low values for “disgust” and “sadness”.

$$F - Score = 2 \times \frac{(Recall \times Precision)}{(Recall + Precision)} \times 100\% \quad (11)$$

Table 5.7: Prototype 3 - F-Score for Each Emotion

Emotion	F-Score
analytical	88.54%
anger	80%
confident	96.43%
disgust	58.06%
fear	100%
joy	93.13%
sadness	55.38%
surprise	-
tentative	74.23%

5.3.3 Performance Measures of the Emotions “joy”, “sadness” and “anger”

According to User Survey 1 as mentioned in [Section 3.2.2](#), the respondents have stated “joy”, “sadness” and “anger” are the important emotions to be tracked to provide adaptive functionalities. The performance measures of these emotions are shown in Table 5.8.

Table 5.8: Prototype 3 - Performance Measures for “joy”, “sadness” and “anger”

Emotion	Accuracy (General)	Precision	Recall	Specificity	F-Score
joy	83.57%	95.31%	91.04%	98.97%	93.13%
sadness		85.71%	40.91%	99.05%	55.38%
anger		88.89%	72.73%	99.71%	80%

The limitation of this model is the low values for Recall and F-score for “sadness”. According to the Recall, “of all observations which are truly sadness only 40.91% was correctly predicted”. The low value of Recall affected the F-Score value of the emotion too. The need of increasing the Recall value for emotion “sadness” can be done if the model can be further trained using more data or implementing a method to

dynamically train the model through Active Learning [64] which paves the way to future research.

Other values show that this model is practically applicable when detecting emotions through facial expressions and text analytics to get an optimal emotion by combining those two techniques, which confirms the first research question as mentioned in [Section 1.2.2](#). Further prototypes will be focusing mainly on the second research question.

5.4 Prototype 4

The 4th iteration of the research design started with the implementation of Prototype 4, which was discussed in [Section 4.4](#). After the implementation, the evaluation was carried out with user participation which will be explained throughout this section.

5.4.1 Participant Selection and Evaluation Procedure

Out of the 40 participants, **nine males** and **nine females** were selected using convenient sampling [29]. Since the consent was taken previously, they agreed to participate in this prototype evaluation too. The main reason for limiting the number from 40 to 18 was the time constraint.

The users were asked to install the application in their mobile devices and to do the registration process as mentioned in [Section 4.4.1](#). After that, they were asked to use the “Emotional Keyboard” as much as possible in their daily use and the data (feedback) was recorded in the Firebase Real-time database as discussed in [Section 4.4.7](#) for 30-45 days.

5.4.2 Analysis

After the user evaluations, the YES count which depicts that the model has suggested a particular function correctly and the NO count which represents that the model has

suggested a wrong function, were plotted with the time instance as shown in Fig. 5.8. The **Time Instance** axis starts from zero and keeps increasing. When a particular user has responded to a notification for the first time, it will be the time instance one. Then, when he/she again responds, it will be the time instance two. Similarly, this keeps on increasing one unit at a time for each response from the user. For different users, the total number of time instances will differ based on their usage of the application.

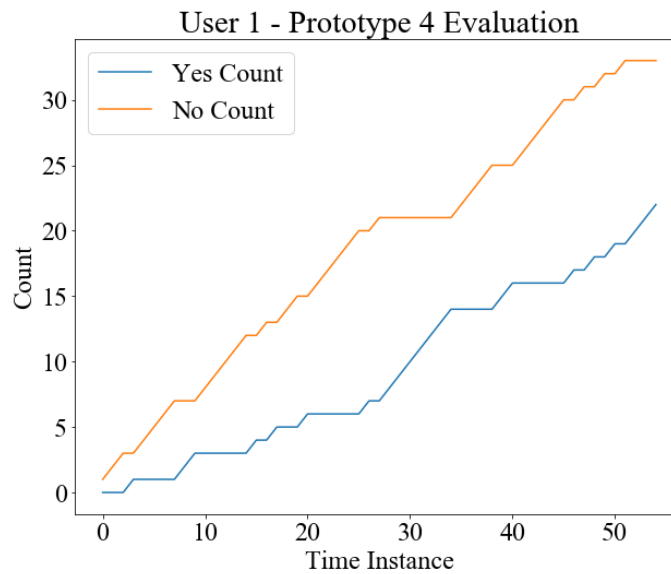


Figure 5.8: Prototype 4 - User 1 evaluation

Fig. 5.8 shows the plot for User 1, which depicts that with time the YES count was increased but it could not exceed the NO count. This means the **model may take more time to suggest the functions correctly**.

Another two graphs for two users have been attached in [Appendix F.4](#), where the experiment was successful for the User 2 by the increment of YES count, but for the User 3, the YES count was increased with time but it could not exceed the NO count (Similar to what happened to User 1 as explained before).

5.4.3 The Need for Prototype 5

The above analysis shows that this model takes more time to produce the correct suggestions which are not practical to evaluate. Due to this limitation, another quick

approach was required which can manipulate this time constraint. It was developed as Prototype 5 as discussed in [Section 4.5](#).

5.5 Prototype 5

The same evaluation procedure was carried out as mentioned in [Section 5.4.1](#) with the same 18 participants. They were asked to use the prototype for 30-45 days and the data was recorded as explained in [Section 4.4.7](#).

5.5.1 Individual Analysis

For each participant, the analysis was done by plotting the YES count (which depicts that the model has suggested a particular function correctly) and the NO count (which represents that the model has suggested a wrong function) with time. The individual analysis can be divided into three subsections as follows.

1. Plots with one intersection point in between
2. Plots with more than one intersection point in between
3. Plots without an intersection point in between (Intersection point is at the end of the plot)

In all these plots, intersection points depicts the moment for a particular user, that the number of YES count (which symbolizes the situations that correctly predicting adaptive activities) have overrun the NO count (which symbolizes the situations that incorrectly predicting adaptive activities). It can be observed that for all the users there have been one or more situations where the YES count has overrun the NO count which will be discussed below.

Plots with one intersection point in between

Out of 18 users, User 1, 3, 13, 14, 17 and 18 have only one intersection point in between. (Please refer [Appendix F.5](#) for all the plots) The following observations can be made when interpreting the plots for the above-mentioned users.

- **User 1** – The intersection point has occurred in a shorter period of time (in between the time instances zero and five). It means that the model has predicted the correct function which the user desires as soon as the user has started to use the prototype. Furthermore, with time the YES count has gone higher and higher while the NO count has gone lower.
- **User 3** – Similar to User 1 the intersection point has occurred as soon as the user has started to use the prototype. Before occurring the intersection point, the YES and NO counts were increasing at a similar rate but at some point, the YES count has overrun the NO count. In between the time instance 20 and 30, there has been a situation where the YES and NO counts again have become almost similar, but again in a short period of time, the YES count has increased even more.
- **User 13** – In the beginning, the model has predicted wrong functions since the NO count has been increased. In between time instances 10 and 20, the YES count has been increased until the end, which means the model has done correct predictions after the intersection point.
- **User 14** – Similar to User 13, in the beginning, the model has done incorrect predictions until the time instance 15. It has been taken a long time to YES count to overrun the NO count for this particular user.
- **User 17** - Similar to User 14, in the beginning, the model has done incorrect predictions until the time instance between 10 and 15. It has been taken a long time to YES count to overrun the NO count for this particular user too.
- **User 18** – Similar to User 3 the intersection point has occurred as soon as the user has started to use the prototype. Before occurring the intersection point, the YES and NO counts were increasing at a similar rate, but at the time instance five, the YES count has overrun the NO count. In between the time instance five and 10, there has been a situation where the YES and NO counts again have become almost similar, but again in a short period of time, the YES count has increased even more.

Plots with more than one intersection point in between

This is the most common observation for most of the users. Out of 18 users, User 2, 4, 6, 7, 8, 9, 10, 11, 12, 15 and 16 have more than one intersection point in between. (Please refer [Appendix F.5](#) for all the plots).

Out of the above-mentioned users, User 4, 5, 7, 10 and 11 only have two intersection points while the others have more than two. The reason for more than one intersection point is the continuous increment of YES and NO counts in nearby time instances. This depicts the dynamic behaviour of the model which was trying to do the predictions correctly as much as possible (when the YES count was increased for a particular moment between two intersection points) but consequently, the user decided that not to do the particular activity (when the NO count was increased for a particular moment between two intersection points). This type of more dynamic behaviour can be seen for Users 6, 8, 12 and 16 since the plots of those users have six or seven intersection points that are closer to each other.

Plots without an intersection point in between (Intersection point is at the end of the plot)

This is a rare scenario. Out of 18 users, only User 5 grouped into this category. It can be seen that there is only one intersection point for this user, which is at the end of the plot near the time instance 20. (Please refer [Appendix F.5](#) for the plot). This shows that until the time instance 20 the YES count was not able to overrun the NO count. At the time instance 20 both YES and NO counts have got the same value. As it can be seen that the rate of the increment of the YES count was getting higher during the time instances 15 and 20, which shows that after the intersection point there is a high probability that the YES count may overrun the NO count at some time instance in future.

As explained under the above-mentioned topics, it has been proven that **the model can dynamically predict the adaptive functions without halting the learning process**. This leads to the concept named “An Adaptive System with User Control” [21] where the user can correct the wrongs that the model has been doing since the user has control over it. This is suitable in dynamic scenarios, as this has been tested when the

users were performing their day to day activities. There can be situations in which this model cannot capture the changes suddenly since those deviates from the users' day to day activities. For example, some days a user may need to “Browse Internet” during his/her day instead of doing his/her usual activities which are the reasons for the dynamic behaviour with more than one intersection point (mostly six to seven intersections points) as explained above. Since this dynamic behaviour of the model is difficult to be interpreted, the overall outcome will be discussed in the next section.

5.5.2 Overall Analysis

For all the users, the total YES count and the total NO count were averaged for each time instance using the eq. (12).

$$\begin{aligned}
 & \text{Average YES or NO Count for time instance } t \\
 &= \frac{\sum_{i=0}^n (\text{YES or NO Count for the user } i \text{ for the time instance } t)}{\text{Total number of users available in time instance } t}
 \end{aligned}
 \tag{12}$$

For all the users the total YES count and the total NO count were averaged for each time instance and were plotted as shown in Fig. 5.9 which consists of the overall analysis.

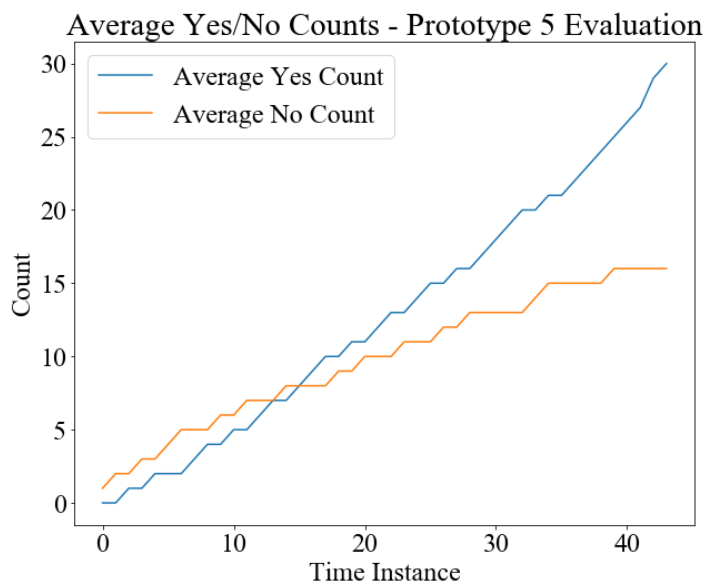


Figure 5.9: Prototype 5 - Overall analysis

This describes the general analysis of Prototype 5 which interprets how the model behaves generally irrespective of the user. Fig. 5.9 is similar to the plots which were plotted for each individual user. (Please refer [Appendix F.5](#) for all the plots)

Intersection points in the overall analysis

Two intersection points have been identified between the time instances 10 and 20. These points depict how much time the average YES count has been generally taken to overrun the average NO count. It means the out of the total 40 time instances (since the maximum of the x-axis is 40), the model has started to do the predictions correctly in a reasonable amount of time (that is between the time instances 10 and 20).

Rate of the YES/NO counts

Until the time instance 10 (between 10 and 20), the average YES count and average NO count have increased at a similar rate, which means the model has done the correct predictions as well as the incorrect predictions with a similar rate. Furthermore, in that range, the average NO count is higher than the average YES count.

Thus, it can be proved that, in between the time instances 10 and 20, average YES count has overrun the average NO count and the rate of the average YES count has gone higher than the average NO count in consequent time instances until the end of the evaluation. It generally gives the inference that the model has started working properly for any user.

5.6 Summary

This chapter mainly explained the results and evaluation of the proof of concept through five prototypes. The first section discussed the results and the decisions taken from the evaluation of Prototype 1, followed by the next section which discussed the improvement of the results using Prototype 2 which eliminated the limitations of Prototype 1 and a comparison of the results of both these prototypes have been performed. Evaluation of Prototype 3 was discussed using a multiclass confusion matrix and performance measures such as accuracy, precision, recall, specificity and f-score. It paved the way to the conclusion of the first research question. The evaluation

of Prototype 4 was discussed next, but it was unsuccessful as that model required more time to do the correct predictions which was not practical. As a solution, the evaluation of the Prototype 5 was done which have concluded this chapter by giving an inference of how the model performs with respect to each individual (individual analysis) and in general (overall analysis). The next chapter will discuss the conclusion of the research from the insights gained through this chapter.

Chapter 6 - Conclusions

This chapter discusses the conclusions of the research that have obtained from the results of the analysis. Section 6.1 will provide the answers to the research questions and will review the aims and objectives. After that, Section 6.2 will highlight the conclusions obtained for the research problem followed by Section 6.3 which will explain the research contribution to the fields of Computer Science and Human-Computer Interaction. Consequently, the limitations will be explained in Section 6.4 and finally, the chapter ends by giving details about the further implications of this research in Section 6.5.

6.1 Conclusions about the Research Questions (Aims and Objectives)

The main aim of this research was to develop a model for a mobile device that will provide the ability to take more optimal decisions about the functionalities, by considering user emotions based on the current context in real-time, which enhances the acceptability and usability of that particular system. Two research questions were identified and the implementation and evaluation of the prototypes were done as discussed in previous chapters.

Three prototypes were used to achieve the objectives specified under the first research question “*How to aggregate different emotion detection techniques to decide the most meaningful emotion?*”. First, User Survey 1 strengthened the decisions to implement the prototypes as mentioned in Section 3.2.2. The base of the prototypes was a keyboard named “Emotional Keyboard” compatible with Android devices, which can track emotions from facial expressions and text. Prototype 1 was implemented by detecting emotions from facial expressions using Affectiva SDK [26] and detecting emotions from text using IBM Watson Tone Analyzer API [25] separately, thus achieving the first two objectives under the first research question. Detecting emotions from facial expressions consisted many drawbacks when practically carrying out the

user evaluations due to the problems in camera orientation, uneven illumination and misjudged emotions as discussed in [Section 5.1.4](#) which led to the implementation of the Prototype 2 which gave more focus to detect emotions from text, rather than facial expressions, by using threshold values. At this point the solution to the first research question was already attained, but an ANN was developed using Keras (with TensorFlow) and embedded into the Prototype 3 as discussed in [Section 4.3](#) using practically collected data during the evaluations of the Prototypes 1 and 2. It has given 75.04% training accuracy and 78.17% validation accuracy. From the data collected during the user evaluations of this prototype, a confusion matrix was developed and calculated the Accuracy, Precision, Recall, Specificity and F-Score values for each emotion. The calculated testing accuracy was 83.57%, which implies the practical usage of this model is better when comparing with training and validation accuracies, thus leads to the completion of the first research question.

Finding the answer to the second research question *“How a new model can be developed to do optimal decision making in order to perform affective interactions by considering user emotions based on the context?”*, was started by conducting User Survey 2 as mentioned in [Section 3.2.4](#). A Preference Tree was created using the data from the User Survey 2 which consists of the probabilities that a particular user performs a function based on his/her current emotion and the context. For each user, a Preference Tree was separately created for his/her profile and the functions were suggested by traversing the branches of the tree as discussed in [Section 4.4.6](#) based on the user emotion and the context. The user evaluations of Prototype 4 have been carried out as mentioned in [Section 5.4.1](#), where the user can respond to the suggested activity by choosing whether he/she likes or not to use that. This model was not practical to use in day to day life since it took more time to produce the correct suggestions. As an improvement, Prototype 5 was developed by keep tracking frequently used functions as discussed in [Section 4.5](#). The user evaluations were carried out with the participation of 18 users (nine males and nine females) and the user feedback was recorded. With time, the model was able to correctly suggest the user activities which can be proved by the increment of the number of “YES”s (User agreed to use the activity) as discussed in [Section 5.5](#) when considering the analysis for each and individual user and as well as generally, thus can come to a conclusion

that this model can be accepted as an adaptive system with user control [21] which can suggest functions based on the user emotion and the context.

6.2 Conclusions about the Research Problem

The research problem was stated in [Section 1.2.1](#) and the justification was thoroughly explained in [Section 1.4](#). According to those, due to the popularity of smart technologies, mobile phones have become a part of our day to day lives and improving the UX of mobile devices has vital importance. To improve usability, this research has suggested an adaptive system with user control [21], which can suggest functions by considering the current emotion of the user and the context. Using Prototype 5 it has been proven that the users are satisfied with the system which represents a model that consists of a Preference Tree and frequent functions. According to the best of our knowledge, no research has been conducted by considering both emotions and context to provide not only adaptive interfaces but also to provide adaptive functionalities, but this research has proven that those factors too can be mixed as multi-dimensional factors when deciding adaptive functions. Since emotion detection by combining both facial expressions and text analytics has not been carried out before, the Prototypes 1 to 3 were formulated to test whether combining these two methods is practically applicable. It has been proven that detecting emotions from facial expressions have many drawbacks than text, but to eliminate those obstacles an ANN was developed which can detect emotions from both of these methods. The full research contribution will be discussed in the next section which will elaborate more on the conclusions of the research problem.

6.3 Research Contribution

6.3.1 Contribution to CS

The main contribution of this research is the introduction of the new model by combining the user emotions with the context to make decisions by providing the users with more affective functionalities. Here the Preference Tree which was developed using User Survey 2 (Refer [Section 4.4.1](#)) was the base of the model that has been

proposed by this research. During the user evaluations of Prototype 4 and 5, it has been proven that generating suggestions by traversing the branches of the Preference Tree and retrieving the probabilities and updating them based on the user feedback (Adaptive system with user control [21]) was effective. Additionally, by incorporating the probabilities by tracking frequent functions [27], more robust suggestions were able to recommend to the user. Together with the Preference Tree, frequent functions and perceiving multi-dimensional factors such as user emotions, location, time and user activity form a model that can be further tested and improved (which will be discussed more in [Section 6.5](#)) in the future.

The above model was mainly formed from the second research question. When discussing the first research question, to the best of my knowledge no one has tested deciding the most meaningful emotion using an ANN which makes it a contribution too. Also, the practical applicability of each of the above findings was validated during the prototype evaluations thus leading to successful HCI research by identifying the user behaviour through observation.

6.3.2 Contribution to HCI

Human-Computer Interaction (HCI) is a multidisciplinary field that involves psychology, cognitive science, ergonomics, sociology, computer science and engineering, business management, arts and writing [65]. Contribution to the field of HCI means that it will cover all the other fields around it. In HCI, a system that was designed according to 5Es [66] can be considered as a usable system. Affective and adaptive systems play a main role when discussing usable systems. Improving the system functionalities by detecting the user emotions based on the context can be a major contribution to this research area, which promotes the characteristics in 5Es.

Since User-Centered Design (UCD) [30] was incorporated in the research methodology as explained in [Section 1.5](#), two gulfs of interactions [67] such as **Gulf of Execution** which improves the discoverability of the functions to the user and the **Gulf of Evaluation** which facilitates the feedback from the system to the user, were also consolidated when performing prototype evaluations using Evolutionary Prototyping

[31] by reducing the gap between the mental model of the user and the mental model of the designer thus leading to a valid HCI research.

6.4 Limitations

Since this research was conducted by introducing a keyboard as a proof of concept, the main focus was on the satisfaction of the user in the UX level, not the performance of the hardware device. Since this prototype uses several background processes and sensors to perceive multi-dimensional factors, the draining of battery happened instantly which should be eliminated if this will be introduced as a final product. Also, the used keyboard layout is deprecated from Android API level 29 (Pie) [68] because the Android Open Source Project (AOSP) insists the developers to use their own widget instead of using their inbuilt one to be more convenient.

6.5 Implications for Further Research

When providing adaptive functions this research only considered the user emotions and the context parameters such as location, time and user activity. Further research can be conducted to test whether how the physical characteristics of the user, sensory abilities, cognition can be considered in the context to make a more personalized profile for the user that can be used to suggest functions based on those. Besides, this research only considered suggesting functions for a particular emotion (along with the context) which the user will agree to perform, but it was not measured whether the negative emotion has changed to a positive one or not after performing that function. Further research can be carried out to discover this level of satisfaction with a more structured user evaluation process.

References

- [1] J. Forlizzi and K. Battarbee, "Understanding experience in interactive systems," *DIS2004 - Designing Interactive Systems: Across the Spectrum*, pp. 261–268, 2004.
- [2] L. Alben, "Defining the criteria for effective interaction design," *Interactions*, vol. 3, no. 3, pp. 11–15, 1996.
- [3] R. W. Picard, "Affective computing for HCI," *HCI*, vol. 1, no. August, pp. 829–833, 1999.
- [4] W. D. Wattarachchi, E. Hettiarachchi, and K. P. Hewagamage, "Critical Success Factors of Analysing User Emotions to Improve the Usability of Systems," in *2019 Twelfth International Conference on Ubi-Media Computing (Ubi-Media)*, 2019, pp. 91–95, doi: 10.1109/Ubi-Media.2019.00026.
- [5] R. W. Picard, "Toward Machines with Emotional Intelligence," pp. 1–22.
- [6] K. Dalvand and M. Kazemifard, "An Adaptive User Interface Based on User's Emotion," pp. 161–166, 2012.
- [7] M. Using, "Intelligent Modeling and Control of Washing," no. Ii, pp. 812–817, 2005.
- [8] E. Paul, "Facial expression and Emotion," *American Pshychologist Association*, vol. 49, pp. 384–392, 1993.
- [9] R. Plutchik, "A psychoevolutionary theory of emotions," *Social Science Information*, vol. 21, no. 4–5, pp. 529–553, 1982, doi: 10.1177/053901882021004003.
- [10] M. Kolodziej, A. Majkowski, R. J. Rak, P. Tarnowski, and T. Pielaszkiwicz, "Analysis of facial features for the use of emotion recognition," in *Proceedings of 2018 19th International Conference Computational Problems of Electrical Engineering, CPEE 2018*, 2018, pp. 1–4, doi: 10.1109/CPEE.2018.8507137.
- [11] C. Corneanu, F. Noroozi, D. Kaminska, T. Sapinski, S. Escalera, and G. Anbarjafari, "Survey on Emotional Body Gesture Recognition," *CoRR*, vol. abs/1801.0, 2018, doi: 10.1109/TAFFC.2018.2874986.
- [12] S. Porcu, A. Floris, and L. Atzori, "Towards the Prediction of the Quality of

- Experience from Facial Expression and Gaze Direction,” in *ICIN 2019 QoE Management Workshop, Paris, 2019*, pp. 82–87, doi: 10.1109/icin.2019.8685917.
- [13] S. Lugovic, I. Dunder, and M. Horvat, “Techniques and applications of emotion recognition in speech,” in *2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2016, pp. 1278–1283, doi: 10.1109/MIPRO.2016.7522336.
- [14] R. Shikder, S. Rahaman, F. Afroze, and A. B. M. A. Al Islam, “Keystroke/mouse usage based emotion detection and user identification,” *2017 International Conference on Networking, Systems and Security (NSysS)*, pp. 96–104, 2017, doi: 10.1109/NSysS.2017.7885808.
- [15] Z. Khalili and M. H. Moradi, “Emotion detection using brain and peripheral signals,” *2008 Cairo International Biomedical Engineering Conference, CIBEC 2008*, pp. 2006–2009, 2008, doi: 10.1109/CIBEC.2008.4786096.
- [16] M. Mukeshimana, X. Ban, N. Karani, and R. Liu, “Multimodal Emotion Recognition for Human-Computer Interaction: A Survey,” vol. 8, no. 4, pp. 1289–1301, 2017.
- [17] S. I. Filing, “Situation-dependent metaphor for personal multimedia information,” vol. 9, no. 6, pp. 725–743, 1999.
- [18] H. Lee, Y. Choi, and Y. Kim, “An Adaptive User Interface based on Spatiotemporal Structure Learning,” in *2011 IEEE Consumer Communications and Networking Conference (CCNC)*, 2011, pp. 923–927, doi: 10.1109/mis.2003.1193657.
- [19] M. Afzal *et al.*, “Model-based adaptive user interface based on context and user experience evaluation,” *Journal on Multimodal User Interfaces*, vol. 12, no. 1, pp. 1–16, 2018, doi: 10.1007/s12193-018-0258-2.
- [20] N. Sebe, I. Cohen, and Thomas S. Huang, “Multimodal Emotion Recognition,” *Handbook of Pattern Recognition and Computer Vision*, vol. 4, pp. 387–419, 2005, doi: 10.1142/9503.
- [21] F. Gullà, S. Ceccacci, M. Germani, and L. Cavalieri, “Design Adaptable and Adaptive User Interfaces: A Method to Manage the Information Design Adaptable and Adaptive User Interfaces: a Method to Manage the Information,” no. May, 2015, doi: 10.1007/978-3-319-18374-9.

- [22] S.-M. Wang, C.-H. Li, Y.-C. Lo, T.-H. K. Huang, and L.-W. Ku, "Sensing Emotions in Text Messages: An Application and Deployment Study of EmotionPush," pp. 1–6, 2016.
- [23] J. Park, S. Han, H. K. Kim, Y. Cho, and W. Park, "Developing Elements of User Experience for Mobile Phones and Services: Survey, Interview, and Observation Approaches," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 23, 2013, doi: 10.1002/hfm.20316.
- [24] M. Hirakawa, P. Hewagamage, and T. Ichikawa, "Situation-dependent browser to explore the information space," *Proceedings. 1998 IEEE Symposium on Visual Languages (Cat. No.98TB100254)*, pp. 108–115, 1998, doi: 10.1109/VL.1998.706153.
- [25] J. M. Garcia-Garcia, V. M. R. Penichet, and M. D. Lozano, "Emotion detection," *Proceedings of the XVIII International Conference on Human Computer Interaction - Interacción '17*, no. October, pp. 1–8, 2017, doi: 10.1145/3123818.3123852.
- [26] R. S. Deshmukh and V. Jagtap, "A survey: Software API and database for emotion recognition," *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 284–289, 2018, doi: 10.1109/ICCONS.2017.8250727.
- [27] Y. Fukazawa, M. Hara, M. Onogi, and H. Ueno, "Automatic cell phone menu customization based on user operation history," 2009, doi: 10.1527/tjsai.25.68.
- [28] "Mobile vs Desktop Traffic in 2019 | Perficient Digital | Perficient Digital Agency." <https://www.perficientdigital.com/insights/our-research/mobile-vs-desktop-usage-study> (accessed Dec. 27, 2019).
- [29] A. Bhattacharjee, *Social Science Research: Principles, Methods, and Practices*. 2012.
- [30] D. A. Norman and S. W. Draper, *User Centered System Design; New Perspectives on Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1986.
- [31] M. Carr and J. Verner, "Prototyping and Software Development Approaches," *Prototyping and Software Development Approaches*, no. 3, pp. 1–16, 2004.
- [32] A. Landowska, G. Brodny, and M. R. Wrobel, "Limitations of emotion recognition from facial expressions in e-learning context," *CSEDU 2017 -*

- Proceedings of the 9th International Conference on Computer Supported Education*, vol. 2, no. Csedu, pp. 383–389, 2017, doi: 10.5220/0006357903830389.
- [33] S. N. Shivhare and S. Khethawat, “Emotion Detection from Text,” *Computer Science & Information Technology*, vol. 2, 2012, doi: 10.5121/csit.2012.2237.
- [34] N. M. Hakak, M. Mohd, M. Kirmani, and M. Mohd, “Emotion analysis: A survey,” in *2017 International Conference on Computer, Communications and Electronics (COMPTELIX)*, 2017, pp. 397–402, doi: 10.1109/COMPTELIX.2017.8004002.
- [35] J. A. Russell, “A circumplex model of affect,” *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714.
- [36] A. Tellegen and D. Watson, “Toward a consensual structure of mood,” *Psychological Bulletin*, vol. 98, no. 2, pp. 219–235, 1985, doi: 10.1037/0033-2909.98.2.219.
- [37] S. Anwar and M. Milanova, “Real Time Face Expression Recognition of Children with Autism,” *International Academy of Engineering and Medical Research (IAEMR)*, no. 1, 2016.
- [38] M. Dubey and L. Singh, “Automatic Emotion Recognition Using Facial Expression: A Review,” *International Research Journal of Engineering and Technology (IRJET)*, vol. 3, no. 2, pp. 488–492, 2016.
- [39] K. J. Kantharia and G. I. Prajapati, “Facial behavior recognition using soft computing techniques: A survey,” in *2015 Fifth International Conference on Advanced Computing Communication Technologies (ACCT)*, 2015, pp. 30–34, doi: 10.1109/ACCT.2015.132.
- [40] M. A. Sayette, J. F. Cohn, J. M. Wertz, M. A. Perrott, and D. J. Parrott, “A Psychometric Evaluation of the Facial Action Coding System for Assessing Spontaneous Expression,” *Journal of Nonverbal Behavior*, vol. 25, pp. 167–186, 2002.
- [41] H. A., T. A., and A. T., “Face Behavior Recognition Through Support Vector Machines,” *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 1, pp. 101–108, 2016, doi: 10.14569/ijacsa.2016.070115.
- [42] L. A. M. Bostan and R. Klinger, “An Analysis of Annotated Corpora for

- Emotion Classification in Text Title and Abstract in German,” in *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 2104–2119.
- [43] C. Strapparava and R. Mihalcea, “Affective Text,” 2007.
- [44] A. F. M. N. H. Nahin, J. M. Alam, H. Mahmud, and K. Hasan, “Identifying emotion by keystroke dynamics and text pattern analysis,” *Behaviour and Information Technology*, vol. 33, no. 9, pp. 987–996, 2014, doi: 10.1080/0144929X.2014.907343.
- [45] A. Neviarouskaya, H. Prendinger, and M. Ishizuka, “Textual Attitude Analysis Model,” *Proc. of the NAACL HLT Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pp. 80–88, 2010.
- [46] P. Khanna and M. Sasikumar, “Recognising Emotions from Keyboard Stroke Pattern,” *International Journal of Computer Applications*, vol. 11, no. 9, pp. 1–5, 2010, doi: 10.5120/1614-2170.
- [47] A. Pentel, “Predicting Age and Gender by Keystroke Dynamics and Mouse Patterns,” no. July, pp. 381–385, 2017, doi: 10.1145/3099023.3099105.
- [48] A. Pentel, “Predicting user age by keystroke dynamics,” *Advances in Intelligent Systems and Computing*, vol. 764, no. April, pp. 336–343, 2019, doi: 10.1007/978-3-319-91189-2_33.
- [49] S. Chaffar and D. Inkpen, “Using a Heterogeneous Dataset for Emotion Analysis in Text,” *Springer*, pp. 62–67, 2011.
- [50] D. Benyon, P. Innocent, and D. Murray, “System Adaptivity and the Modelling of Stereotypes,” in *Human–Computer Interaction–INTERACT ’87*, 2014, pp. 245–253, doi: 10.1016/b978-0-444-70304-0.50047-9.
- [51] L. Rothrock, R. Koubek, F. Fuchs, M. Haas, and G. Salvendy, “Review and reappraisal of adaptive interfaces: Toward biologically inspired paradigms,” *Theoretical Issues in Ergonomics Science*, vol. 3, no. 1, pp. 47–84, 2002, doi: 10.1080/14639220110110342.
- [52] A. Ortony and G. Clore, “Review Reviewed Work (s): The Cognitive Structure of Emotions . by Andrew Ortony , Gerald L . Clore and Allan Collins Review by : B . N . Colby Published by : American Sociological Association Stable URL : <http://www.jstor.org/stable/2074241>,” no. January, 1988, doi: 10.2307/2074241.

- [53] I. Stathopoulou, E. Alepis, G. A. Tsihrintzis, and M. Virvou, “Knowledge-Based Systems On assisting a visual-facial affect recognition system with keyboard-stroke pattern information,” *Knowledge-Based Systems*, vol. 23, no. 4, pp. 350–356, 2010, doi: 10.1016/j.knosys.2009.11.007.
- [54] J. Ben Schafer, D. Frankowski, J. Herlocker, and S. Sen, “Collaborative Filtering Recommender Systems,” 2007, no. November.
- [55] “Get the last known location | Android Developers.” <https://developer.android.com/training/location/retrieve-current> (accessed Jan. 10, 2020).
- [56] “ActivityRecognitionClient | Google APIs for Android.” <https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognitionClient> (accessed Jan. 10, 2020).
- [57] E. Grossi and M. Buscema, “Introduction to artificial neural networks,” *European Journal of Gastroenterology and Hepatology*, vol. 19, no. 12, pp. 1046–1054, 2007, doi: 10.1097/MEG.0b013e3282f198a0.
- [58] “Request App Permissions | Android Developers.” <https://developer.android.com/training/permissions/requesting> (accessed Jan. 10, 2020).
- [59] K. Potdar, T. S., and C. D., “A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers,” *International Journal of Computer Applications*, vol. 175, no. 4, pp. 7–9, 2017, doi: 10.5120/ijca2017915495.
- [60] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of trends in Practice and Research for Deep Learning,” pp. 1–20, 2018.
- [61] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” pp. 1–15, 2014.
- [62] D. Martin Ward Powers, “Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation,” *Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011, doi: 10.9735/2229-3981.
- [63] C. Manliguez, “Generalized Confusion Matrix for Multiple Classes,” no. November, pp. 4–6, 2016, doi: 10.13140/RG.2.2.31150.51523.
- [64] K. Konyushkova, S. Raphael, and P. Fua, “Learning active learning from data,”

Advances in Neural Information Processing Systems, vol. 2017-Decem, no. Nips, pp. 4226–4236, 2017.

- [65] I. S. MacKenzie, *Human-Computer Interaction: An Empirical Research Perspective*, 1st ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2013.
- [66] W. Quesenbery, “Balancing the 5Es : Usability set your mind at Es,” *Cutter IT Journal*, vol. 17, no. 2, pp. 4–11, 2004.
- [67] D. A. Norman, *The Design of Everyday Things*. New York, NY, USA: Basic Books, Inc., 2002.
- [68] “Keyboard | Android Developers.” <https://developer.android.com/reference/android/inputmethodservice/Keyboard> (accessed Jan. 25, 2020).

Appendix A: User Survey 1

A.1 Questionnaire (Google Form)

Improving User Experience based on User Emotions when using Mobile Devices

It would be much appreciated if you can fill this simple questionnaire as a support for my final year research based on improving user experience based on user emotions when using mobile devices.

* Required

Age *

- Below 18 years
- 18 to 24 years
- 25 to 35 years
- Above 35 years

Gender *

- Female
- Male

Occupation *

- Student
- Working

When you are using a mobile device, you can feel many emotions due to various reasons. Assume there is a system which can provide you with feedback and guidance based on your emotion. Choose the emotions which you would like the system to track and provide feedback and guidance? *

- Happy
- Surprise
- Sad
- Angry
- Disgust
- Fear

In order to change your mood (emotion), what functionalities of a mobile device would you like to use? *

- Listen to Music
- Watch a Video (offline/online)
- Browse the gallery/photos of your phone
- Play a Game
- Chat with Friends (Viber, Whatsapp, Messenger)
- Make a Phone Call
- Browse Internet (Googling)
- Other: _____

Do those functionalities have an impact based on the time of the day? (Eg: In the morning you like to listen to music, but at night you like to chat with your friends) *

- Yes
- No

Do those functionalities have an impact on the location where you are at in that moment? (Eg: Work/University or Home) *

Yes

No

SUBMIT

Never submit passwords through Google Forms.

A.2 Results of User Survey 1

Question 1, 2, and 3 - Age vs. Gender vs. Occupation

Analysis of age, gender, and occupation of the respondents are shown in the Fig. A.1.

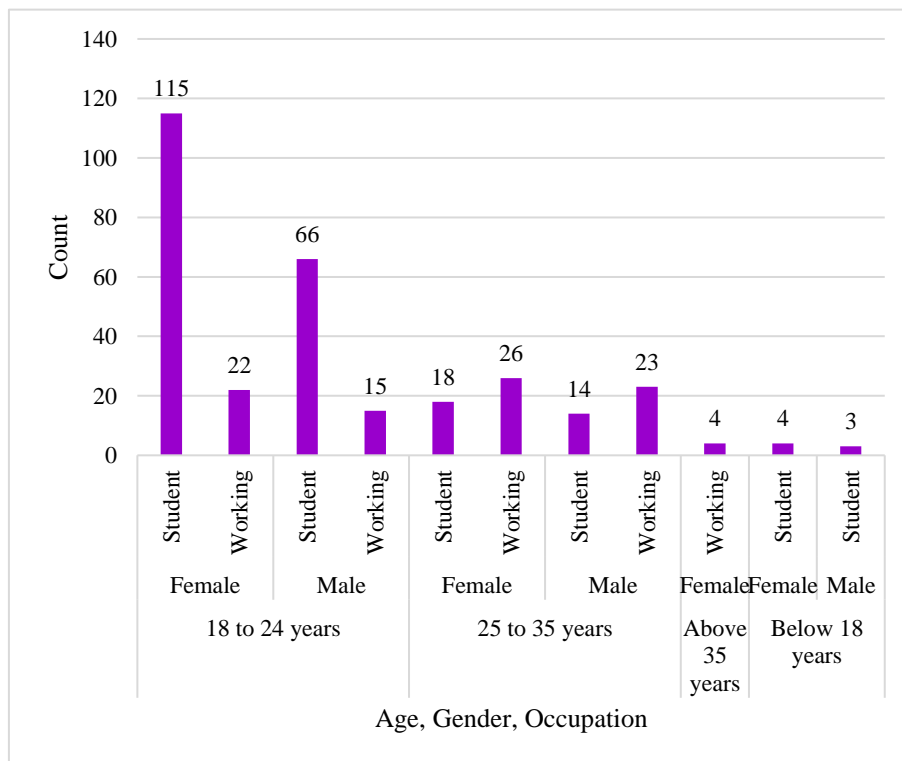


Figure A.1: Analysis of User Survey 1 - Questions 1,2 and 3

Question 4 - When you are using a mobile device, you can feel many emotions due to various reasons. Assume there is a system which can provide you with feedback and guidance based on your emotion. Choose the emotions which you would like the system to track and provide feedback and guidance?

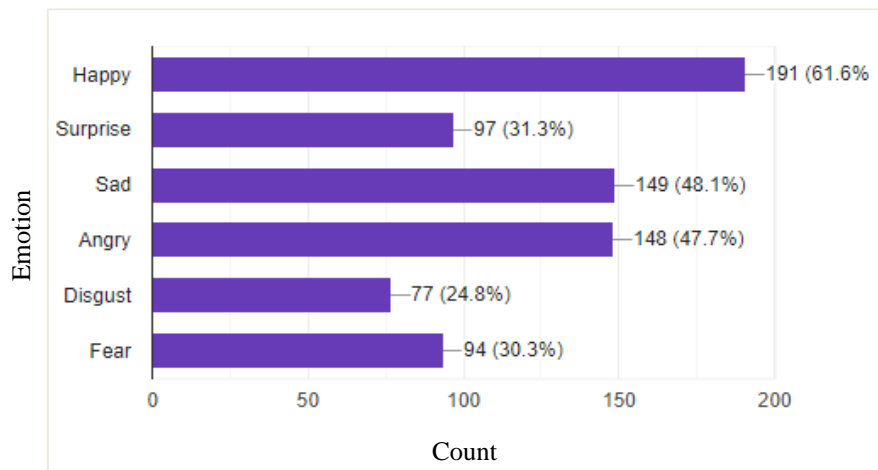


Figure A.2: Analysis of User Survey 1 - Question 4

Question 5 - In order to change your mood (emotion), what functionalities of a mobile device would you like to use?

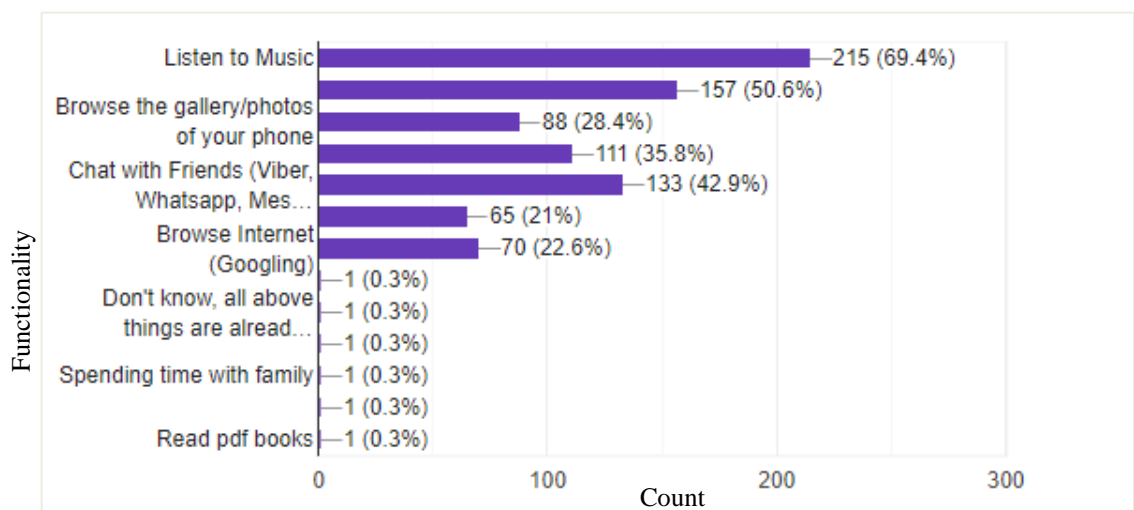


Figure A.3: Analysis of User Survey 1 - Question 5

**Question 6 - Do those functionalities have an impact based on the time of the day?
(Eg: In the morning you like to listen to music, but at night you like to chat with
your friends)**

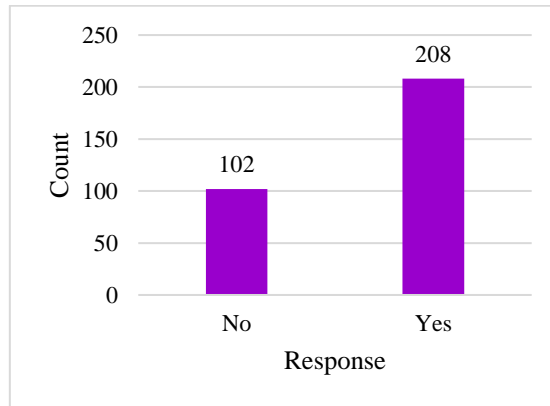


Figure A.4: Analysis of User Survey 1 - Question 6

**Question 7 - Do those functionalities have an impact on the location where you
are at that moment?**

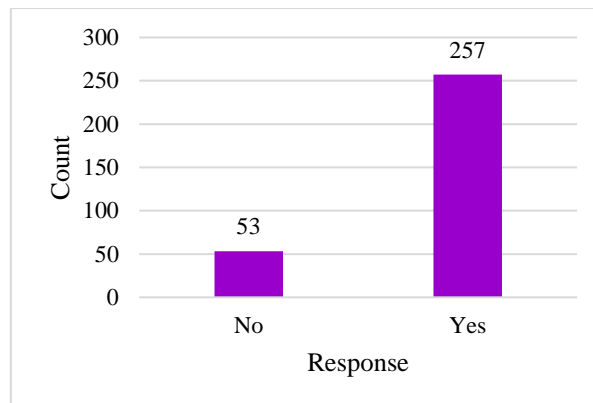


Figure A.5: Analysis of User Survey 1 - Question 7

Appendix B: User Survey 2

B.1 Questionnaire (Google Form)

Collecting Data to Improve User Experience based on User Emotions when using Mobile Devices

This survey is conducted with the aim of gathering information for improving the experience of users based on their emotions when using mobile devices.

When interacting with a mobile phone (browsing social media, reading news, texting a friend, checking emails, etc.) the user may feel negative emotions such as SAD or ANGRY. In order to change his/her mood, they may sometimes change the activity which they are currently doing to something else, so they can cheer up.

My intention is to capture what the users like to do in order to change their SAD or ANGRY mood while doing some activity using a mobile phone.

The survey is performed as part of the data gathering process for my Final Year Computer Science undergraduate research project.

Please make sure to use a laptop or a desktop computer to fill this questionnaire. Or else hold your mobile phone horizontally (in landscape mode) so all the options for the questions will be visible from section 2.

* Required

Age

- Below 18 years
- 18 to 24 years
- 25 to 35 years
- Above 35 years

Gender *

- Female
- Male

Occupation *

- Student
- Employed

Next

If Occupation=" Student"

Reactions in the early MORNING, when using your mobile phone?

Choose the option that you like to do, if something makes you SAD/ANGRY when using a mobile device in the MORNING.

Suppose you just woke up and you are using the mobile phone to check the latest news, emails, etc. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suppose you get ready to go to university now. You are walking to the bus halt while checking messages in your phone and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now while walking? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Now, you are waiting at the bus stop (standing) waiting for a bus. You started checking messages or emails and you become SAD/ANGRY. What would you like to do to change your mood now? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

You got into the bus but has a long ride to the university. Therefore, you started checking news feed, read blogs, etc. using your mobile phone. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Assume you are in the university and in a free time before lunchtime, you are checking your phone and you get SAD/ANGRY by reading a message. What would you prefer to do now? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Back](#)

[Next](#)

Reactions in the EVENING, when using your mobile phone?

Choose the option that you like to do if something makes you SAD/ANGRY when using a mobile device in the EVENING.

Suppose you have finished your lunch and you are using the mobile phone to check the news, gossips, etc. Then, you see something that makes you SAD/ANGRY. Since you are still at the university and the lectures will be starting soon, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After finishing university studies, you are walking to the bus stop to catch the bus. While walking you are checking messages and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now while walking? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

While waiting in the bus stop (standing), you started checking your messages until the bus comes. Then, you get SAD/ANGRY while reading a message. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After you get into the bus until you reach home, assume you are checking news feed, read blogs, etc. using your mobile phone. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood when you are in the bus? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suppose you are at home now and at leisure time you are either checking emails, watching tutorials, preparing for assignments, etc. using the mobile phone. Assume you become SAD/ANGRY while doing these things. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Back Submit

If Occupation=" Employed"

Reactions in the early MORNING, when using your mobile phone?

Choose the option that you like to do if something makes you SAD/ANGRY when using a mobile device in the MORNING.

Suppose you just woke up and you are using the mobile phone to check the latest news, emails, etc. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Suppose you are walking to the bus stop to go to work. while walking you are checking messages and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now while walking? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

While waiting at the bus stop (standing) for a bus, you are checking your messages and it makes you SAD/ANGRY. What would you like to do to change your mood now? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After you get into the bus until you reach your workplace, assume you are checking news feed, read blogs, etc.. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Assume you are at your workplace and it is before lunchtime. You check your phone and you get SAD/ANGRY by reading a message. What would you prefer to do now? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Back

Next

Reactions in the EVENING, when using your mobile phone?

Choose the option that you like to do if something makes you SAD/ANGRY when using a mobile device in the EVENING.

Suppose after having lunch at your workplace, you are using the mobile phone to check the news, gossips, etc. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After work, you are walking to the bus stop to catch the bus. While walking your checking messages and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Again you check your messages in your phone while waiting at the bus stop (standing) until the bus comes. Then, you get SAD/ANGRY while reading a message. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

After you get into the bus until you reach home, assume you are checking news feed, read blogs, etc. using your mobile phone. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Suppose you come home now and at leisure time you are either checking emails, reading news, doing work related stuff, etc. using your mobile phone. Assume you become SAD/ANGRY while doing these things. Now, what would you like to do to change your mood? *

	Listen to Music	Use Social Media (Facebook, Twitter, Instagram)	Chat with Friends (Whatsapp, Viber, Messenger)	Play a Game	Browse the Gallery/Photos in your phone	Browse Internet (Read news, other information searching etc.)
Sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

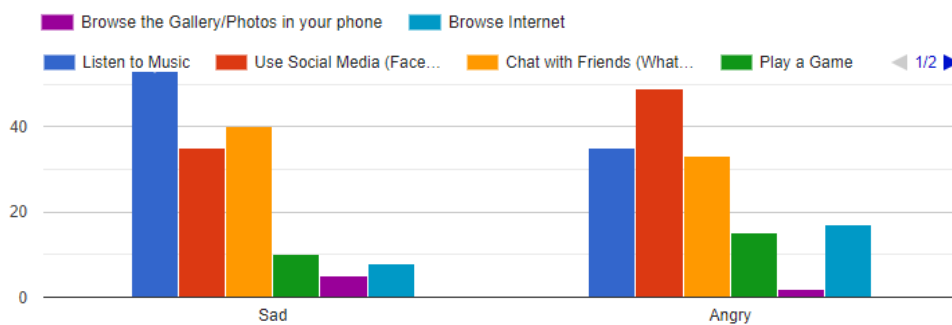
Back Submit

B.2 Results of User Survey 2

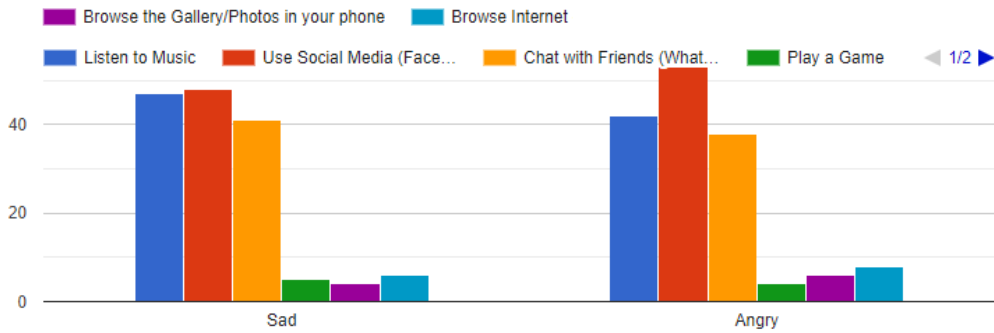
Student Category

Reactions in the early MORNING, when using your mobile phone?

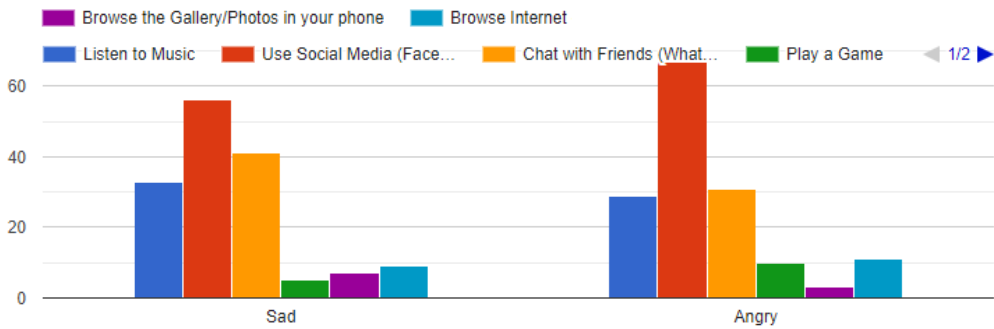
Suppose you just woke up and you are using the mobile phone to check the latest news, emails, etc. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood?



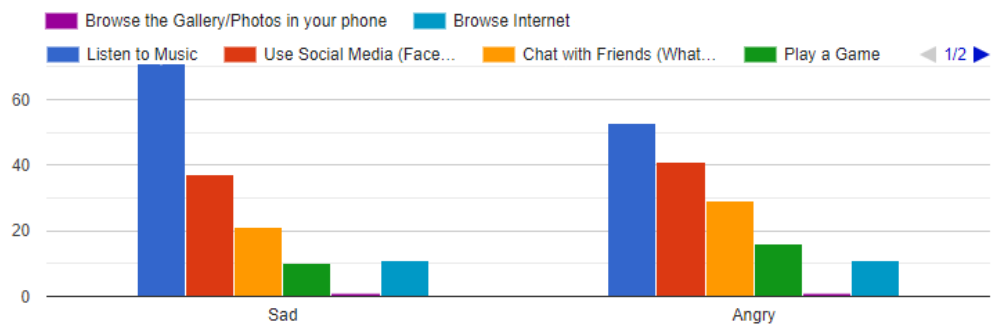
Suppose you get ready to go to university now. You are walking to the bus halt while checking messages in your phone and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now while walking?



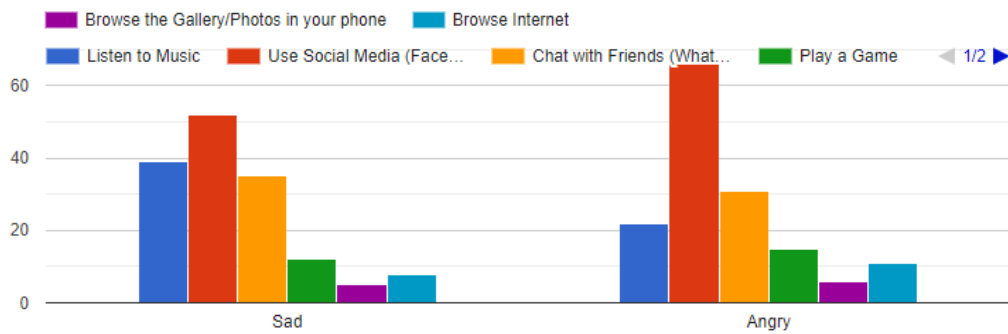
Now, you are waiting at the bus stop (standing) waiting for a bus. You started checking messages or emails and you become SAD/ANGRY. What would you like to do to change your mood now?



You got into the bus but has a long ride to the university. Therefore, you started checking news feed, read blogs, etc. using your mobile phone. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood?

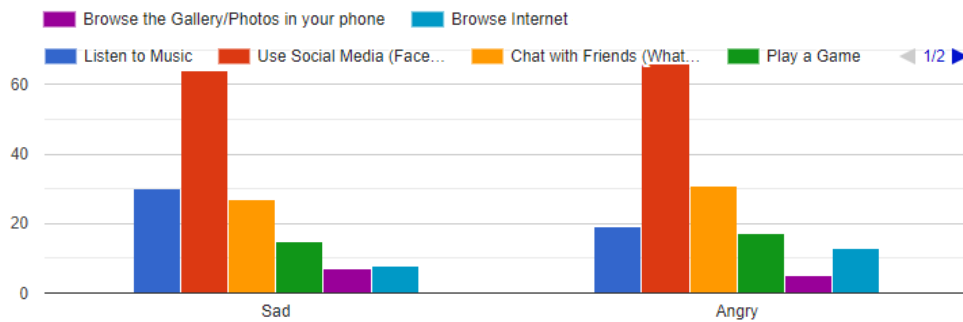


Assume you are in the university and in a free time before lunchtime, you are checking your phone and you get SAD/ANGRY by reading a message. What would you prefer to do now?

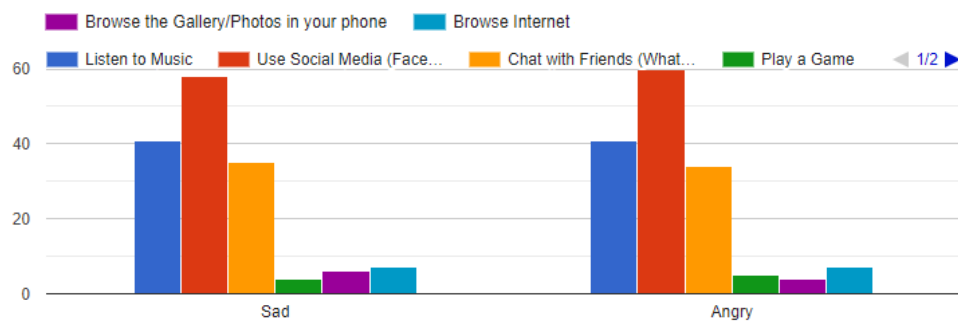


Reactions in the EVENING, when using your mobile phone?

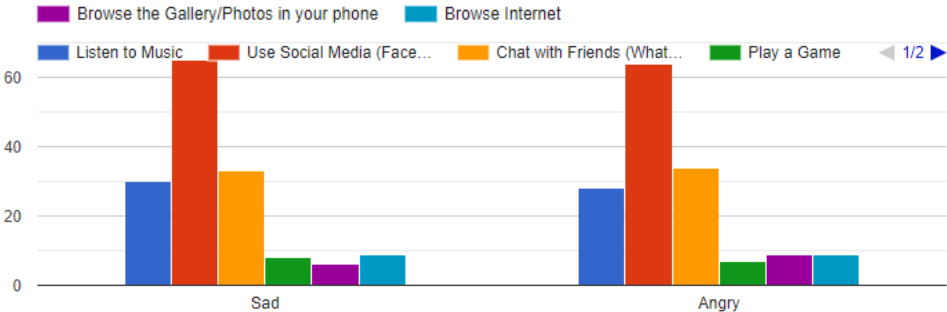
Suppose you have finished your lunch and you are using the mobile phone to check the news, gossips, etc. Then, you see something that makes you SAD/ANGRY. Since you are still at the university and the lectures will be starting soon, what would you like to do to change your mood?



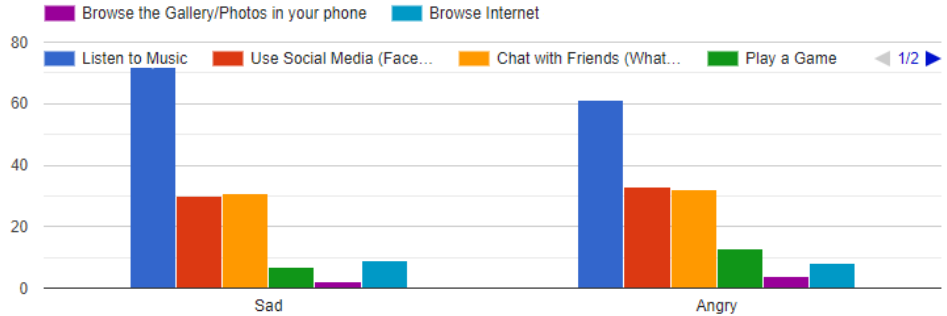
After finishing university studies, you are walking to the bus stop to catch the bus. while walking you are checking messages and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now while walking?



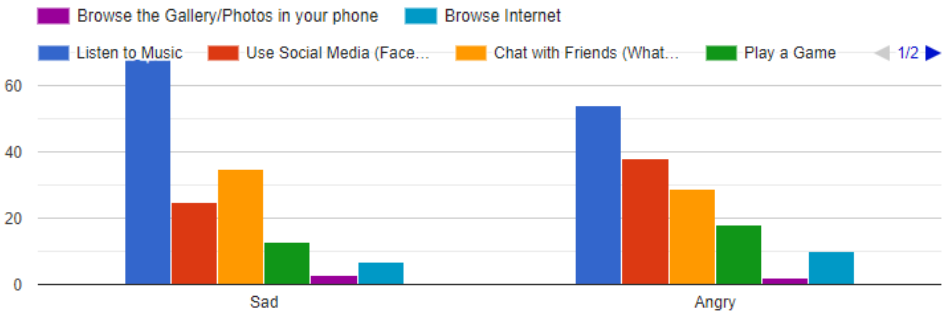
While waiting in the bus stop (standing), you started checking your messages until the bus comes. Then, you get SAD/ANGRY while reading a message. Now, what would you like to do to change your mood?



After you get into the bus until you reach home, assume you are checking news feed, read blogs, etc. using your mobile phone. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood when you are in the bus?



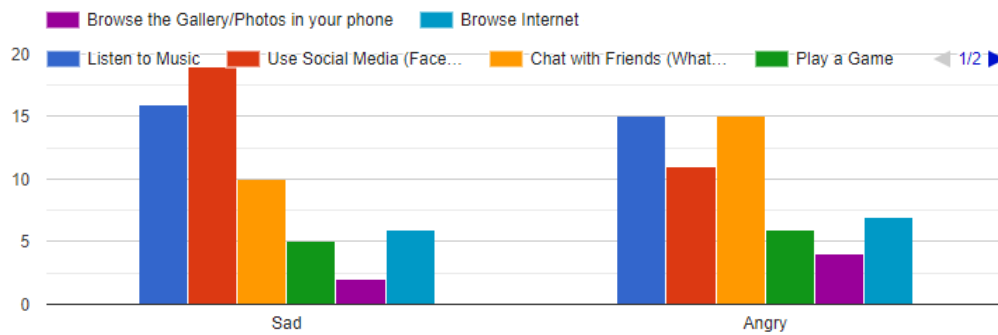
Suppose you are at home now and at leisure time you are either checking emails, watching tutorials, preparing for assignments, etc. using the mobile phone. Assume you become SAD/ANGRY while doing these things. Now, what would you like to do to change your mood?



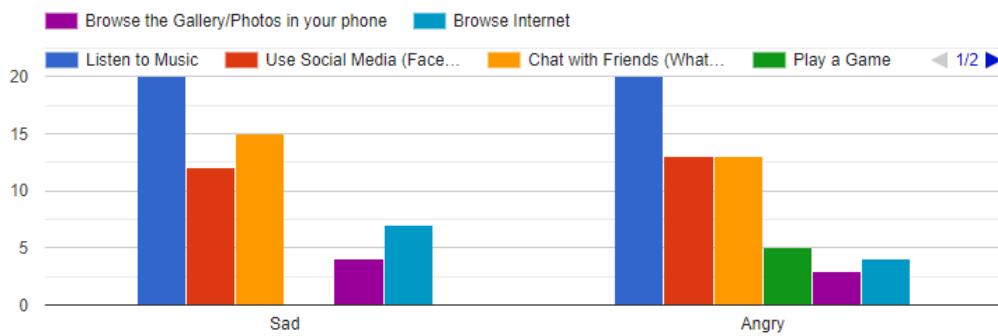
Employed Category

Reactions in the early MORNING, when using your mobile phone?

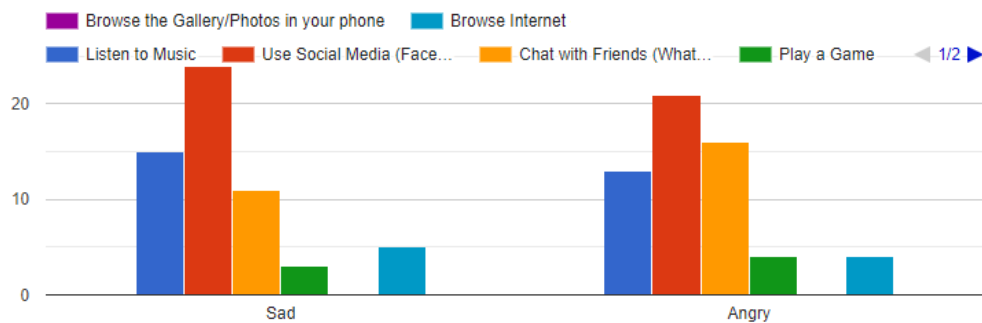
Suppose you just woke up and you are using the mobile phone to check the latest news, emails, etc. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood?



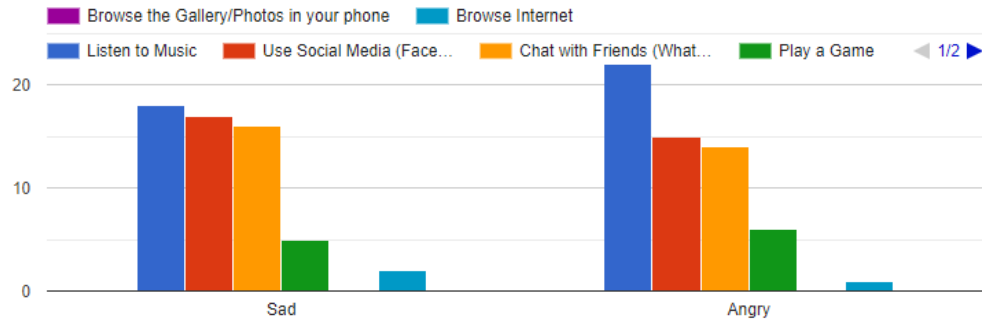
Suppose you are walking to the bus stop to go to work. while walking you are checking messages and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now while walking?



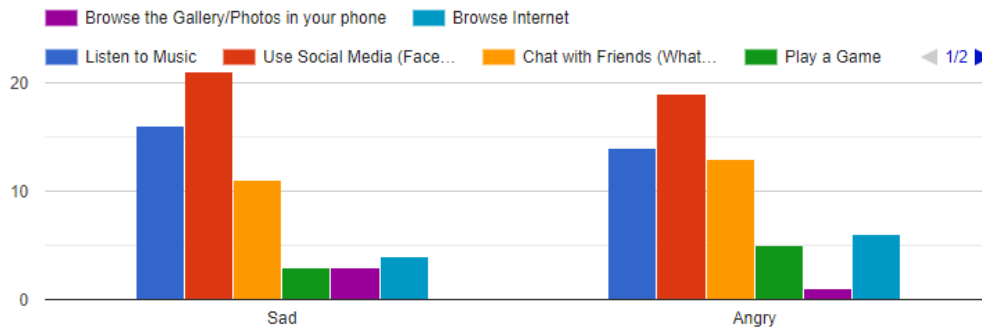
While waiting at the bus stop (standing) for a bus, you are checking your messages and it makes you SAD/ANGRY. What would you like to do to change your mood now?



After you get into the bus until you reach your workplace, assume you are checking news feed, read blogs, etc.. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood?

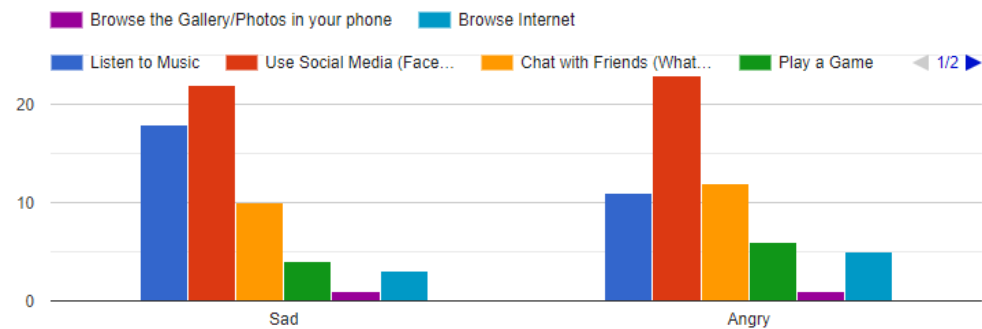


Assume you are at your workplace and it is before lunchtime. You check your phone and you get SAD/ANGRY by reading a message. What would you prefer to do now?

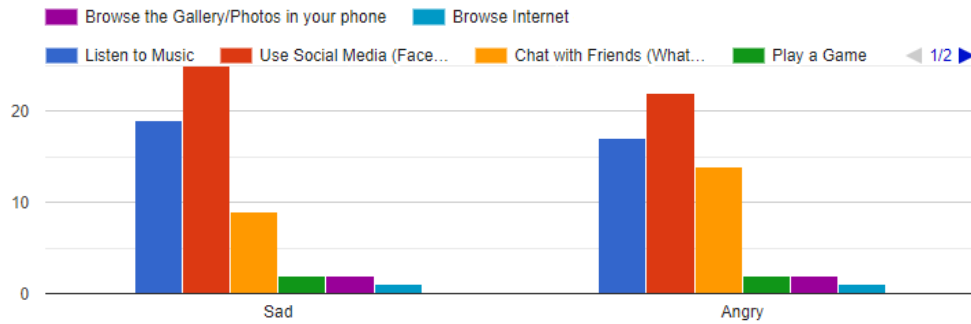


Reactions in the EVENING, when using your mobile phone?

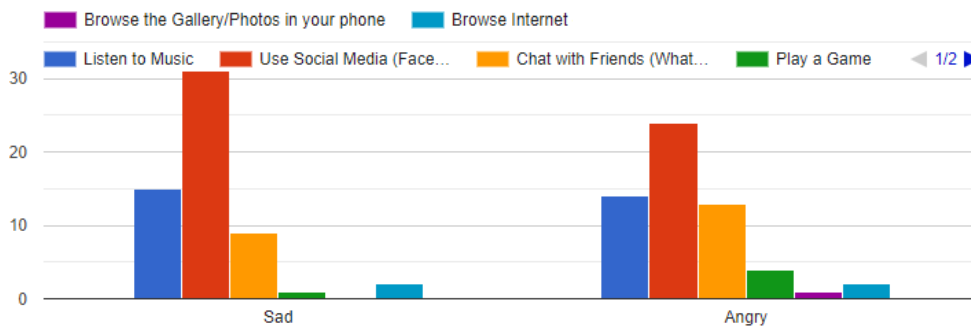
Suppose after having lunch at your workplace, you are using the mobile phone to check the news, gossips, etc. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood?



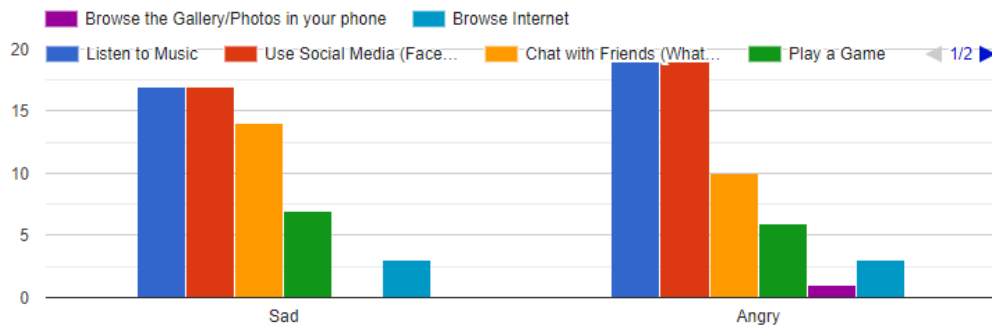
After work, you are walking to the bus stop to catch the bus. While walking your checking messages and suddenly you get SAD/ANGRY by reading a message. What would you prefer to do now?



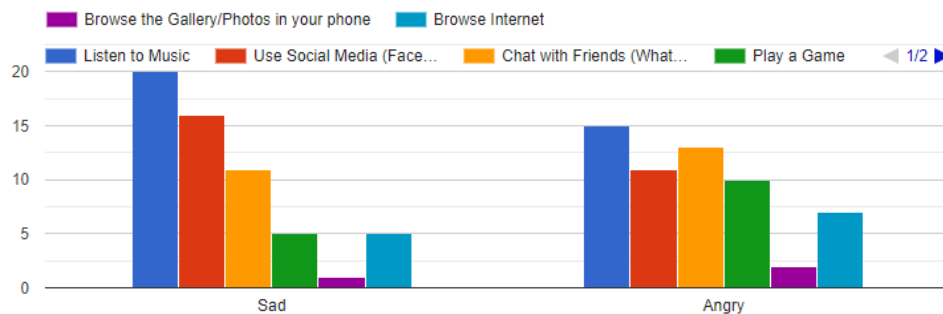
Again you check your messages in your phone while waiting at the bus stop (standing) until the bus comes. Then, you get SAD/ANGRY while reading a message. Now, what would you like to do to change your mood?



After you get into the bus until you reach home, assume you are checking news feed, read blogs, etc. using your mobile phone. Then, you see something that makes you SAD/ANGRY. Now, what would you like to do to change your mood?



Suppose you come home now and at leisure time you are either checking emails, reading news, doing work related stuff, etc. using your mobile phone. Assume you become SAD/ANGRY while doing these things. Now, what would you like to do to change your mood?



Appendix C: Informed Consent Form

CONSENT TO PARTICIPATE IN A USER STUDY

Improving User Experience with Multi-Dimensional Factors for Effective Interactions in a Dynamic Environment

It would be highly appreciated if you can participate in this research study as a support for my final year research in Computer Science at the University of Colombo School of Computing (UCSC). The goal of this research is to improve the experience of users when using mobile devices based on the emotions and the context in real-time. Your contribution is important and invaluable because you are familiar with smartphones and you have some proficiency in English.

Ask the research team questions. If you have questions later, the contact information for the research investigator in charge of the study is given below.

YOUR PRIVACY?

We will not store your confidential details and when we write about or share the results from the study, the combined and analyzed information will be used. After the completion of the research, we can give you an overall conclusion that we will derive based on the data which we will collect.

WHAT IF YOU HAVE QUESTIONS, SUGGESTIONS OR CONCERNS?

The person in charge of this study is **W. D. Watearachchi**, a final year undergraduate of the University of Colombo School of Computing (UCSC). If you have questions, suggestions, or concerns regarding this study or you want to withdraw from the study his contact information is:

- Email: wasuradananjith@gmail.com
- Contact No: +94775396038

If you have any questions, suggestions or concerns about your rights as a volunteer in this research, contact,

- Supervisor: **Prof. K.P. Hewagamage** (Director of University of Colombo School of Computing)
Email: kph@ucsc.cmb.ac.lk
- Co-Supervisor: **Dr. K.H.E.L.W. Hettiarachchi** (Senior lecturer at University of Colombo School of Computing)
Email: eno@ucsc.cmb.ac.lk

ARE YOU WILLING TO PARTICIPATE IN FUTURE STUDIES TOO?

Since this is a prototype evaluation, it would be much appreciated if you can join with us until the end of this research, so that we can do the future evaluations of the prototype in next stages with your contribution.

Are you willing to participate in future prototype evaluations?		
Yes	No	Signature _____

INFORMED CONSENT SIGNATURES

You will receive a copy of this consent form after it has been signed.

_____ Signature of participant		_____ Date
_____ Name of participant		
_____ Email address of participant		
_____ Signature of person obtaining informed consent		_____ Date
Mr. Wasura Dananjith Watterachchi _____ Printed name of person obtaining informed consent		

Appendix D: Code Listings

D.1 Code Listings of Prototype 1

Following is the XML code segment for QWERTY layout of the keyboard.

```
<?xml version="1.0" encoding="utf-8" ?>
<Keyboard xmlns:android="http://schemas.android.com/apk/res/android"
    android:keyWidth="10%p"
    android:horizontalGap="0px"
    android:verticalGap="0px"
    android:keyHeight="60dp">

    <Row>
        <Key android:keyLabel="1"
            android:keyEdgeFlags="left"
            android:codes="49"/>
        <Key android:keyLabel="2"
            android:codes="50"/>
        <Key android:keyLabel="3"
            android:codes="51"/>
        <Key android:keyLabel="4"
            android:codes="52"/>
        <Key android:keyLabel="5"
            android:codes="53"/>
        <Key android:keyLabel="6"
            android:codes="54"/>
        <Key android:keyLabel="7"
            android:codes="55"/>
        <Key android:keyLabel="8"
            android:codes="56"/>
        <Key android:keyLabel="9"
            android:codes="57"/>
        <Key android:keyLabel="0"
            android:keyEdgeFlags="right"
            android:codes="48"/>
    </Row>

    <Row>
        <Key android:keyLabel="q" android:keyEdgeFlags="left"
        android:codes="113"/>
        <Key android:keyLabel="w" android:codes="119"/>
        <Key android:keyLabel="e" android:codes="101"/>
        <Key android:keyLabel="r" android:codes="114"/>
        <Key android:keyLabel="t" android:codes="116"/>
        <Key android:keyLabel="y" android:codes="121"/>
        <Key android:keyLabel="u" android:codes="117"/>
        <Key android:keyLabel="i" android:codes="105"/>
        <Key android:keyLabel="o" android:codes="111"/>
        <Key android:keyLabel="p" android:keyEdgeFlags="right"
        android:codes="112"/>
    </Row>

    <Row>
        <Key android:keyLabel="a" android:keyEdgeFlags="left"
        android:codes="97"/>
    </Row>
</Keyboard>
```

```

        <Key android:keyLabel="s" android:codes="115"/>
        <Key android:keyLabel="d" android:codes="100"/>
        <Key android:keyLabel="f" android:codes="102"/>
        <Key android:keyLabel="g" android:codes="103"/>
        <Key android:keyLabel="h" android:codes="104"/>
        <Key android:keyLabel="j" android:codes="106"/>
        <Key android:keyLabel="k" android:codes="107"/>
        <Key android:keyLabel="l" android:codes="108"/>
        <Key android:keyLabel="\# \@" android:keyEdgeFlags="right"
android:codes="35,64"/>
    </Row>

    <Row>
        <Key android:keyLabel="CAPS" android:keyEdgeFlags="left"
android:codes="-1"/>
        <Key android:keyLabel="z" android:codes="122"/>
        <Key android:keyLabel="x" android:codes="120"/>
        <Key android:keyLabel="c" android:codes="99"/>
        <Key android:keyLabel="v" android:codes="118"/>
        <Key android:keyLabel="b" android:codes="98"/>
        <Key android:keyLabel="n" android:codes="110"/>
        <Key android:keyLabel="m" android:codes="109"/>
        <Key android:keyLabel="." android:codes="46"/>
        <Key android:keyLabel="\? ! :" android:keyEdgeFlags="right"
android:codes="63,33,58"/>
    </Row>

    <Row android:rowEdgeFlags="bottom">
        <Key android:keyLabel=",*" android:keyWidth="10%p"
android:keyEdgeFlags="left" android:codes="44,42"/>
        <Key android:keyLabel="/" android:keyWidth="10%p"
android:codes="47"/>
        <Key android:keyLabel="SPACE" android:keyWidth="50%p"
android:isRepeatable="true" android:codes="32"/>
        <Key android:keyLabel="DEL" android:keyWidth="20%p"
android:isRepeatable="true" android:codes="-5"/>
        <Key android:keyLabel="DONE" android:keyWidth="10%p"
android:keyEdgeFlags="right" android:codes="-4"/>

    </Row>
</Keyboard>

```

The Java code segment to retrieve the maximum emotions from facial expressions and text data by 60 second time intervals is shown below.

```

Runnable faceEmotionChecker = new Runnable() {
    @Override
    public void run() {
        try {
            // This will returns the maximum face emotion
            Double maxValueInFaceEmotions =
                (Collections.max(faceEmotions.values()));
            // Iterate through HashMap
            for (Map.Entry<String, Double> entry : faceEmotions.entrySet()) {
                if (entry.getValue().equals(maxValueInFaceEmotions)) {
                    // Get the index of the maximum face emotion
                    maxFaceEmotion = entry.getKey();
                }
            }
        }
    }
}

```

```

// This will returns the maximum text emotion
Double maxValueInTextEmotions =
    (Collections.max(textEmotions.values()));
// Iterate through HashMap
for (Map.Entry<String, Double> entry : textEmotions.entrySet()) {
    if (entry.getValue().equals(maxValueInTextEmotions)) {
        // Get the index of the maximum text emotion
        maxTextEmotion = entry.getKey();
    }
}

showNotification(maxFaceEmotion, maxTextEmotion, faceEmotions,
                textEmotions);

initializeFaceEmotions();
initializeTextEmotions();

} finally {
    // Run in 60 seconds intervals
    mHandler.postDelayed(faceEmotionChecker, 60000);
}
};

```

D.2 Code Listings of Prototype 2

Following is the Java code segment to decide emotions using the threshold values by giving priority to emotions from text.

```

Runnable faceEmotionChecker = new Runnable() {
    @Override
    public void run() {
        try {
            // This will returns the maximum face emotion
            Double maxValueInFaceEmotions =
                (Collections.max(faceEmotions.values()));
            // Iterate through HashMap
            if (maxValueInFaceEmotions >= 50) { // Check for the threshold
                for (Map.Entry<String, Double> entry :
                    faceEmotions.entrySet()) {
                    if (entry.getValue().equals(maxValueInFaceEmotions)) {
                        maxFaceEmotion = entry.getKey();
                    }
                }
            }

            // This will returns the maximum text emotion
            Double maxValueInTextEmotions =
                (Collections.max(textEmotions.values()));

            if (maxValueInTextEmotions == 0) {
                maxTextEmotion = "No Emotion";
            }
            else {
                // Iterate through HashMap
                if (maxValueInTextEmotions >= 50) { // Check for the threshold
                    for (Map.Entry<String, Double> entry :
                        textEmotions.entrySet()) {
                        if (entry.getValue().equals(maxValueInTextEmotions))

```

```

        {
            maxTextEmotion = entry.getKey();
        }
    }
}

showNotification(maxFaceEmotion, maxTextEmotion, faceEmotions,
                textEmotions);

initializeFaceEmotions();
initializeTextEmotions();

} finally {
    mHandler.postDelayed(faceEmotionChecker, 60000);
}
};

```

D.3 Code Listings of Prototype 3

Below is the Python code which was used to train the ANN.

```

# Import Libraries.
import pandas as pd
import numpy as np
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import OneHotEncoder
from sklearn.model_selection import StratifiedKFold
from keras.utils import plot_model
import matplotlib.pyplot as plt

# Import the dataset.
dataset = pd.read_csv('data.csv')

# creating input features and target variables.
X= dataset.iloc[:,0:13]
y= dataset.iloc[:,13]

# Normalize data.
X = np.divide(X,100)

# Convert the categorical string labels to categorical integer labels.
labelencoder = LabelEncoder()
y = labelencoder.fit_transform(y)

# Apply one hot encoding for above categorical labels.
one_hot_labels_y = tf.keras.utils.to_categorical(y, num_classes=9)

# Instantiate the model.
model = tf.keras.Sequential()
model.add(layers.Dense(30, activation=tf.nn.relu, input_dim=13))
model.add(layers.Dropout(rate=0.5))

```



```

model.add(layers.Dense(9, activation=tf.nn.softmax))

# Configure the model for training.
model.compile(optimizer=tf.train.AdamOptimizer(0.001),
              loss='categorical_crossentropy',
              metrics=['accuracy'])

# Train the model for a given number of epochs.
history = model.fit(X, one_hot_labels_y, validation_split=0.3, epochs=200, batch_size=10)

# Plot training & validation accuracy values
plt.plot(history.history['acc'])
plt.plot(history.history['val_acc'])
plt.title('Model accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Train', 'Test'], loc='upper left')
plt.show()

# Plot training & validation loss values
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('Model loss')
plt.ylabel('Loss')
plt.xlabel('Epoch')
plt.legend(['Train', 'Test'], loc='upper left')
plt.show()

```

Below is the Java code which was used to incorporate the above trained ANN into the Android app and to pass inputs in order to get the most meaningful emotion as the output.

```

public static String[] outputEmotion = {
    "analytical",
    "anger",
    "confident",
    "disgust",
    "fear",
    "joy",
    "sadness",
    "surprise",
    "tentative"};

try {

    Log.d("ML", "faceEmotions " + faceEmotions.toString());
    Log.d("ML", "textEmotions " + textEmotions.toString());

    try {
        // loading the TensorFlow Lite model
        tfliteModel = loadModelFile();
        tflite = new Interpreter(tfliteModel);

        // inputs to the ANN
        float [][] input = new float[][] {{
            faceEmotions.get("anger"),
            faceEmotions.get("disgust"),
            faceEmotions.get("fear"),

```

```

        faceEmotions.get("joy"),
        faceEmotions.get("sadness"),
        faceEmotions.get("surprise"),
        textEmotions.get("analytical"),
        textEmotions.get("anger"),
        textEmotions.get("confident"),
        textEmotions.get("fear"),
        textEmotions.get("joy"),
        textEmotions.get("sadness"),
        textEmotions.get("tentative")});

// outputs to be retrieved
float [][] output = new float[][] {{0,0,0,0,0,0,0,0,0,0}};
tflite.run(input, output);

float max = output[0][0];
int index = 0;

for (int i = 0; i < output[0].length; i++)
{
    if (max < output[0][i])
    {
        max = output[0][i];
        index = i;
    }
}

// keep tracking of maximum emotion from face and text for evaluation
// purposes
Float maxValueInFaceEmotions =
    (Collections.max(faceEmotions.values()));
Float maxValueInTextEmotions =
    (Collections.max(textEmotions.values()));

if ((maxValueInFaceEmotions != 0) || (maxValueInTextEmotions != 0)){
    Log.d("ML", "Max emotion " + outputEmotion[index]);
    showNotification(outputEmotion[index], faceEmotions,
        textEmotions);

    initializeFaceEmotions();
    initializeTextEmotions();
}
} catch (IOException e) {
    Log.d("FaceDetector", "no");
    e.printStackTrace();
}

} finally {
    mHandler.postDelayed(faceEmotionChecker, 60000);
}

```

D.4 Code Listings of Prototype 4

The probabilities from User Survey 2 was stored as the **Preference Tree** in JSON format for each user in his/her account after the registration as follows.

```

{
  "anger" : {
    "evening" : {
      "home" : {
        "still" : {

```

```

        "browseInternet" : 9.35,
        "browseTheGallery" : 2.35,
        "chatWithFriends" : 20.8,
        "listenToMusic" : 30.85,
        "playAGame" : 14.55,
        "useSocialMedia" : 22.1
    },
    "vehicle" : {
        "browseInternet" : 9.35,
        "browseTheGallery" : 2.35,
        "chatWithFriends" : 20.8,
        "listenToMusic" : 30.85,
        "playAGame" : 14.55,
        "useSocialMedia" : 22.1
    },
    "walking" : {
        "browseInternet" : 9.35,
        "browseTheGallery" : 2.35,
        "chatWithFriends" : 20.8,
        "listenToMusic" : 30.85,
        "playAGame" : 14.55,
        "useSocialMedia" : 22.1
    }
},
"outside" : {
    "still" : {
        "browseInternet" : 4.7,
        "browseTheGallery" : 3.9,
        "chatWithFriends" : 22.45,
        "listenToMusic" : 21.3,
        "playAGame" : 5.75,
        "useSocialMedia" : 41.9
    },
    "vehicle" : {
        "browseInternet" : 5.25,
        "browseTheGallery" : 2.15,
        "chatWithFriends" : 19.2,
        "listenToMusic" : 36.6,
        "playAGame" : 9.45,
        "useSocialMedia" : 27.35
    },
    "walking" : {
        "browseInternet" : 3.2,
        "browseTheGallery" : 3,
        "chatWithFriends" : 23.3,
        "listenToMusic" : 28.3,
        "playAGame" : 3.4,
        "useSocialMedia" : 38.8
    }
},
"work" : {
    "still" : {
        "browseInternet" : 8.6,
        "browseTheGallery" : 2.5,
        "chatWithFriends" : 20.6,
        "listenToMusic" : 15.8,
        "playAGame" : 10.8,
        "useSocialMedia" : 41.7
    },
    "vehicle" : {
        "browseInternet" : 8.6,
        "browseTheGallery" : 2.5,
        "chatWithFriends" : 20.6,
        "listenToMusic" : 15.8,
        "playAGame" : 10.8,
        "useSocialMedia" : 41.7
    },
    "walking" : {

```

```

        "browseInternet" : 8.6,
        "browseTheGallery" : 2.5,
        "chatWithFriends" : 20.6,
        "listenToMusic" : 15.8,
        "playAGame" : 10.8,
        "useSocialMedia" : 41.7
    }
}
},
"morning" : {
    "home" : {
        "still" : {
            "browseInternet" : 11.7,
            "browseTheGallery" : 4.1,
            "chatWithFriends" : 23.9,
            "listenToMusic" : 24.5,
            "playAGame" : 10.1,
            "useSocialMedia" : 25.7
        },
        "vehicle" : {
            "browseInternet" : 11.7,
            "browseTheGallery" : 4.1,
            "chatWithFriends" : 23.9,
            "listenToMusic" : 24.5,
            "playAGame" : 10.1,
            "useSocialMedia" : 25.7
        },
        "walking" : {
            "browseInternet" : 11.7,
            "browseTheGallery" : 4.1,
            "chatWithFriends" : 23.9,
            "listenToMusic" : 24.5,
            "playAGame" : 10.1,
            "useSocialMedia" : 25.7
        }
    },
    "outside" : {
        "still" : {
            "browseInternet" : 7.1,
            "browseTheGallery" : 1,
            "chatWithFriends" : 24.05,
            "listenToMusic" : 20.8,
            "playAGame" : 6.75,
            "useSocialMedia" : 40.3
        },
        "vehicle" : {
            "browseInternet" : 7.1,
            "browseTheGallery" : 1,
            "chatWithFriends" : 24.05,
            "listenToMusic" : 20.8,
            "playAGame" : 6.75,
            "useSocialMedia" : 40.3
        },
        "walking" : {
            "browseInternet" : 6.1,
            "browseTheGallery" : 4.6,
            "chatWithFriends" : 23.8,
            "listenToMusic" : 31.15,
            "playAGame" : 5.6,
            "useSocialMedia" : 28.75
        }
    },
    "work" : {
        "still" : {
            "browseInternet" : 8.8,
            "browseTheGallery" : 2.9,
            "chatWithFriends" : 21.45,
            "listenToMusic" : 19.35,

```

```

        "playAGame" : 9.25,
        "useSocialMedia" : 38.25
    },
    "vehicle" : {
        "browseInternet" : 8.8,
        "browseTheGallery" : 2.9,
        "chatWithFriends" : 21.45,
        "listenToMusic" : 19.35,
        "playAGame" : 9.25,
        "useSocialMedia" : 38.25
    },
    "walking" : {
        "browseInternet" : 8.8,
        "browseTheGallery" : 2.9,
        "chatWithFriends" : 21.45,
        "listenToMusic" : 19.35,
        "playAGame" : 9.25,
        "useSocialMedia" : 38.25
    }
}
}
},
"sadness" : {
    "evening" : {
        "home" : {
            "still" : {
                "browseInternet" : 6.6,
                "browseTheGallery" : 1.85,
                "chatWithFriends" : 21.1,
                "listenToMusic" : 39.75,
                "playAGame" : 8.6,
                "useSocialMedia" : 22.1
            },
            "vehicle" : {
                "browseInternet" : 6.6,
                "browseTheGallery" : 1.85,
                "chatWithFriends" : 21.1,
                "listenToMusic" : 39.75,
                "playAGame" : 8.6,
                "useSocialMedia" : 22.1
            },
            "walking" : {
                "browseInternet" : 6.6,
                "browseTheGallery" : 1.85,
                "chatWithFriends" : 21.1,
                "listenToMusic" : 39.75,
                "playAGame" : 8.6,
                "useSocialMedia" : 22.1
            }
        },
        "outside" : {
            "still" : {
                "browseInternet" : 4.7,
                "browseTheGallery" : 2,
                "chatWithFriends" : 18.7,
                "listenToMusic" : 22.9,
                "playAGame" : 3.5,
                "useSocialMedia" : 48.2
            },
            "vehicle" : {
                "browseInternet" : 5.6,
                "browseTheGallery" : 0.65,
                "chatWithFriends" : 22.3,
                "listenToMusic" : 38.5,
                "playAGame" : 8.35,
                "useSocialMedia" : 24.6
            },
            "walking" : {

```

```

        "browseInternet" : 3.2,
        "browseTheGallery" : 3.7,
        "chatWithFriends" : 19.35,
        "listenToMusic" : 30,
        "playAGame" : 3,
        "useSocialMedia" : 40.75
    }
},
"work" : {
    "still" : {
        "browseInternet" : 5.25,
        "browseTheGallery" : 3.2,
        "chatWithFriends" : 17.55,
        "listenToMusic" : 25.45,
        "playAGame" : 8.4,
        "useSocialMedia" : 40.15
    },
    "vehicle" : {
        "browseInternet" : 5.25,
        "browseTheGallery" : 3.2,
        "chatWithFriends" : 17.55,
        "listenToMusic" : 25.45,
        "playAGame" : 8.4,
        "useSocialMedia" : 40.15
    },
    "walking" : {
        "browseInternet" : 5.25,
        "browseTheGallery" : 3.2,
        "chatWithFriends" : 17.55,
        "listenToMusic" : 25.45,
        "playAGame" : 8.4,
        "useSocialMedia" : 40.15
    }
}
},
"morning" : {
    "home" : {
        "still" : {
            "browseInternet" : 7.8,
            "browseTheGallery" : 3.4,
            "chatWithFriends" : 21.85,
            "listenToMusic" : 31.35,
            "playAGame" : 7.6,
            "useSocialMedia" : 28
        },
        "vehicle" : {
            "browseInternet" : 7.8,
            "browseTheGallery" : 3.4,
            "chatWithFriends" : 21.85,
            "listenToMusic" : 31.35,
            "playAGame" : 7.6,
            "useSocialMedia" : 28
        },
        "walking" : {
            "browseInternet" : 7.8,
            "browseTheGallery" : 3.4,
            "chatWithFriends" : 21.85,
            "listenToMusic" : 31.35,
            "playAGame" : 7.6,
            "useSocialMedia" : 28
        }
    }
},
"outside" : {
    "still" : {
        "browseInternet" : 7.3,
        "browseTheGallery" : 2.3,
        "chatWithFriends" : 23.1,
        "listenToMusic" : 23.9,

```



```

currentLocation.setLongitude(currentLongitude);

Location homeLocation = new Location("");
homeLocation.setLatitude(homeLatitude);
homeLocation.setLongitude(homeLongitude);

Location workLocation = new Location("");
workLocation.setLatitude(workLatitude);
workLocation.setLongitude(workLongitude);

float distanceBetweenHomeAndCurrentLocation =
    currentLocation.distanceTo(homeLocation);
float distanceBetweenWorkAndCurrentLocation =
    currentLocation.distanceTo(workLocation);

if (distanceBetweenHomeAndCurrentLocation <
    distanceBetweenWorkAndCurrentLocation){
    // If current location is close to home check a radius of 20m
    if (distanceBetweenHomeAndCurrentLocation <= 20){
        location = "home";
    }
    else{
        location = "outside";
    }
}
else{
    // If current location is close to workplace check a radius of 20m
    if (distanceBetweenWorkAndCurrentLocation <= 20){
        location = "work";
    }
    else{
        location = "outside";
    }
}
}
}

```

The Java implementation for detecting user activities are shown below.

```

private void handleDetectedActivities (List<DetectedActivity>
probableActivities) {
    StringBuilder str = new StringBuilder();

    for(DetectedActivity activity : probableActivities) {
        switch( activity.getType() ) {
            case DetectedActivity.IN_VEHICLE: {
                userActivities.put("vehicle", activity.getConfidence());
                break;
            }
            case DetectedActivity.ON_BICYCLE: {
                userActivities.put("vehicle", activity.getConfidence());
                break;
            }
            case DetectedActivity.ON_FOOT: {
                userActivities.put("walking", activity.getConfidence());
                break;
            }
            case DetectedActivity.RUNNING: {
                userActivities.put("walking", activity.getConfidence());
                break;
            }
            case DetectedActivity.STILL: {
                userActivities.put("still", activity.getConfidence());
                break;
            }
            case DetectedActivity.TILTING: {

```



```

        userActivities.put("still", activity.getConfidence());
        break;
    }
    case DetectedActivity.WALKING: {
        userActivities.put("walking", activity.getConfidence());
        break;
    }
}
}
Log.d("UserActivity", userActivities.toString());
Intent intent = new Intent(LOCAL_BROADCAST_NAME);
intent.putExtra(LOCAL_BROADCAST_EXTRA, str.toString());
LocalBroadcastManager.getInstance(this).sendBroadcast(intent);
}

// This will returns the user activity with the highest confidence
Integer maxConfidenceInUserActivity =
    (Collections.max(ActivityRecognizedService.userActivities.values()));

Map.Entry<String, Integer> maxEntry = null;

for (Map.Entry<String, Integer> entry :
    ActivityRecognizedService.userActivities.entrySet()) {

    if (maxEntry == null
        || entry.getValue().compareTo(maxEntry.getValue()) > 0) {
        maxEntry = entry;
    }
}

maxUserActivity = maxEntry.getKey();
Toast.makeText(getApplicationContext(),
    maxUserActivity + " " +
        ActivityRecognizedService.maxUserActivityConfidence
    , Toast.LENGTH_LONG).show();

```

The Java implementation for detecting maximum adaptive function are shown below.

```

private void findMaxAdaptiveFunction(final String maxEmotion) {
    SharedPreferences pref = getSharedPreferences("loginData", MODE_PRIVATE);
    String name = pref.getString("name", null);

    FirebaseDatabase.getInstance().getReference("users/" + name
        + "/adaptiveFunctions");
    FirebaseDatabase.getInstance().addListenerForSingleValueEvent(new ValueEventListener() {

@Override
public void onDataChange(@NonNull DataSnapshot dataSnapshot) {

    // Retrieve the probabilities of adaptive functions for the
    // particular user
    AdaptiveFunction adaptiveFunction =
        dataSnapshot.child(maxEmotion)
            .child(timeString).child(location)
            .child(maxUserActivity).getValue(AdaptiveFunction.class);
    adaptiveFunctions = new HashMap<>();
    adaptiveFunctions.put("listen to
        music", adaptiveFunction.getListenToMusic());
    adaptiveFunctions.put("use social
        media", adaptiveFunction.getUseSocialMedia());
    adaptiveFunctions.put("chat with

```

```

        friends", adaptiveFunction.getChatWithFriends());
adaptiveFunctions.put("play a
        game", adaptiveFunction.getPlayAGame());
adaptiveFunctions.put("browse the
        gallery", adaptiveFunction.getBrowseTheGallery());
adaptiveFunctions.put("browse
        internet", adaptiveFunction.getBrowseInternet());

// This will returns the maximum adaptive function
Double maxValueInAdaptiveFunctions =
    (Collections.max(adaptiveFunctions.values()));
// Iterate through HashMap
for (Map.Entry<String, Double> entry :
        adaptiveFunctions.entrySet()) {
    if (entry.getValue().equals(maxValueInAdaptiveFunctions)) {
        maxAdaptiveFunction = entry.getKey();
    }
}

switch(maxAdaptiveFunction) {
    case "listen to music":
        adaptiveDecision = "Would you like to listen to music?";
        break;
    case "use social media":
        adaptiveDecision = "Would you like to use
                            facebook/twitter?";
        break;
    case "chat with friends":
        adaptiveDecision = "Would you like to chat with
                            friends?";
        break;
    case "play a game":
        adaptiveDecision = "Would you like to play a game?";
        break;
    case "browse the gallery":
        adaptiveDecision = "Would you like to browse the
                            gallery?";
        break;
    case "browse internet":
        adaptiveDecision = "Would you like to browse internet?";
        break;
    default:
        adaptiveDecision = "Would you like to close the app?";
        break;
}
}

@Override
public void onCancelled(@NonNull DatabaseError databaseError) {
}
});
}
}

```

Following are some example display screens suggesting functions.



Figure D.1: Suggesting apps to browse internet

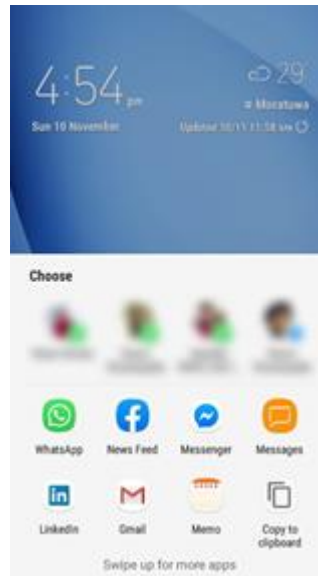


Figure D.2: Suggesting apps to chat with friends



Figure D.3: Suggesting apps to listen to music

D.5 Code Listings of Prototype 5

Following is the Java code segment which increases the count of the particular chosen adaptive function in “frequentFunctions” field in the database.

```

FrequentFunction frequentFunction = new FrequentFunction();
frequentFunction.increaseCount(chosenFunction, name);

// Increase the count of currently used function
public void increaseCount(final String passedFunction, String name){
    FirebaseDatabase mDatabase = FirebaseDatabase.getInstance().getReference("users/" + name +
                                                                              "/" + databaseRef);
    mDatabase.addListenerForSingleValueEvent(new ValueEventListener() {
        @Override
        public void onDataChange(@NonNull DataSnapshot dataSnapshot) {
            FrequentFunction frequentFunction =
                dataSnapshot.getValue(FrequentFunction.class);

            int count = 0;
            switch(passedFunction) {
                case "listen to music":
                    count = frequentFunction.getListenToMusic()+1;
                    frequentFunction.setListenToMusic(count);
                    break;
                case "use social media":
                    count = frequentFunction.getUseSocialMedia()+1;
                    frequentFunction.setUseSocialMedia(count);
                    break;
                case "chat with friends":
                    count = frequentFunction.getChatWithFriends()+1;

```

```

        frequentFunction.setChatWithFriends(count);
        break;
    case "play a game":
        count = frequentFunction.getPlayAGame()+1;
        frequentFunction.setPlayAGame(count);
        break;
    case "browse the gallery":
        count = frequentFunction.getBrowseTheGallery()+1;
        frequentFunction.setBrowseTheGallery(count);
        break;
    case "browse internet":
        count = frequentFunction.getBrowseInternet()+1;
        frequentFunction.setBrowseInternet(count);
        break;
    default:
        break;
}
mDatabase.setValue(frequentFunction);
});
}

```

Below is the Java code segment to calculate the Frequent Probabilities and to find the maximum adaptive functions to be suggested.

```

// Finding the maximum adaptive function based on the frequent probabilities
// and the probabilities from the Preference Tree (adaptiveFunctions)
private void findMaxAdaptiveFunction(final String maxEmotion) {
    SharedPreferences pref = getSharedPreferences("loginData", MODE_PRIVATE);
    String name = pref.getString("name", null);

    // Retrieve data from frequentFunctions field
    mDatabase1 = FirebaseDatabase.getInstance().getReference("users/" + name
        + "/frequentFunctions");
    mDatabase1.addListenerForSingleValueEvent(new ValueEventListener() {
        @Override
        public void onDataChange(@NonNull DataSnapshot dataSnapshot) {
            // Retrieve the counts of frequently used functions for the
            // particular user
            frequentFunction = dataSnapshot.getValue(FrequentFunction.class);

            SharedPreferences pref = getSharedPreferences("loginData",
                MODE_PRIVATE);
            String name = pref.getString("name", null);

            mDatabase2 = FirebaseDatabase.getInstance().getReference("users/"
                + name + "/adaptiveFunctions");
            mDatabase2.addListenerForSingleValueEvent(new
                ValueEventListener() {
                    @Override
                    public void onDataChange(@NonNull DataSnapshot dataSnapshot)
                    {
                        int totalFrequentFunctionCount =
                            frequentFunction.getBrowseInternet() +
                            frequentFunction.getBrowseTheGallery() +
                            frequentFunction.getChatWithFriends() +
                            frequentFunction.getListenToMusic() +
                            frequentFunction.getPlayAGame() +
                            frequentFunction.getUseSocialMedia();

                        // Retrieve the probabilities of adaptive functions for
                        // the particular user from the Preference Tree -
                        // adaptiveFunctions
                        AdaptiveFunction adaptiveFunction =

```

```

        dataSnapshot.child(maxEmotion)
            .child(timeString).child(location)
            .child(maxUserActivity)
            .getValue(AdaptiveFunction.class);

// This is only to store the probabilities from the
// Preference Tree
adaptiveFunctions = new HashMap<>();
adaptiveFunctions.put("listen to
    music",adaptiveFunction.getListenToMusic());
adaptiveFunctions.put("use social
    media",adaptiveFunction.getUseSocialMedia());
adaptiveFunctions.put("chat with
    friends",adaptiveFunction.getChatWithFriends());
adaptiveFunctions.put("play a
    game",adaptiveFunction.getPlayAGame());
adaptiveFunctions.put("browse the
    gallery",adaptiveFunction.getBrowseTheGallery());
adaptiveFunctions.put("browse
    internet",adaptiveFunction.getBrowseInternet());

// This is to store the frequent probabilities
frequentAdaptiveFunctions = new HashMap<>();

// If the user is using the app for the first time
if (totalFrequentFunctionCount == 0){
    frequentAdaptiveFunctions.put("listen to music",
        adaptiveFunction.getListenToMusic());
    frequentAdaptiveFunctions.put("use social media",
        adaptiveFunction.getUseSocialMedia());
    frequentAdaptiveFunctions.put("chat with friends",
        adaptiveFunction.getChatWithFriends());
    frequentAdaptiveFunctions.put("play a game",
        adaptiveFunction.getPlayAGame());
    frequentAdaptiveFunctions.put("browse the gallery",
        adaptiveFunction.getBrowseTheGallery());
    frequentAdaptiveFunctions.put("browse internet",
        adaptiveFunction.getBrowseInternet());
}
else{
    double listenToMusicProbability =
        (((double) frequentFunction.getListenToMusic() /
            totalFrequentFunctionCount) *100+
            adaptiveFunction.getListenToMusic()) /2;
    frequentAdaptiveFunctions.put("listen to music",
        listenToMusicProbability);

    double useSocialMediaProbability =
        (((double) frequentFunction.getUseSocialMedia() /
            totalFrequentFunctionCount) *100+
            adaptiveFunction.getUseSocialMedia()) /2;
    frequentAdaptiveFunctions.put("use social media",
        useSocialMediaProbability);

    double chatWithFriendsProbability =
        (((double) frequentFunction.getChatWithFriends() /
            totalFrequentFunctionCount) *100+
            adaptiveFunction.getChatWithFriends()) /2;
    frequentAdaptiveFunctions.put("chat with friends",
        chatWithFriendsProbability);

    double playAGameProbability =
        (((double) frequentFunction.getPlayAGame() /
            totalFrequentFunctionCount) *100+
            adaptiveFunction.getPlayAGame()) /2;
    frequentAdaptiveFunctions.put("play a game",
        playAGameProbability);
}
}

```

```

    double browseTheGalleryProbability =
        (((double) frequentFunction.getBrowseTheGallery() /
          totalFrequentFunctionCount) * 100 +
          adaptiveFunction.getBrowseTheGallery()) / 2;
    frequentAdaptiveFunctions.put("browse the gallery",
        browseTheGalleryProbability);

    double browseInternetProbability =
        (((double) frequentFunction.getBrowseInternet() /
          totalFrequentFunctionCount) * 100 +
          adaptiveFunction.getBrowseInternet()) / 2;
    frequentAdaptiveFunctions.put("browse internet",
        browseInternetProbability);
}

Log.d("FrequentFunctions",
    frequentAdaptiveFunctions.toString());

// This will returns the maximum adaptive function
Double maxValueInAdaptiveFunctions =
    (Collections.max(frequentAdaptiveFunctions.values()));
// Iterate through HashMap
for (Map.Entry<String, Double> entry :
    frequentAdaptiveFunctions.entrySet()) {
    if (entry.getValue()
        .equals(maxValueInAdaptiveFunctions)) {
        maxAdaptiveFunction = entry.getKey();
    }
}

switch (maxAdaptiveFunction) {
    case "listen to music":
        adaptiveDecision = "Would you like to listen to
            music?";
        break;
    case "use social media":
        adaptiveDecision = "Would you like to use
            facebook/twitter?";
        break;
    case "chat with friends":
        adaptiveDecision = "Would you like to chat with
            friends?";
        break;
    case "play a game":
        adaptiveDecision = "Would you like to play a
            game?";
        break;
    case "browse the gallery":
        adaptiveDecision = "Would you like to browse the
            gallery?";
        break;
    case "browse internet":
        adaptiveDecision = "Would you like to browse
            internet?";
        break;
    default:
        adaptiveDecision = "Would you like to close the
            app?";
        break;
}
Log.d("AdaptiveDecision", adaptiveDecision);
});
}

```

Appendix E: Deciding the Thresholds

E.1 Deciding Threshold 1 (T1)

Table E.1 shows the values that have been used to decide T1 for Prototype 2. When considering one row, the “Chosen Emotion” field consists of the emotion that the user has agreed to choose. The value which is underlined is the maximum value emotion and if the column name of the maximum value and “Chosen Emotion” matches, it is correctly predicted.

Table E.1: Percentage Values When Detecting Emotions from Facial Expressions

Chosen Emotion	Anger	Disgust	Fear	Joy	Sadness	Surprise
Sadness	0.91254741	0.428543389	0.484867573	3.69E-06	<u>99.2312851</u>	0.004062998
Joy	1.84E-05	0.011084231	9.31E-06	<u>94.37898254</u>	4.85E-06	4.268239498
Joy	2.01E-05	29.16969872	4.40E-05	<u>60.90101624</u>	0.00022977	0.201457232
Anger	<u>54.36896543</u>	0.025465896	0.052457965	0.000125654	10.45896548	0.000014569
Sadness	0.254563258	0.001215896	0.005214539	0.000048965	<u>65.45865716</u>	0.000505248
Anger	<u>60.25796855</u>	15.24865837	35.79586523	0.000475896	10.49568759	0.001457859
Fear	10.04524789	0.001254789	<u>56.45895236</u>	0.007854963	7.589651452	0.289651458
Disgust	0.152487563	<u>61.75682459</u>	1.256324856	34.45865585	31.14586459	8.43E-06

Consider the 4th row of Table E.1. The “Chosen Emotion” is “Anger” and the value for that is “54.36896543” which is the maximum in the row. When considering each and every maximum value (underlined value) in each row, the lowest is “54.36896543” which is greater than 50. Thus, 50 was chosen as T1.

E.2 Deciding Threshold 2 (T2)

Table E.2 shows the values that have been used to decide T2 for Prototype 2. When considering one row, the “Chosen Emotion” field consists of the emotion that the user has agreed to choose. The value which is underlined is the maximum value emotion and if the column name of the maximum value and “Chosen Emotion” matches, it is correctly predicted.

Table E.2: Percentage Values When Detecting Emotions from Text

Chosen Emotion	Analytical	Anger	Confident	Fear	Joy	Sadness	Tentative
Analytical	<u>58.9295</u>	0	0	0	0	0	0
Tentative	0	0	0	0	0	51.5331	<u>99.8976</u>
Sadness	0	0	0	0	0	<u>51.6147</u>	0
Analytical	<u>76.2356</u>	0	0	0	0	55.6758	0
Confident	0	0	<u>96.1633</u>	0	0	55.4018	0
Sadness	0	0	0	0	0	<u>54.4112</u>	0
Anger	0	<u>62.5291</u>	0	0	0	0	0
Confident	0	0	<u>75.1512</u>	0	0	0	0

Consider the 3rd row of Table E.2. The “Chosen Emotion” is “Sadness” and the value for that is “51.6147” which is the maximum in the row. When considering each and every maximum value (underlined value) in each row, the lowest is “51.6147” which is greater than 50. Thus, 50 was chosen as T2.

Appendix F: Prototype Evaluations

F.1 Prototype 1 Evaluation

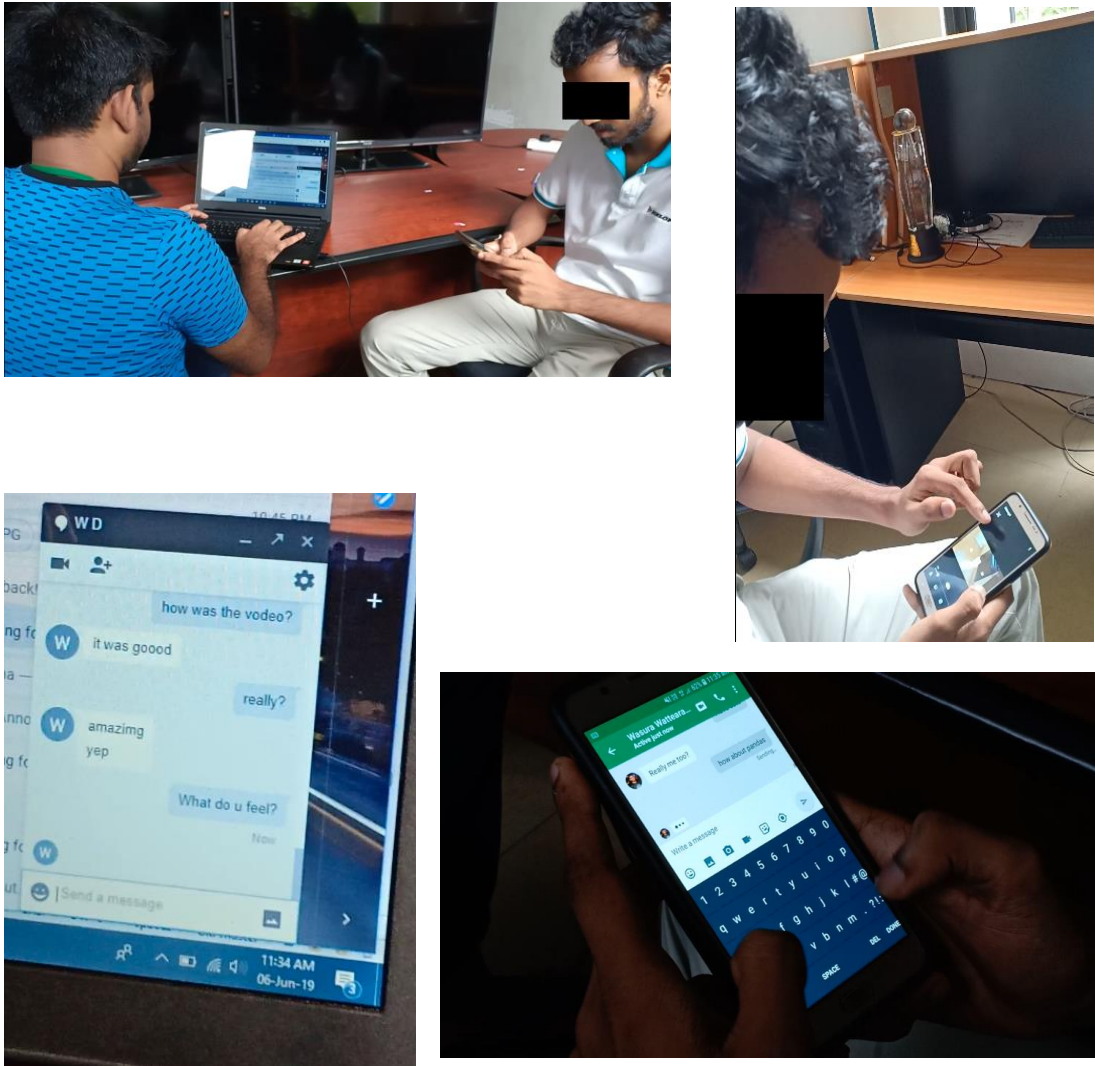


Figure F.1: Evaluating Prototype 1

F.2 Prototype 2 Evaluation

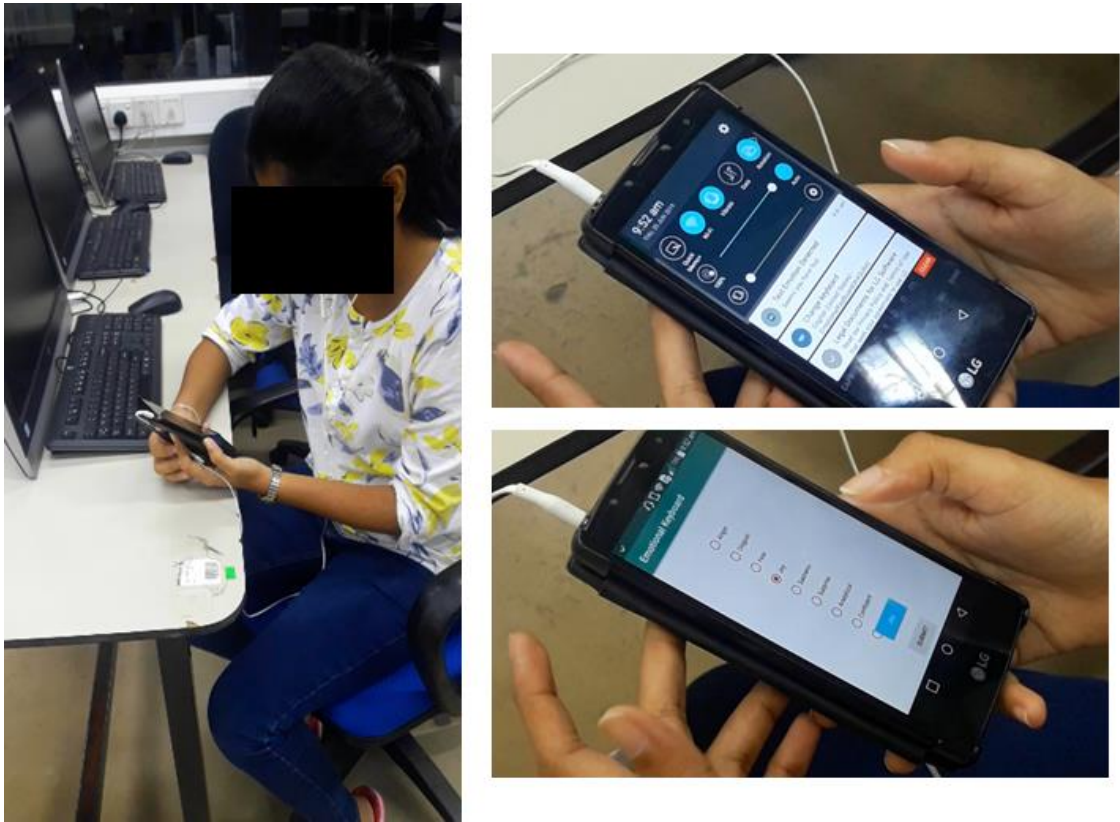


Figure F.2: Evaluating Prototype 2

F.3 Prototype 3 Evaluation

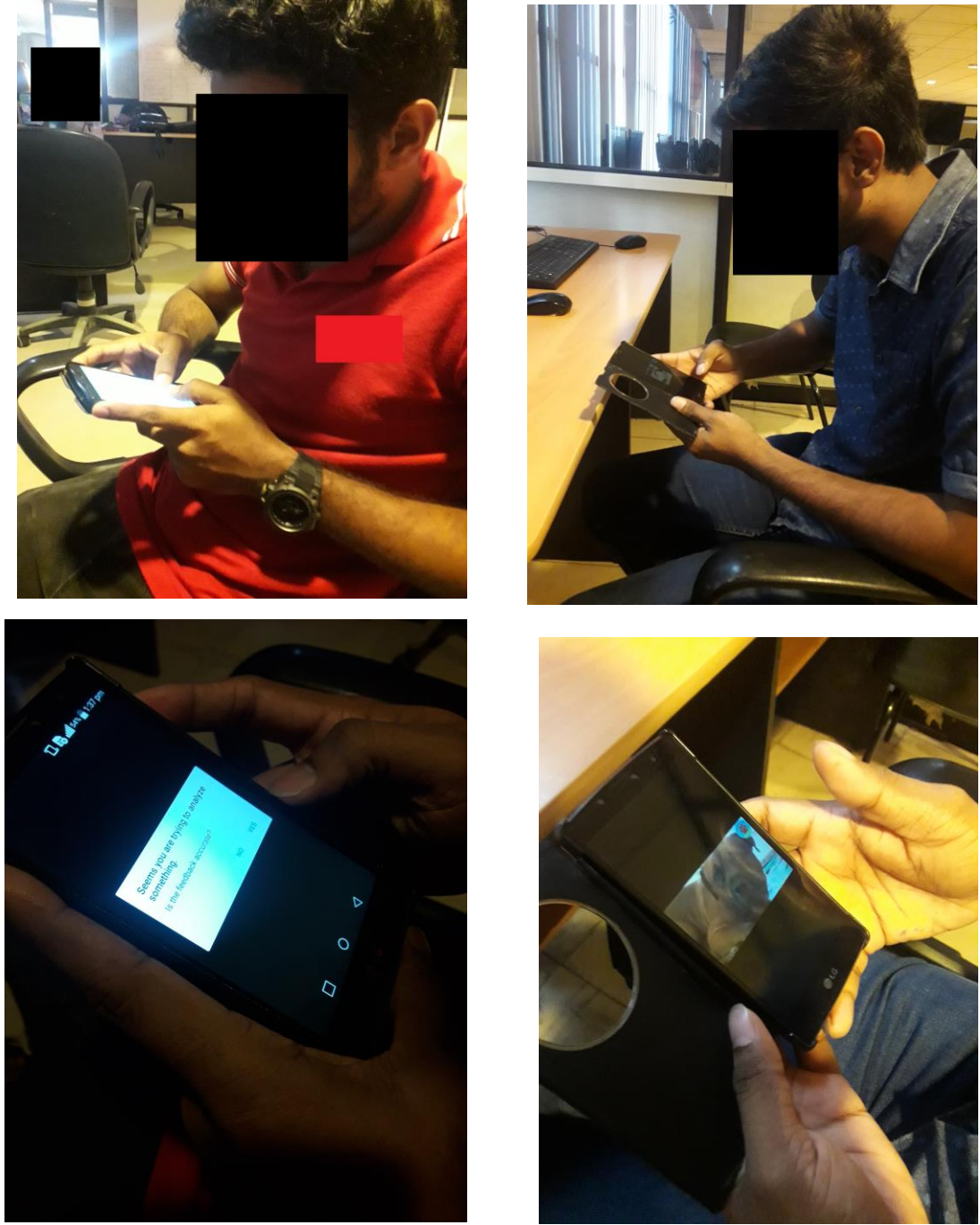


Figure F.3: Evaluating Prototype 3

F.4 Prototype 4 Evaluation

The graphs which were plotted for 2 users where YES count and NO count plotted against the time as shown in Fig. F.4 and Fig. F.5.

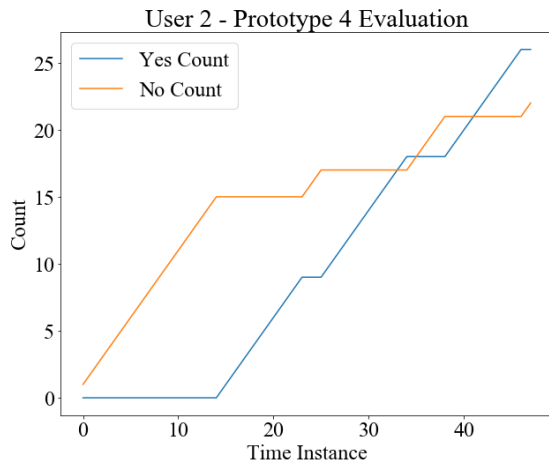


Figure F.4: Prototype 4 - User 2 evaluation

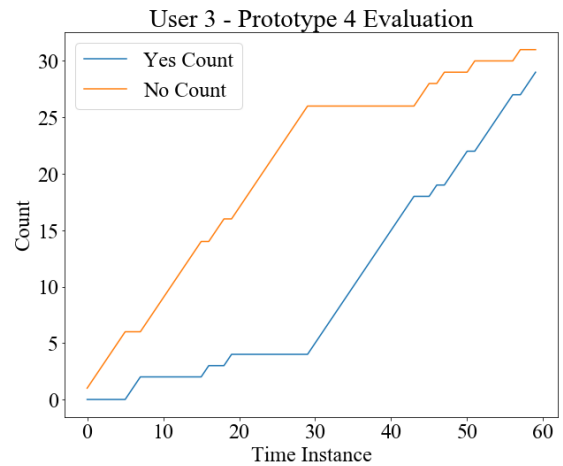


Figure F.5: Prototype 4 - User 3 evaluation

G.5 Prototype 5 Evaluation

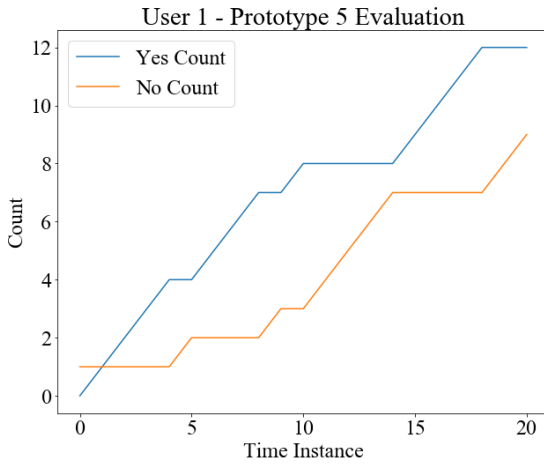


Figure F.6: Prototype 5 - User 1 evaluation

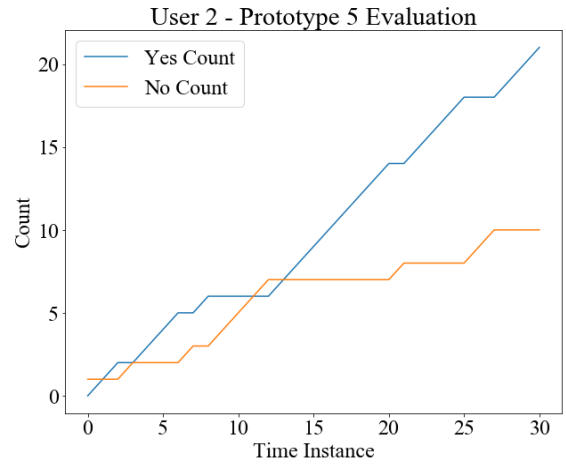


Figure F.7: Prototype 5 - User 2 evaluation

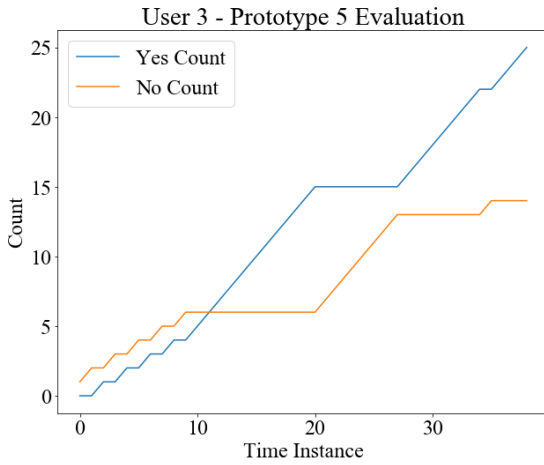


Figure F.8: Prototype 5 - User 3 evaluation

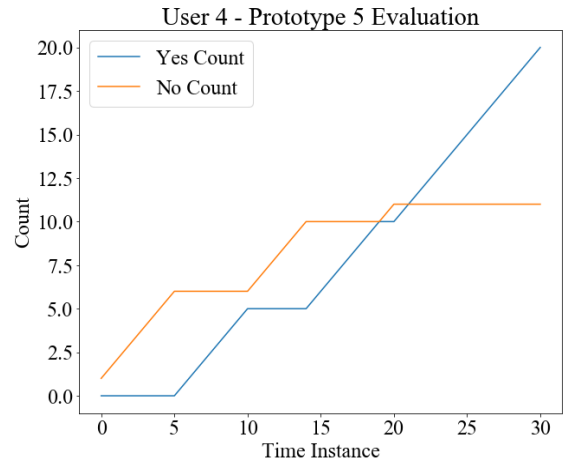


Figure F.9: Prototype 5 - User 4 evaluation

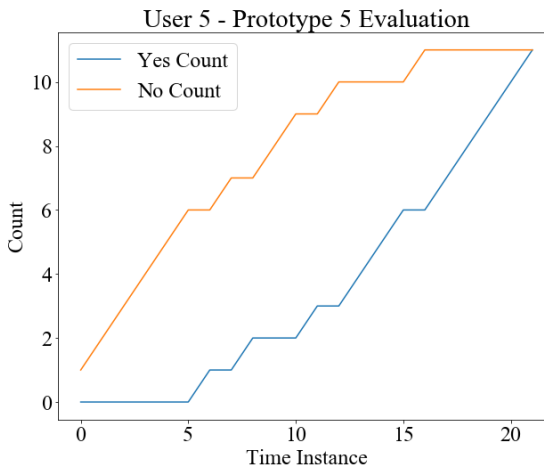


Figure F.10: Prototype 5 - User 5 evaluation

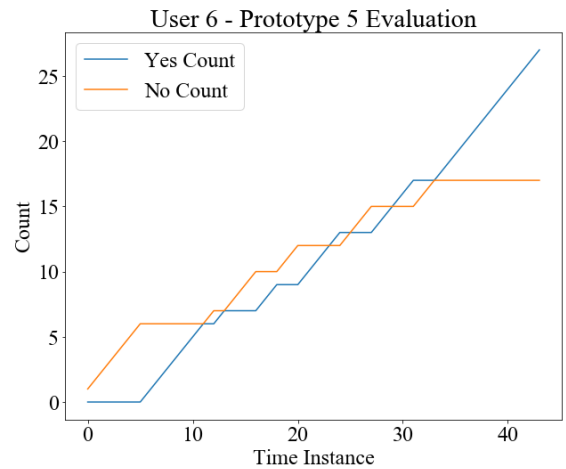


Figure F.11: Prototype 5 - User 6 evaluation

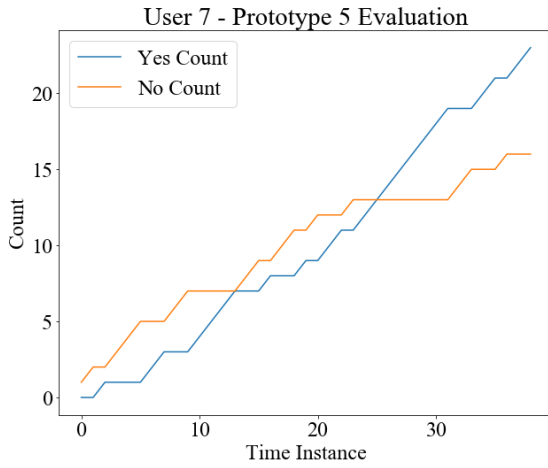


Figure F.12: Prototype 5 - User 7 evaluation

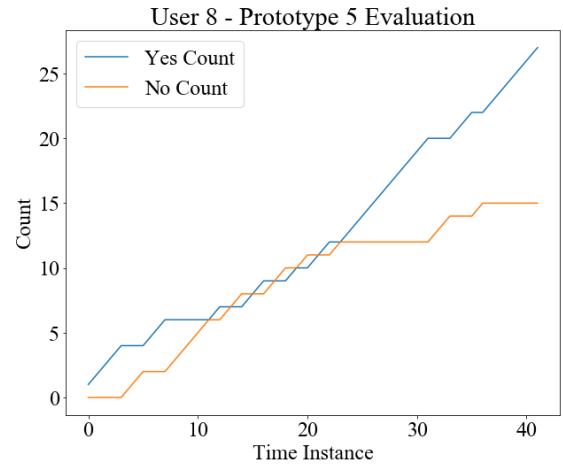


Figure F.13: Prototype 5 - User 8 evaluation

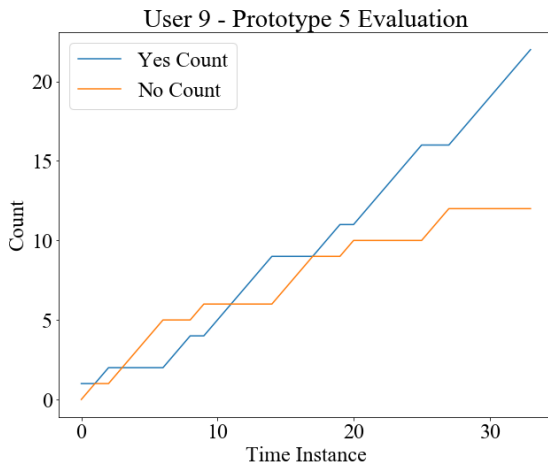


Figure F.14: Prototype 5 - User 9 evaluation

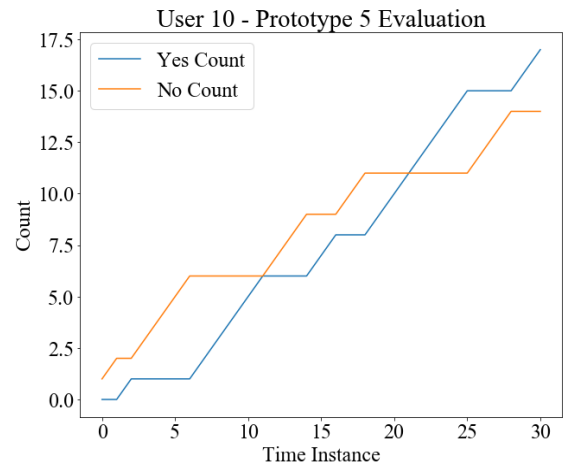


Figure F.15: Prototype 5 - User 10 evaluation

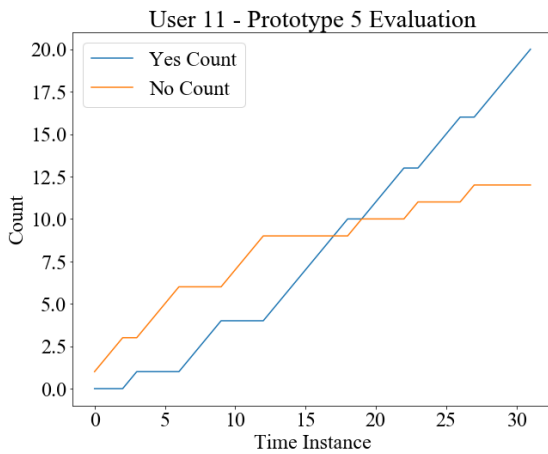
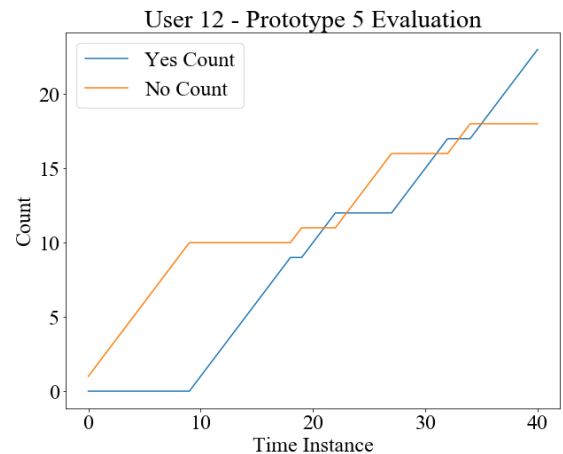
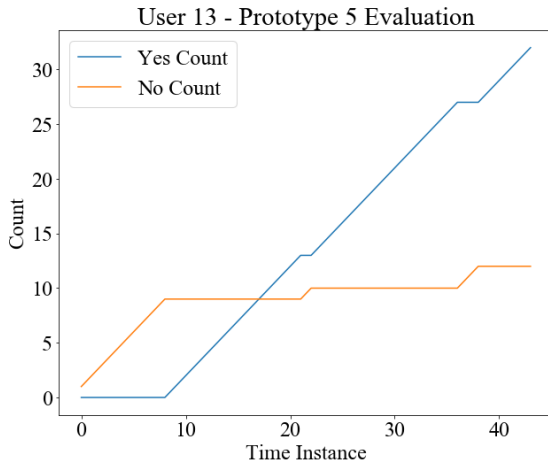


Figure F.16: Prototype 5 - User 11 evaluation





Appendix G: Publications

G.1 Publication 1

A paper titled “**Critical Success Factors of Analysing User Emotions to Improve Usability of Systems**” was presented in the **12th International Conference on Ubi-Media Computing in Bali, Indonesia (August 6th - 9th, 2019)** is attached below.

Critical Success Factors of Analysing User Emotions to Improve the Usability of Systems

Wasura. D. Watterarachchi
University of Colombo School of Computing
Colombo 7, Sri Lanka
wasuradananjith@gmail.com

Enosha Hettiarachchi
University of Colombo School of Computing
Colombo 7, Sri Lanka
eno@ucsc.cmb.ac.lk

K. P. Hewagamage
University of Colombo School of Computing
Colombo 7, Sri Lanka
kph@ucsc.cmb.ac.lk

Abstract— Emotions of a person depend on several factors like personality, relationships, mental health, and among them, context also plays a major role. The acceptability and usability of a system can be enhanced when the user emotions are considered based on the context. Most of the past research has been carried out related to the detection of emotions using facial expressions, gaze direction, heart rate analysis, Electroencephalogram (EEG) signals, text analytics, and keystroke dynamics along with mouse movements. This paper is a literature review which summarizes the previous research work in order to identify the critical success factors of the existing approaches when detecting user emotions through facial expressions and text data along with the context. Finally, as a conclusion, considering the advantages and drawbacks of the existing applications, the novel approaches that need to be followed to analyse user emotions to improve the usability of systems are proposed.

Keywords— Facial Behavior Recognition, Emotion Analysis in Text, Emotion Detection, Human-Computer Interaction (HCI)

I. INTRODUCTION

Emotions of a person is a factor which contributes to the improvement of a system by providing a positive and comfortable experience while using the system. Detecting user emotions can be done through several methods such as using facial expressions [1], pupil dilation [2], gestures [3], voice [4], typing pattern and keystroke analysis, and text analytics [5]. When detecting the emotions, one can apply the above methods individually. However, the accuracy of detecting the correct emotion can be improved when two or more of the above methods are combined [6].

Apart from the emotions, by identifying the context and the environment of the user, more dynamic and affective system functionalities such as changing the colour, font size, widgets of the system can be introduced and implemented in the systems [7] in order to improve the User Experience (UX). Due to the availability of smart technologies, these functionalities can be embedded into mobile devices [8] rather than incorporating them with traditional desktop systems [7].

Several types of research have been carried out to detect user emotions through facial expressions [9]–[11] and text data [5], [8], [12]. Therefore, this paper is a literature review which summarizes the previous research work with respect to detecting user emotions through facial expressions and text data along with the context by also identifying the critical success factors of some of the applications developed. The main aim is identifying the best methods to detect emotions and increase the

acceptability and usability of the system by reviewing previous work related to adaptive systems. In order to achieve this aim, approximately 30 research papers which were written in English, during the period of 1969 to 2019 were analysed.

The rest of the paper is structured as follows. Section 2 covers the background of the paper which gives a brief introduction about the terms that are related to research. Section 3 reviews the related techniques that have been used to detect emotions using facial expressions and text data. Section 4 discusses some applications designed based on user emotions and context to provide adaptable systems. Section 5 concludes our discussion and future work are stated in Section 6.

II. BACKGROUND

This section provides a brief overview of the terms which are related to this paper.

A. Emotions

There are different types of emotion models which can be used to label the emotions in a discrete manner or as a mixture of two or more emotions [12]. Emotion models can be divided into two types [13] as Emotion Categories and Emotion Dimensions.

Most researchers have conducted experiments using emotion categories. Ekman's model [14] is the most popularly discussed model among many researchers. It identifies 6 discrete emotions, namely, anger, disgust, fear, joy, sadness, and surprise. Plutchik's wheel of emotions [15] is another example of emotion categorization and it consists of 4 couples of emotions – joy-sadness, anger-fear, trust-disgust, and anticipation-surprise.

When discussing Emotion Dimensions the circumplex model is an example which states that every emotion can be expressed in a linear manner by combining valence, arousal, and dominance [12]. Some researchers have used both 2D and 3D dimensions to measure the level of pleasure, relaxation, strain, attention, rejection, pleasant, and level of activation, in addition to arousal and valence [13].

B. Context

Context can be divided into sub-categories as shown in Fig. 1 [16].

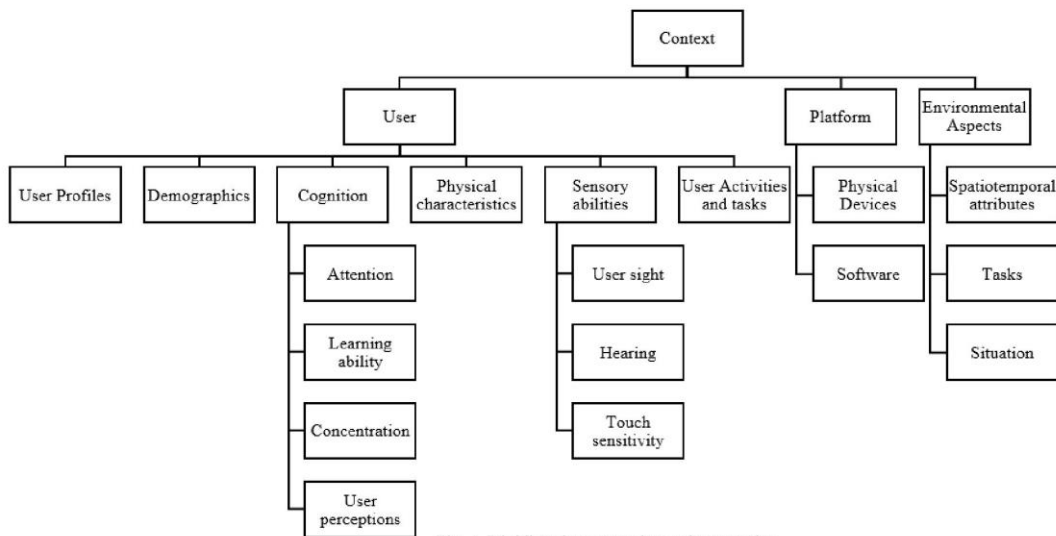


Fig. 1. Dividing the context into sub categories

The term context has a broad meaning and it can be used to take various decisions when making an affective system. Most of the researchers have taken physical location, time, weather and user activities as the context, thus leading to concepts like situated computing [7] and spatiotemporal structure learning [17].

III. REVIEW OF RELATED WORK

A review of emotion detection techniques using facial expressions and text analytics while considering some of their real-world applications are presented below.

A. Emotion Detection using Facial Expressions

As the statement, “*Your Face speaks louder than your voice*”, states, the best way to capture user emotions is through analysis of facial expressions [18]. Most researches are based on identifying emotions through analysing the facial muscle movements using the Action Units (AU) as explained in Facial Action Coding System (FACS) [19] which was originally developed by Carl-Herman Hjortsjö [20], and applied to experiments with both Single Label Classification and Multi Label Classification of emotions.

As shown in Fig. 2, facial expressions detection can be mainly decomposed into four steps such as Data Acquisition, Pre-Processing, Feature Extraction and Selection, and Prediction [1], [9], [10].

In Data Acquisition, both static images and videos are used as data. Some image databases such as Bosphorus, JAFFE (Japanese Female Facial Expressions) [9], and datasets like KDEF (Karolinska Directed Emotional Faces) [1] and CAFÉ (Child Affective Facial Expression) [10] are some examples for Data Acquisition models.

During the Pre-Processing stage, different sizes of images were converted to the same resolution or size and the brightness (converted to grayscale) was adjusted where necessary [9].

During Feature Extraction and Selection, different types of methods were used. Most of the researches have used the Viola-Jones detector to identify the regions of interest [1], [10]. Apart from that, the Euclidean distance [9] had been used to measure the distance between different points in the face, and the most appropriate set of features to work with had been chosen.

After the necessary features were selected and a model was trained, Support Vector Machines (SVM), k-NN, Bagged Trees, RUSBoost Trees were used to perform predictions [1], [2], [9] where most of the researchers have proved that Bagged Trees [1] and the RUSBoost Trees [2] are the best.

Most of the researches have gained high accuracies when detecting emotions using facial expressions and popular APIs, and SDKs such as Microsoft cognitive services, Kairos, Eyris EmoVu, and Affectiva [21] were developed.

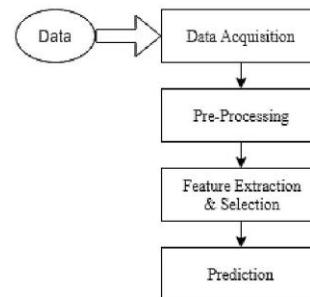


Fig. 2. Steps in facial expression detection

B. Emotion Detection using Text Analytics

Expression of emotions can be identified through two methods, namely, vocabulary-based and affective terms based. Words such as sad, love, happy, hate can be considered under vocabulary-based methods while terms such as ugh, eww, yuhuu can be considered under affective terms based methods [13].

There are free datasets where the data has been categorized using Ekman's model as well as non-standard methods [12], [22]–[27] as shown in Table I.

Experiments were mainly performed under two directions such as Rule-Based and Machine Learning Based. Fig. 3 shows some popular approaches to conduct text analysis [12], [13], [27] from which the Machine Learning Based approaches have gained the highest popularity.

In Rule-based methods, popular dictionaries like WordNetAffect and SentiWordNet were used which were developed by accumulating the emotional words [12]. Some of the lexicons have built using crowdsourcing where the particular set of users were asked to rate the words according to Ekman's model [14] and Plutchik's emotions [15].

TABLE I. DATASETS FOR TEXT ANALYTICS

Dataset	Description	Emotion Categories	Availability
AffectiveText [12], [27]	News headlines	Ekman's Model	Free
Blogs [12], [26]	Sentences from blogs	Ekman's Model	Contact Authors
DailyDialogs [12], [28]	Conversation data	Ekman's Model	Free
Tales [12]	Sentences from fairy tales	Ekman's Model	Free
TEC [12]	Twitter hashtags	Ekman's Model	Free
Emotion-Stimulus [12]	Sentences annotated with emotion and cause	Ekman's Model + shame	Free
CrowdFlower [12]	Tweets	Non-standard	Free
ISEAR [12]	Questionnaires answered by people from different cultures	Non-standard	Free
Electoral-Tweets [12]	Tweets on elections	Non-standard	Free
EmoInt [12]	Tweets	Non-standard	Free
fb-valence-arousal [12]	From Facebook posts	Valence, Arousal	Free
Neviarouskay a et al.'s dataset 1 [22]	Stores in 13 different categories	Non-standard	Not sure
Neviarouskay a et al.'s dataset 2 [23]	Sentences from diary-like posts	Non-standard	Not sure
Alm's dataset [25]	Sentences from fairy tales	5 from Ekman's Model	Not sure
Aman's Dataset [26]	Sentences from blogs	Ekman's Model + neutral	Not sure

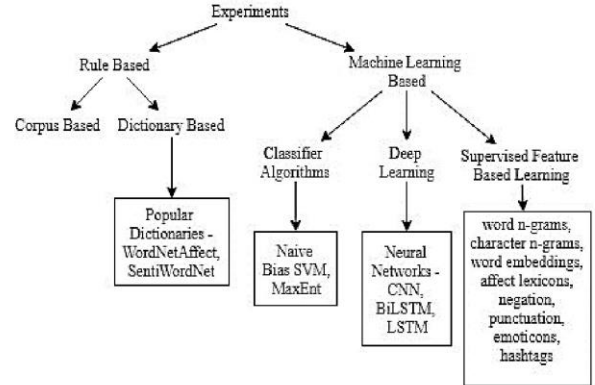


Fig. 3. Text analysis methods

When comparing with dictionary-based methods in Rule-based experiments, Machine Learning based methods like Supervised Feature-based learning approaches have improved the performance by using the features such as word n-grams, character n-grams, word embeddings, affect lexicons, negation, punctuation, emoticons, or hashtags with the classifiers such as Naive Bayes, SVM [29]. Neural Networks such as CNNs (Convolutional Neural Networks), BiLSTM (Bidirectional Long Short-Term Memory) [30] and LSTM (Long Short-Term Memory) [31] have compared against linear classifiers such as SVM (Support Vector Machine) and MaxEnt (Maximum Entropy), where the BiLSTM has given the better performance.

C. Emotion Detection by analysing the Typing Pattern

Keystroke dynamics and mouse movement analysis play a huge role when analysing the typing pattern [5], [22], [32].

Experiments were mainly performed using 2 methods; reading paragraphs and watching videos. A set of participants were given a paragraph to read and were subsequently asked to type something on a computer [33]. The same procedure was followed for experiments based on video clips as well [5]. The main drawback of using these methods is that the participants can manipulate their emotions in order to suit the expected results of the experiment. For example, someone who does not feel happiness or sadness after watching an emotional video would pretend by expressing a fake emotion.

Further, some experiments had been conducted to analyse the mouse scrolling pattern by combining the above two approaches. After watching a video, the participants were asked to read a paragraph by scrolling down and then, type something at the end. The disadvantage of this method is while reading the paragraph the emotion might fade away thus leading the researcher to detect neutral emotions [5].

In those experiments, the following behaviors were identified.

- Participant types in a normal way
- Participant types in a quick manner

- Participant types in a slow manner
- Participant uses backspace key often
- Participants hit unrelated keys often
- Participant does not use the keyboard

By analysing these behaviors some researchers have extracted typing speed, number of characters typed in the 5-second interval, the time taken to type a given text, number of times the backspace was used, and idle time between typing and not typing [33]. Also, features such as average time of left/right/double clicks, mouse scroll, cursor x/y distances, key down to key up time and vice versa, key down to key down time, regular key press, and enter/arrow/backspace/function keys pressed had been extracted [5].

By analysing the extracted features, negative, positive or neutral emotion levels have been identified [33]. It has been stated that, when there are negative feelings, mistakes happen often due to the clicking of unnecessary keys which leads to the usage of backspaces to correct them, as well as it decreases the typing speed. User types normally when he/she is in a neutral state and typing speed will increase when in a positive state.

When compared with facial expression recognition, this method has less accuracy due to biased experiment results which were discussed earlier such as faded emotions with time and manipulation of fake emotions [5]. Further, user emotions are not the only factor which influences the typing pattern, but also the novelty of the system to the user, familiar words and text, and also the old age has a considerable effect on the typing pattern [34].

IV. APPLICATIONS DESIGNED TO IMPROVE USABILITY BASED ON EMOTIONS AND CONTEXT

Many applications were developed by taking user emotions into account with the aim of producing effective and affective usable systems through personalization and customization. Some of these applications are stated below.

A. EmotionPush

EmptionPush is an Android app with the functionality of detecting the emotion of messages in Facebook Messenger and producing push notifications which state the emotions [8]. Machine Learning techniques based on LibSVM had been used for the implementation of the model. The deployment was done through the Google Play Store where 8 native English speakers were chosen as participants. They have observed that the wrong prediction of emotions did not interfere with users' chatting experience. During the analysis, they have seen that user behavior changes accordingly after seeing the push notifications. Also, the users have instantly read the messages which had feelings of sadness and anger, while the messages with joy were kept away to be read later.

This research has received high performance for the emotions of joy, sadness, and anger, but low performance for neutral and fear which is a drawback. So, they have planned to improve the performance of predicting emotions in the future while providing an improved interface for ease of use.

B. An Adaptive User Interface based on Spatiotemporal Structure Learning for Samsung Galaxy S

Spatiotemporal structure learning means creating a probability model to express the information which was captured through the context of the user [17]. In here, the context stands for time, location, weather, emotion, and the activities of the user. It is important to note that they have taken the emotions embedded into the context to carry out this study.

To provide the user with a more personalized interface they have developed a system to recommend the user with the most suitable apps to be used by analysing the context. They have considered 3-axis accelerometer data and the user had been asked to manually enter the emotion to the system. For future work, they have planned to detect emotions through facial expressions and voice data.

They have given a questionnaire to collect necessary information to decide the most preferred mobile apps by considering the emotion (positive, negative, neutral) of a mobile phone user when performing some activity (walk, run, ride a vehicle) at particular time (morning, afternoon or evening), at a particular location (home, company) on a particular weather condition (sunny, cloudy, rainy).

They have used naïve bias probabilistic model for the learning phase and have identified meaningful associations which have higher conditional probabilities and have used them for predictions.

C. Situation-dependent Browser

In this application, the situation represents the context which is denoted by $S(\mathbf{L}, \mathbf{t}, \mathbf{a}, I)$ where \mathbf{L} stands for location, \mathbf{t} stands for the time period, \mathbf{a} stands for a set of user activities, and I stands for a set of other significant factors [7].

They have come up with a browser which provides the functionality to do browsing based on the situation (context). For example, someone can think of a video file which he/she watched last week at a friend's home. To search that video, he/she can use this situation-dependent browser.

The main advantage of this research is that it helps the user in an intelligent manner to manage the data and improves the efficiency of searching a file in a desktop system.

Most of the applications were not be able to identify all emotions accurately. For some emotions such as joy, sadness, and anger, they have given high accuracy but for emotions such as neutral and fear the accuracy was low. This can be due to less amount of data which was used for training the machine learning models. Therefore, systems with higher accuracies could be developed by providing more data and carry out the experiments with more user participation to generalize the decisions.

V. CONCLUSION

User emotions and the context play a major role when it comes to developing a system which is more suitable for the user while having the ability to adapt according to the person who is using it, at the same time. There are both advantages and drawbacks of using different kinds of techniques when detecting emotion using facial expressions or text data. The context of the user can also be mixed with the user emotions in order to get

better results. Existing applications like EmotionPush [8], Adaptive Interface for Samsung Galaxy S [17], Situation-dependent Browser [7] are examples for this.

These applications were not able to give high performance for some emotions such as neutral and fear, and also required more data for the training process.

VI. FUTURE WORK

In order to facilitate further improvement, models could be created to gather user data and emotions which could then be stored in a profile. It can be used later to predict the most affective action to be performed by the system based on the context data and also to provide effective feedback.

REFERENCES

- [1] M. Kolodziej, A. Majkowski, R. J. Rak, P. Tarnowski, and T. Pielaszkiewicz, "Analysis of facial features for the use of emotion recognition," in *Proceedings of 2018 19th International Conference Computational Problems of Electrical Engineering, CPEE 2018*, 2018, pp. 1–4.
- [2] S. Porcu, A. Floris, and L. Atzori, "Towards the Prediction of the Quality of Experience from Facial Expression and Gaze Direction," in *ICIN 2019 QoE Management Workshop, Paris*, 2019, pp. 82–87.
- [3] C. Corneanu, F. Noroozi, D. Kaminska, T. Sapinski, S. Escalera, and G. Anbarjafari, "Survey on Emotional Body Gesture Recognition," *CoRR*, vol. abs/1801.0, 2018.
- [4] S. Lugovic, I. Dunder, and M. Horvat, "Techniques and applications of emotion recognition in speech," in *2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2016, pp. 1278–1283.
- [5] R. Shikder, S. Rahaman, F. Afroze, and A. B. M. A. Al Islam, "Keystroke/mouse usage based emotion detection and user identification," *2017 Int. Conf. Networking, Syst. Secur.*, pp. 96–104, 2017.
- [6] G. A. Tshrintzis, M. Virvou, E. Alepis, and I. O. Stathopoulou, "Towards improving visual-facial emotion recognition through use of complementary keyboard-stroke pattern information," *Fifth Int. Conf. Inf. Technol. New Gener. (ITNG 2008)*, pp. 32–37, 2008.
- [7] M. Hirakawa, P. Hewagamage, and T. Ichikawa, "Situation-dependent browser to explore the information space," *Proceedings. 1998 IEEE Symp. Vis. Lang. (Cat. No.98TB100254)*, pp. 108–115, 1998.
- [8] S.-M. Wang, C.-H. Li, Y.-C. Lo, T.-H. K. Huang, and L.-W. Ku, "Sensing Emotions in Text Messages: An Application and Deployment Study of EmotionPush," pp. 1–6, 2016.
- [9] H. A., T. A., and A. T., "Face Behavior Recognition Through Support Vector Machines," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 1, pp. 101–108, 2016.
- [10] S. Anwar and M. Milanova, "Real Time Face Expression Recognition of Children with Autism," *Int. Acadmy Eng. Med. Res.*, no. 1, 2016.
- [11] M. Dubey and L. Singh, "Automatic Emotion Recognition Using Facial Expression: A Review," *Int. Res. J. Eng. Technol.*, vol. 3, no. 2, pp. 488–492, 2016.
- [12] L. A. M. Bostan and R. Klinger, "An Analysis of Annotated Corpora for Emotion Classification in Text Title and Abstract in German," in *Proceedings of the 27th International Conference on Computational Linguistics*, 2018, pp. 2104–2119.
- [13] N. M. Hakak, M. Mohd, M. Kirmani, and M. Mohd, "Emotion analysis: A survey," in *2017 International Conference on Computer, Communications and Electronics (COMPELX)*, 2017, pp. 397–402.
- [14] E. Paul, "Facial expression and Emotion," *Am. Pshychologist Assoc.*, vol. 49, pp. 384–392, 1993.
- [15] R. Plutchik, "A psychoevolutionary theory of emotions," *Soc. Sci. Inf.*, vol. 21, no. 4–5, pp. 529–553, 1982.
- [16] M. Afzal *et al.*, "Model-based adaptive user interface based on context and user experience evaluation," *J. Multimodal User Interfaces*, vol. 12, no. 1, pp. 1–16, 2018.
- [17] H. Lee, Y. Choi, and Y. Kim, "An Adaptive User Interface based on Spatiotemporal Structure Learning," in *2011 IEEE Consumer Communications and Networking Conference (CCNC)*, 2011, pp. 923–927.
- [18] K. J. Kantharia and G. I. Prajapati, "Facial behavior recognition using soft computing techniques: A survey," in *2015 Fifth International Conference on Advanced Computing Communication Technologies (ACCT)*, 2015, pp. 30–34.
- [19] M. A. Sayette, J. F. Cohn, J. M. Wertz, M. A. Perrott, and D. J. Parrott, "A Psychometric Evaluation of the Facial Action Coding System for Assessing Spontaneous Expression," *J. Nonverbal Behav.*, vol. 25, pp. 167–186, 2002.
- [20] C. H. Hjortsjö, *Man's Face and Mimic Language*. 1969.
- [21] R. S. Deshmukh and V. Jagtap, "A survey: Software API and database for emotion recognition," *2017 Int. Conf. Intell. Comput. Control Syst.*, pp. 284–289, 2018.
- [22] A. Neviarouskaya, H. Prendinger, and M. Ishizuka, "Textual Attitude Analysis Model," *Proc. NAACL HLT Work. Comput. Approaches to Anal. Gener. Emot. Text*, pp. 80–88, 2010.
- [23] A. Neviarouskaya, "Recognition of Fine-Grained Emotions from Text: An Approach Based on the Compositionality Principle Recognition of Affect, Judgment, and Appreciation in Text," *Smart Innov. Syst. Technol.*, pp. 179–207, 1970.
- [24] S. Chaffar and D. Inkpen, "Using a Heterogeneous Dataset for Emotion Analysis in Text," *Springer*, pp. 62–67, 2011.
- [25] E. C. O. Alm, "Affect in Text and Speech," 2008.
- [26] S. Aman and S. Szpakowicz, "Identifying Expressions of Emotion in Text," in *Text, Speech and Dialogue*, 2007, pp. 196–205.
- [27] C. Strapparava and R. Mihalcea, "Affective Text," 2007.
- [28] Y. Li and H. Su, "DailyDialog: A Manually Labelled Multi-turn Dialogue Dataset," pp. 986–995, 2017.
- [29] S. Mohammad, "SemEval '12: Proceedings of the First Joint Conference on Lexical and Computational Semantics - Volume 1: Proceedings of the Main Conference and the Shared Task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation," 2012, pp. 246–255.
- [30] M. Schuster and K. K. Paliwal, "Bidirectional Recurrent Neural Networks," vol. 45, no. 11, p. 6757, 1997.
- [31] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [32] A. F. M. N. H. Nahin, J. M. Alam, H. Mahmud, and K. Hasan, "Identifying emotion by keystroke dynamics and text pattern analysis," *Behav. Inf. Technol.*, vol. 33, no. 9, pp. 987–996, 2014.
- [33] P. Khanna and M. Sasikumar, "Recognising Emotions from Keyboard Stroke Pattern," *Int. J. Comput. Appl.*, vol. 11, no. 9, pp. 1–5, 2010.
- [34] A. Pentel, "Predicting Age and Gender by Keystroke Dynamics and Mouse Patterns," no. July, pp. 381–385, 2017.

G.2 Publication 2

Another publication titled **“Analysing the Compatibility of Identifying Emotions by Facial Expressions and Text Analytics when Using Mobile Devices”** has been accepted by the **2020 International Conference on Machine Learning and Human-Computer Interaction (MLHMI) in Singapore** and was presented in March 2020. The acceptance notification is attached below.



50 Nanyang Ave, Singapore 639798 (<http://www.acm-sg.org/>)

Notification of Acceptance of the MLHMI 2020

March 20-22, 2020 in Singapore

<http://www.mlhmi.org/>

Paper ID: ML0001

Paper Title: Analysing the Compatibility of Identifying Emotions by Facial Expressions and Text Analytics when Using Mobile Devices

Dear Wasura Wattarachchi, Enosha Hettiarachchi and Kamalanath Hewagamage,

First of all, thank you for your concern. **2020 International Conference on Machine Learning and Human-Computer Interaction** received papers from different countries and regions during the submission period which were reviewed by international experts in related field. According to the recommendations from reviewers and technical program committees, we are glad to inform you that your paper identified above has been selected for publication and oral presentation. You are invited to present your paper and studies during our MLHMI conference that would be held March 20-22, 2020 in Singapore. Your paper was blind-reviewed and based on the evaluation. The reviewers' comments are enclosed.

Paper of MLHMI 2020 will be published in International Conference Proceedings Series by ACM, which will be indexed by **Ei Compendex, Scopus** and submitted to be reviewed by and Thomson Reuters Conference Proceedings Citation Index (CPCI).

(Important Steps for your registration): Please do finish all the 6 steps on time to guarantee the paper published in the proceedings successfully:

1. Revise your paper according to the Review Comments in the attachment carefully. (Five authors at most each paper)

2. Format your paper according to the Formatting Instructions carefully.

<http://www.mlhmi.org/template.docx> (DOC Format)

3. Download and complete the Registration Form.

<http://www.mlhmi.org/reg.doc> (English)

4. Finish the payment of Registration fee. (The information can be found in the Registration form)

<http://www.mlhmi.org/reg.doc>(English)

5. Add the CCS Concepts part and XML code in your paper.

<http://www.mlhmi.org/CCS&XML.docx>



50 Nanyang Ave, Singapore 639798 (<http://www.acm-sg.org/>)

6. Send your final papers (both .doc and .pdf format), filled registration form (.doc format), and the scanned payment (in .jpg format) to us at mlhmi.conference@gmail.com (**Before December 20, 2019**) (**Very important**)

Note: Conference secretary will contact you later about the **copyright**, please pay attention to your e-mail box.

APIT 2020 will check the format of all the registered papers first, so the authors don't need to upload the paper to the system. After the registration, we will send all qualified papers to the publish house and index organization for publishing directly.

Please strictly adhere to the format specified in the conference template while preparing your final paper. If you have any problem, please feel free to contact us via mlhmi.conference@gmail.com. For the most updated information on the conference, please check the conference website at <http://www.mlhmi.org/>. The Conference Program will be available at the website in **late February, 2020**.

Again, congratulations. We are looking forward to seeing you in Singapore.

Yours sincerely,



MLHMI 2020