BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Project Id: 2021-073

Project Proposal Report

De Silva L. S. R

B.Sc. (Hons) Degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka institute of Information Technology

Sri Lanka

February 2021

BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Project Id: 2021-073

Project Proposal Report

Dr. Pradeepa Samarasinghe

B.Sc. (Hons) Degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka institute of Information Technology

Sri Lanka

February 2021

DECLARATION

I declare that this is my own work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student ID	Signature
De Silva L. S. R.	IT18113914	

The supervisor/s should certify the proposal report with the following declaration. The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

…….. …….

Signature of the Supervisor: Date

[Dr. Pradeepa Samarasinghe]

…….. …….

Signature of the Co-Supervisor: Date

[Ms. Vijani Piyawardana]

ABSTRACT

Depression is a chronic disease that can negatively affect the emotional and physical states of an individual. It is necessary to identify the early signs of major depressive disorder to ensure the affected person seeks immediate attention. Psychotherapy assesses depression using standard questionnaires in clinical interviews. However, the assessment may bias with the different perspectives of individuals. Social networks facilitate sharing opinions with others, which reflects the sentiments they are going through. Accordingly, depression can be analyzed using social media. Currently, the mobile applications associated with depression are using standard questionnaires to assess the depression level. In Sri Lanka, people may obscure the state of their minds due to social stigma and fear. However, a self-awareness mobile application that can evaluate own mental health and identify early depression symptoms, can be more valuable to remit the danger. This study aims to develop an automated system that can identify early symptoms of depression users from their public social media activity. Also, the proposed system considers the chatbot conversation with having the individual. In the proposed model, Natural Language Processing (NLP) techniques such as topic Modeling, Sentiment Analysis study the social-medial content and chatbot conversation to identify early signs that may lead to being a depressive person.

Keywords: Depression, Machine Learning, Social media sentiment analysis, Feature extraction

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Globally, mental health problems are the leading cause of morbidity and mortality [1]. Depression is a common mental disorder that primarily contributes to the global burden of disease [2]. According to the World Health Organization's recent reports, the estimation states that more than 264 million people of all generations suffer from depression [3]. A key concern for increasing the number of depressed individuals is limited resources and support services. As mentioned in [4], Psychological and pharmacological treatments lacked resources in underdeveloped countries. Even 50% of depressed individuals do not receive treatments in developed countries due to limited or no support available with mental health disorders [5].

As an underdeveloped country, mental wellbeing is becoming an ongoing epidemic in the Sri Lankan community since the deficiency of expert assistance and community stigma limits mental care. Hence, depressed individuals are reluctant to get treated by engaging with mental care professionals. The treatment limitations arisen from social stigma can eliminate with Social media behavior analysis for depression. Social media behavior can reveal extensive patterns that might be related to depressive disorder. An accurate analysis on social media may help to monitor early depression signs and get treated in advance. Although a medical examination indicates the general state of health of a patient, a specialist may use other aspects related to making a diagnosis of depression. Monitoring the early depression signs would be beneficial in reducing expensive and long-lasting treatment procedures.

1.1 Background & Literature Survey

Previous studies were based on social media content to detect depression rather than identifying early depression signs. De Choudhury M et al. demonstrated a statistical model to estimate depression risk through the Twitter activities of the users. English-speaking Twitter users considered measuring and predicting depression in individuals. The study Proposed the SVM classifier using behavioral cues (linguistic styles, depressive language, ego network, etc.) to estimate the depression risk with an accuracy of approximately 70% [6]. Tsugawa et al. investigated a method to recognize

depression using various features (frequencies of word usage, the ratio of positive/ negative affect words, number of users following, number of users followed, etc.) obtained from Twitter history activities [7]. The study extended the De Choudhury M et al. prediction framework to Japanese-speaking Twitter users [6]. Delahunty et al. developed a machine learning classifier to identify the depression level of the user by considering the chatbot conversation [8]. C. S. A. Raza et al. focused on developing a web application to analyze posts of Twitter users for detecting personal depression attitude using machine learning. The study used three different techniques (Naive Bayes Classifier technique, NLP techniques, Deep Learning technique) to analyze the depression risk by considering positive and negative tweets [9].

Few studies examined the link between depression and social media activities. The effectiveness of using language and behavior data from social media to assess depression has been shown in these studies. Md. Rafiqul Islam et al. focused on analyzing depression through the Facebook activities of the users. The study observed the decision tree algorithm as the highest accuracy in emotional process and linguistic style comparing to other machine learning techniques (KNN, SVM, Ensemble). Facebook comments were considered in predicting depression among Facebook users [10]. Schwartz et al. studied a shortlist of words, topics, phrases to analyze depression. Further, the study focused on seasonal fluctuations of depression [11].

1.2 Research Gap

The previous studies in the domain of mental illnesses mainly focused on social media activities. Few studies explored depression analysis based on chatbot conversation, while the other researchers considered the social media content in analyzing depressive disorder features. The study of [8] focused on depressive disorder based on chatbot conversation, while both the studies in [6, 7] explored social media content-based depression analysis. Correspondingly, the researchers of [9-11] analyzed social media content in identifying depressive features. However, none of the studies focused on implementing an automated analysis in identifying depressive disorder using both the social media content and chatbot conversation. Besides, none of the previous studies focused on the probability of early depressive features based on social media content

and chatbot conversation. The following Table 1.1 summarizes the limitations in previous studies.

Table 1. 1: Comparison between previous studies

	[6]	[7]	[8]	[9]	[10]	[11]	Proposed System
Social Media Content based Depression Analysis	✓	✓	✗	✓	✓	✓	✓
Chatbot conversation-based Depression Analysis	✗	✗	✓	✗	✗	✗	✓
Probability of early depression features based on social media content	✗	✗	✗	✗	✗	✗	✓
Probability of early depression features based on Chatbot conversation	✗	✗	✗	✗	✗	✗	✓

A mobile application that identifies the symptoms of depression helps depressed individuals to have a broad view of their mental health status.

“Happify” is a mobile application including effective tools and programs that helps the user to understand their feelings and thoughts. The science-based activities and games uplift the positivity of individuals. The application includes an AI coach to help individuals. Application is generating extensive mental health reports providing a synopsis of your emotional state. However, the Happify application does not include the mechanism to analyze the content, which leads to depression [12, 13].

The mobile application, “Wysa: Mental Health Support”, offers a range of tools to manage stress and wellness. The Artificial Intelligent (AI) chatbot is the core feature that reacts and responds to the user with video, articles, and exercise suggestions based on the expressed emotions throughout the conversation. Although the feelings are analyzed based on the conversation, the app does not consider a depression-related analysis [14, 15].

Woebot is a fully automated agent for the conversation that design to manage depression and anxiety feelings. It helps people to manage feelings of depression and

anxiety. Woebot offers guidance with interactive quizzes and videos based on Cognitive Behavioral Therapy (CBT). However, Woebot is not cable of analyzing the content to identify the individuals who may lead to depression [16-19].

Mobile Application Youper is designed as a self-help AI-guided application to support individuals to improve their mental health. Youper consists of 4 main features: AI chat-bot, Emotional health Assessment, Mood logs, and journal logs. Although a comprehensive summary of mental health is provided to users, this is not considered the depression symptoms recognition [20-22].

The following Table 1.2 compares the existing mobile application with the proposed system.

Table 1. 2: Existing Mobile Applications for mental health

	Happify [12, 13]	Wysa [14, 15]	Woebot [16-19]	Youper [20-22]	Proposed App
AI chatbot	✓	✓	✓	✓	✓
Social Media content analysis	✗	✗	✗	✗	✓
Emotion analysis based on conversation.	✗	✓	✓	✗	✓
Depression analysis based on conversation	✗	✗	✗	✗	✓
Depression analysis based on social media content	✗	✗	✗	✗	✓
Notify Guardian	✗	✗	✗	✗	✓
Comprehensive report Summary	✓	✗	✗	✓	✓

1.3 Research Problem

Depressed individuals with self-awareness would receive effective therapy for mental health. A mobile application, which analyzes the social media behavior to identify depressive disorder early signs, may help the individual to have self-awareness on mental health. Generally, the existing mobile applications focused on self-monitoring of depressive signs by analyzing the user logged information on questionnaires. The analysis might be more biased since the data is providing the user. Few mobile applications include Chatbot to have the conversation with individuals. However, conversation content is not analyzing to recognize depressive symptoms. According to the above literature, an approach of automated analysis of depressive disorder by identifying early depression signs using social media is not yet researched in Sri Lanka.

Depression diagnosis shortcoming in the industry summarized as follows.

- Difficult to monitor individuals manually for two weeks.
- Lack the automated solution to identify early depression symptoms using Social media platforms.
- Lack the automated solution to analyze conversations to identify early depression symptoms.

2. OBJECTIVES

Research is mainly focused on analyzing the early signs that may lead to being a depressive person. The study proposes to focus on a mobile application to identify behaviors that could be highly anticipated to contribute to depressive disorder.

2.1 Main Objective

The study mainly focuses on social media content analysis for the early identification of depressive disorder.

2.2 Specific Objectives

The social media content analysis for early identification of depressive disorder is divided into specific objectives as follows.

- **Build a database based on Social Media Content.**
  - Initially, the social media content should be acquired, and build a database to save the acquired social media content.

- **Identify social media texting patterns based on the built database.**
  - Social media texting patterns will be identified using the content in the built database.

- **Build a classifier to model, abnormal social activities based on social media content.**
  - A classifier will be built to identify abnormal social activities based on the social media content.

- **Predict the probability of abnormal social activities towards depression with the designed model.**
  - The probability of abnormal social activities towards depression will be predicted using the designed model.

- **Design and implement ChatBot to initiate conversations, identify depressive thoughts based on conversations.**
    - ChatBot will be designed to initiate conversation. The conversation will be retrieved and analyze depressive thoughts. Conversation with the ChatBot will predict the probability of abnormal thoughts towards depression.

Additionally, the following objectives need to be completed to develop the mobile application.

- **Integrate the model with the application.**
    - The built model will be integrated with the mobile application with the help of the Application Programming Interface (API).

- **Implementing interfaces.**
    - Frontend development will be focused on the interface of the proposed system.

- **Mobile Application development for social media analysis.**
    - The implementation of the mobile application will be considered in the mobile application development.

3. METHODOLOGY

3.1 Introduction

The implementation framework for detecting the signs of early depression is explained in this section. The illustration of the system diagram explains the procedure of performing research goals as seen in Figure 3.1.

Collect Data - Data needs to be in text format to train and evaluate the model. Text with its sentiment will use in building the model. Linguistic Inquiry and Word Count (LIWC2015) will use to identify the sentiment in the text [23].

Data Preprocessing - Data preprocessing is the initial step of the model building process. The aim is to acquire a cleaned dataset. The quality of the raw data can be obtained by applying preprocessing techniques as data cleansing, data transformation. Data cleansing consists of an unwanted string and emoji elimination, rows with missing values elimination. Data normalization will be considered to acquire standardized data [24].

Data Split - Split the preprocessed data into two categories as training data to build the model and test data to train the built model. Training data is required than test data.

Feature extraction - The most relevant data should be extracted to build the model accurately. Natural Learning Processing needs feature extraction techniques to convert text into a feature vector(matrix). TF-IDF and Bag-of-Words are the most popular methods of feature extraction in NLP [24].

Feature Selection - The most appropriate feature will be captured using the feature selection. Techniques as TF-IDF score, Chi-square value, Information Gain are the most relevant features for classification.

Classification - Classifications such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision trees (DT), Ensemble considered in previous studies [10]. However, the proposed system will consider the most accurate model that detects the early depression symptoms based on sentiment analysis.

Databases - Kaggle database will use to train the model. Also, raw data used by the study [10] consider in the implementation.

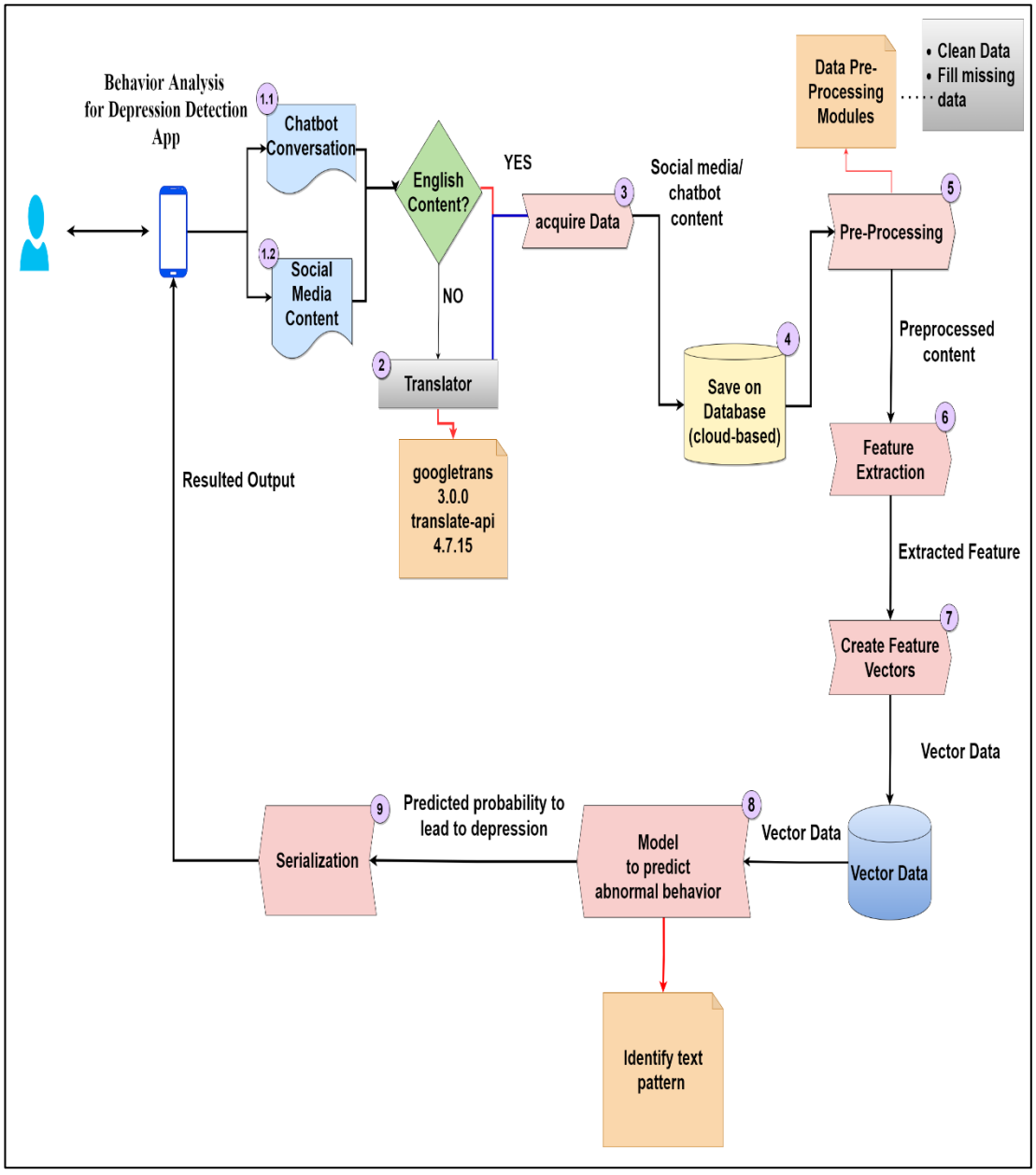


Figure 3. 1: System Diagram of abnormal behavior analysis

3.2 Project Flow

Figure 3.2 describes the process of training the model.

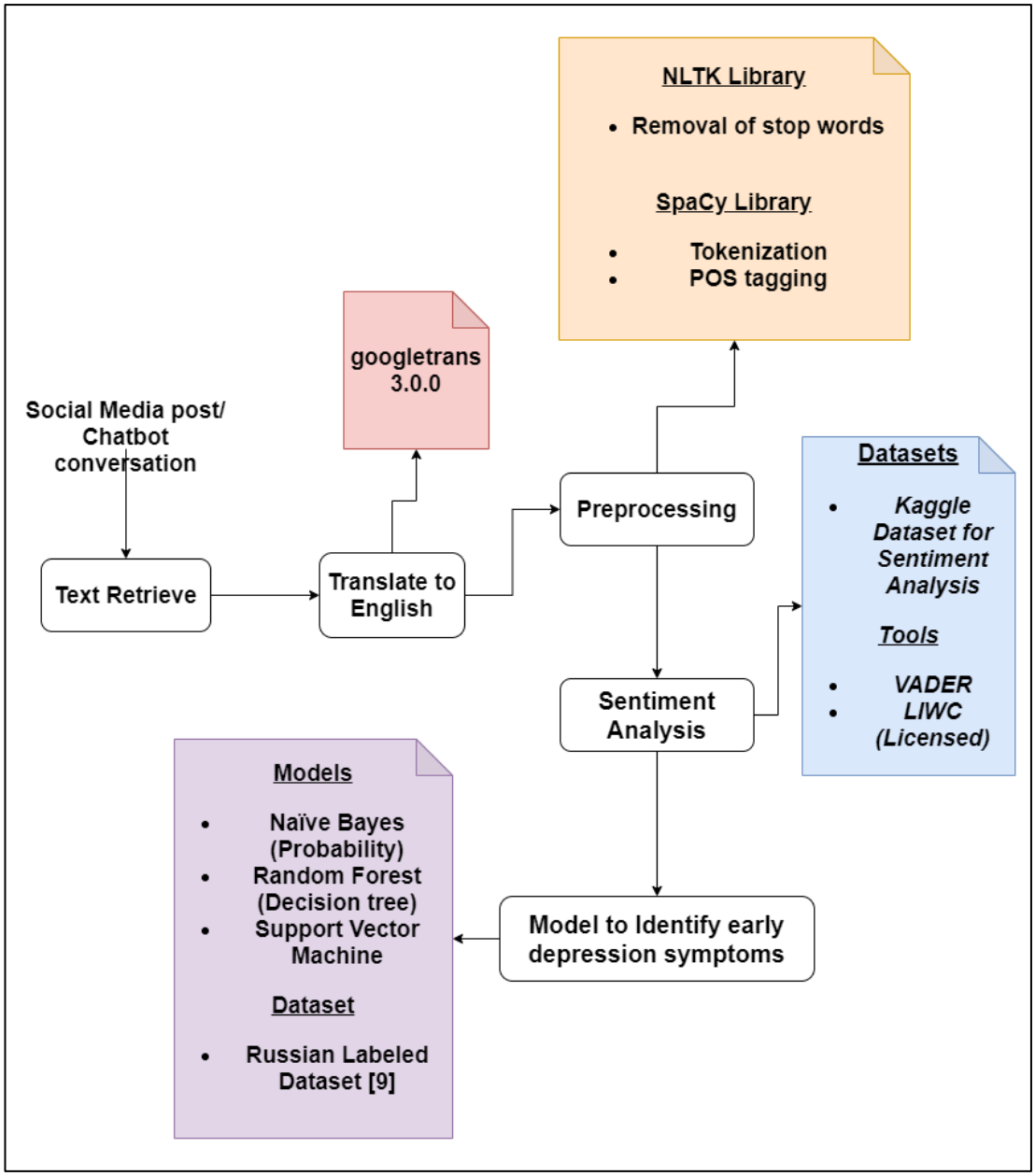


Figure 3. 2: Implementation Process for abnormal social activities analysis model.

Figure 3.3 shows the interface design of the mobile application. The following interfaces include in the proposed application.

Step 01 – The user should complete the registration process to open the application.

Step 02 – Login interface provides the user to login into the system.

Step 03 – User agreement describes the terms and conditions.

Step 04 – Summary displays in the application dashboard.

Step 05 – The user can initiate the chatbot through the application.

Figure 3. 3: Mobile Application Design Flow

3.3 Technologies

The tools, libraries, and programming languages that will use in the proposed system implementation are described as follows.

Table 3. 1: Technologies, tools, libraries for the proposed system.

Technology	**Description**
Anaconda Environment	all the packages required to implement the model will be managed by the Anaconda environment. Packages can be installed, upgraded downgraded easily.
NLTK	Natural Learning Processing Toolkit to build python programs.
Python Programming Language	The model will be implemented using high-level, object-oriented, and generic purpose language, Python.
Flutter	Flutter UI software development kit will use for Frontend development.
Visual Studio Code	Freeware source-code editor used for implementation of the mobile application.
TensorFlow	Open-source library, TensorFlow will use to train the model.
Firebase	Processed Data will be saved in the database, “Firebase”.

3.4 System Testing

To ensure the expected performance from the proposed system, it will undergo various testing phases as unit testing, component testing, integration testing, and system testing.

3.4.1 Model Testing

The model will be tested using the test data to validate the performance of the trained model. Classification accuracy is the most basic way to evaluate the training model [26].

3.4.2 Mobile Application Testing

The proposed mobile application must undergo a testing phase to provide a better user experience. The functionalities should be properly evaluated to uplift the commercial value of the product.

- **User interfaces testing –** The interfaces will be tested to identify the compatibility for various devices. The testing includes the layout of the user interface, performance in various screen resolutions and screen orientations. Furthermore, the elements of the interface will be tested to maximize the commercial value of the product.

- **functionality testing –** The user input is critical since the analysis depends on the user input. The proposed system should provide the intended functionality to acquire data. The functionality to initialize the chatbot and access to social media platforms needs to be implemented accurately.

4. PROJECT REQUIREMENT

The following section is explained the required functionalities for the implementation of the mobile-based application.

4.1 Functional Requirements

- Implement a chatbot to converse with the user.
- Analyze the patterns of social media accessing to identify individuals with abnormal social activities towards depression.
- Analyze an individual's conversation with the chatbot to identify individuals with abnormal social activities towards depression.
- Predict the probability of having early depression symptoms based on analysis of the social media content and conversation with the chatbot.
- Generate the comprehensive report summary of probability on early depressive symptoms.
- Notify the guardian or mental health care provider.

4.2 Non-functional Requirements

- **Privacy** – The system needs to protect privacy since it accesses the social media accounts of the user.
- **Availability** – Users are accessing social media platforms continuously. Therefore, the system needs to monitor social activities regularly to analyze the depressive symptoms.
- **Compatibility**– The users and mental health care providers should be able to use the application using well-known mobile platforms.

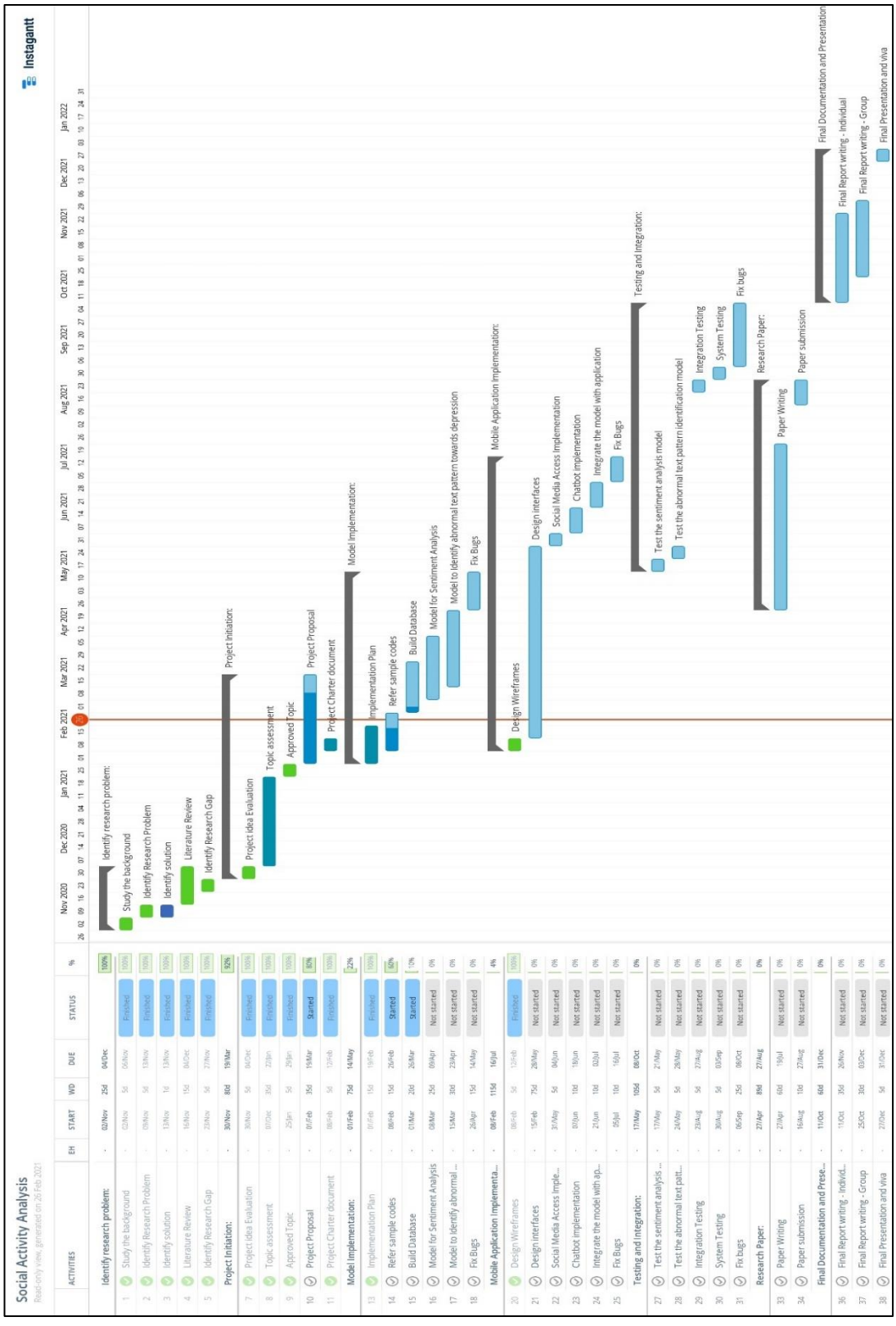


Figure 4. 1: Gantt Chart

The following Figure 4.2 depicts the work breakdown structure.

Figure 4. 2: Work Breakdown Structure

5. COMMERCIALIZATION POTENTIAL

Behavior changes need to monitor for providing effective treatments on mental health. Early identification of depression helps the individuals and the mental health care provider remitting the possible damage to life. However, social stigma hinders individuals for adequate treatments by holding them back from appointing healthcare providers.

The proposed system will accommodate individuals to receive treatment from mental health care providers by notifying the responsible person. The economic value of the proposed mobile application describes as follows.

5.1 Community Value

5.1.1 Value to the public

- Extensive observation of social activities without direct involvement.
- A comprehensive view of daily social activities on social media platforms.
- Notify family and healthcare provider on abnormal social activities.
- Biweekly appointments from the psychiatrist to monitor behavior changes will not be needed with the proposed system.

5.1.2 Value to the Mental Healthcare Provider

- Monitor abnormal social activities of the patients using the proposed system.
- Immediate treatment can be provided to the patients.
- The mental health of the patient will be received with a comprehensive summary.
- Notifications about patients' mental health.

5.2. Scientific Value in Research

- An automated analysis on abnormal social activities towards depression is not yet studied in Sri Lanka.

- The analysis of social media activities to identify early depressive symptoms is not yet considered in the field of study.
- The conversation with the chatbot is not yet analyzed to identify early depressive symptoms.

5.3. Risk and Benefits Assessment

5.3.1. Risk of the research

- The personal details may disclose to the public.
    - Personal information will hide from unauthorized sources.
    - The individual will have the authority to decide the trusted people to share the reports.

5.3.2. Benefits of the research

- The proposed system monitors abnormal social activities of the patients with high accuracy.
- Immediate treatments will be given to patients.
- A comprehensive summary of the patient's mental health will be shared with the Mental Healthcare Provider.

5.4.Confidentiality and privacy

5.4.1. Personal information and Report

- The proposed system will hide the Personal details from the public.
- The extensive Reports on patient's mental health will be shared with trusted sources.

5.5.Budget and Budget Description

The following Table 5. 1 includes the overall estimated cost for project completion. Monthly Revenue estimation explains in Table 5.2.

Table 5. 1: Cost Estimation

Event	**Amount (LKR)**
Appointments for Mental Healthcare Provider	2500.00 (For an appointment)
Internet	500.00 (Per Month)
Total	9500.00 (Approximately)

Table 5. 2: Monthly Revenue Estimation

Users	**Package**	**Price (Monthly)**	**Product Description**
End-Users	Basic	-	• Installation through "Play Store" (Free of Charge). • Summary of daily social activity. • Monitor social activity for two weeks. • Summary report on the probability of depressive features.
	Premium	Rs. 500.00	• Installation through "Play Store" (Free of Charge). • Summary of daily social activity. • Monitor social activity for two weeks. • Monitor social activity for four weeks. • A comprehensive report indicating daily social activity and probability of depressive features.
Mental HealthCare Provider	Premium	Rs. 750.00	• Notifications from patients. • A comprehensive report indicating the details of the depressive symptoms of the patient.

6. DESCRIPTION OF PERSONAL AND FACILITIES

Shalindi Pandithakoralage - Clinical Psychologist

A qualified and experienced Clinical Psychologist, who specialized in Cognitive Behavioral Therapy (CBT), Psychological Assessment, social assistance, statistics, and reaches. She has received her first degree from the Missouri University of Science and Technology. She obtained her MPhil focused in Clinical Psychology from the University of Colombo. The field of focus and experience of her includes counseling for clients who have difficulties with obsessive-compulsive disorder, trauma, anxiety, relationship and marital issues, and depression. She considers both Humanistic and Cognitive-Behavioral principles to maintain a client-focused approach.

The research will be externally supervised by her with experience in the industry and involvement in scientific research studies. She will validate the acquired depressed data by using a gold-standard approach. The research team will be having guidance towards the accuracy analysis from her.

REFERENCES

[1] "Mental health statistics: global and nationwide costs," *Org.uk*, 10-May-2018. [Online]. Available: https://www.mentalhealth.org.uk/statistics/mental-health-statistics-global-and-nationwide-costs. [Accessed: 10-Feb-2021].

[2] M. Marcus, M. Taghi Yasamy, M. van Ommeren, D. Chisholm, and S. Saxena, "A Global Public Health Concern," *World Health Organization (WHO)*. [Online]. Available: https://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf. [Accessed: 20-Feb-2021].

[3] "Depression," *World Health Organization (WHO)*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression. [Accessed: 10-Feb-2021].

[4] Y. S. Balhara, N. Anwar, and P. Kuppili, "Depression and physical noncommunicable diseases: The need for an integrated approach," *WHO South East Asia J. Public Health*, vol. 6, no. 1, p. 12, 2017.

[5] "Depression: let's talk" says WHO, as depression tops list of causes of ill health," *World Health Organization (WHO)*. [Online]. Available: https://www.who.int/news-room/headlines/30-03-2017--depression-let-s-talk-says-who-as-depression-tops-list-of-causes-of-ill-health. [Accessed: 10-Feb-2021].

[6] M. De Choudhury M. Gamon S. Counts and E. Horvitz, "Predicting Depression via Social Media," *ICWSM*, vol. 7, no. 1, Jun. 2013.

[7] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh, and H. Ohsaki, "Recognizing depression from twitter activity," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 2015.

[8] Delahunty, Fionn, Wood, Ian D., & Arcan, Mihael., "First insights on a passive major depressive disorder prediction system with incorporated conversational chatbot," in *26th AIAI Irish Conference on Artificial Intelligence and Cognitive Science (AICS 2018)*, Trinity College Dublin, Dublin, 2018.

[9] C. S. A. Razak, M. A. Zulkarnain, S. H. A. Hamid, N. B. Anuar, M. Z. Jali, and H. Meon, "Tweep: A system development to detect depression in twitter posts," in *Lecture Notes in Electrical Engineering*, Singapore: Springer Singapore, 2020, pp. 543–552.

[10] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health Inf. Sci. Syst.*, vol. 6, no. 1, p. 8, 2018.

[11] H. A. Schwartz, J. Eichstaedt, M. L. Kern, G. Park, M. Sap, D. Stillwell, M. Kosinski and L. Ungar, "Towards assessing changes in degree of depression through Facebook," in *Proceedings of the Workshop on Computational Linguistics*

and Clinical Psychology: From Linguistic Signal to Clinical Reality, 2014, pp. 118–125.

[12] "Happify: Science-based activities and games," *Happify.com*. [Online]. Available: https://my.happify.com/. [Accessed: 09-Feb-2021].

[13] "Happify App Review: What's Offered, Cost, & Who It's Right For," *Choosingtherapy.com*. [Online]. Available: https://www.choosingtherapy.com/happify-app-review/. [Accessed: 09-Feb-2021].

[14] "Wysa: anxiety, depression & sleep therapy chatbot," *Google.com*. [Online]. Available: https://play.google.com/store/apps/details?id=bot.touchkin&hl=en&gl=US. [Accessed: 11-Feb-2021].

[15] A. Beltran, "Wysa: Mental Health Support," *Commonsensemedia.org*, 08-Jul-2019. [Online]. Available: https://www.commonsensemedia.org/app-reviews/wysa-mental-health-support. [Accessed: 11-Feb-2021].

[16] "Mental health chatbot," *Woebothealth.com*, 08-Dec-2020. [Online]. Available: https://woebothealth.com/. [Accessed: 14-Feb-2021].

[17] "Woebot - Your Self-Care Expert," *Apple.com*. [Online]. Available: https://apps.apple.com/us/app/woebot-your-self-care-expert/id1305375832. [Accessed: 14-Feb-2021].

[18] "Woebot: Your Self-Care Expert," *Google.com*. [Online]. Available: https://play.google.com/store/apps/details?id=com.woebot&hl=en&gl=US. [Accessed: 14-Feb-2021].

[19] E. Brodwin, "I spent 2 weeks texting a bot about my anxiety — and found it to be surprisingly helpful," Business Insider, 30-Jan-2018.

[20] "Youper terms of use," *Youper.ai*, 24-Mar-2016. [Online]. Available: https://www.youper.ai/terms-of-use. [Accessed: 15-Feb-2021].

[21] "Youper mental health AI app review," *Onemindpsyberguide.org*, 16-Aug-2018. [Online]. Available: https://onemindpsyberguide.org/apps/youper-app-review/. [Accessed: 15-Feb-2021].

[22] "TechCrunch is now a part of Verizon Media," *techcrunch.com.* [Online]. Available: https://techcrunch.com/2019/06/18/youper-a-chatbot-that-helps-users-navigate-their-emotions-raises-3-million-in-seed-funding/. [Accessed: 15-Feb-2021].

[23] "liwc-text-analysis," *Pypi.org*. [Online]. Available: https://pypi.org/project/liwc-text-analysis/. [Accessed: 17-Feb-2021].

[24] "What is data normalization?," *Bmc.com*. [Online]. Available: https://www.bmc.com/blogs/data-normalization/. [Accessed: 18-Feb-2021].

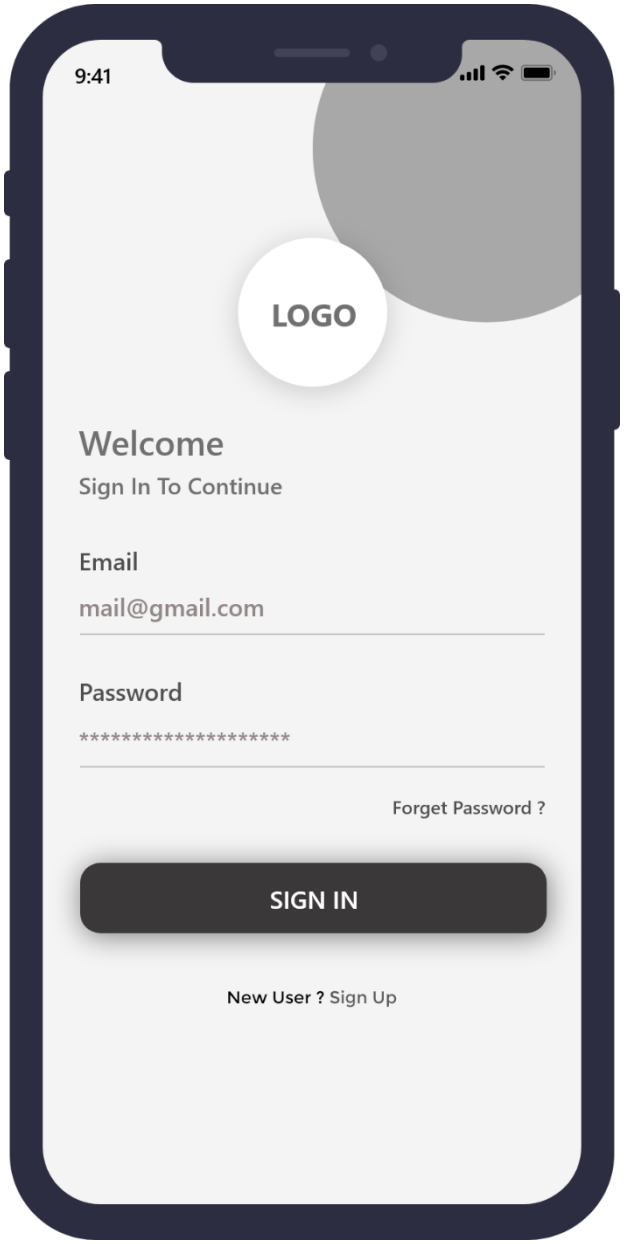
[25] "Feature Extraction Techniques - NLP - GeeksforGeeks," *Geeksforgeeks.org*, 03-Mar-2020. [Online]. Available: https://www.geeksforgeeks.org/feature-extraction-techniques-nlp/. [Accessed: 17-Feb-2021].

[26] M. Billa, “Testers guide for testing machine learning models - analytics Vidhya - medium,” *Analytics Vidhya*, 12-Oct-2019. [Online]. Available: https://medium.com/analytics-vidhya/testers-guide-for-testing-machine-learning-models-e7e5cea81264. [Accessed: 23-Feb-2021].

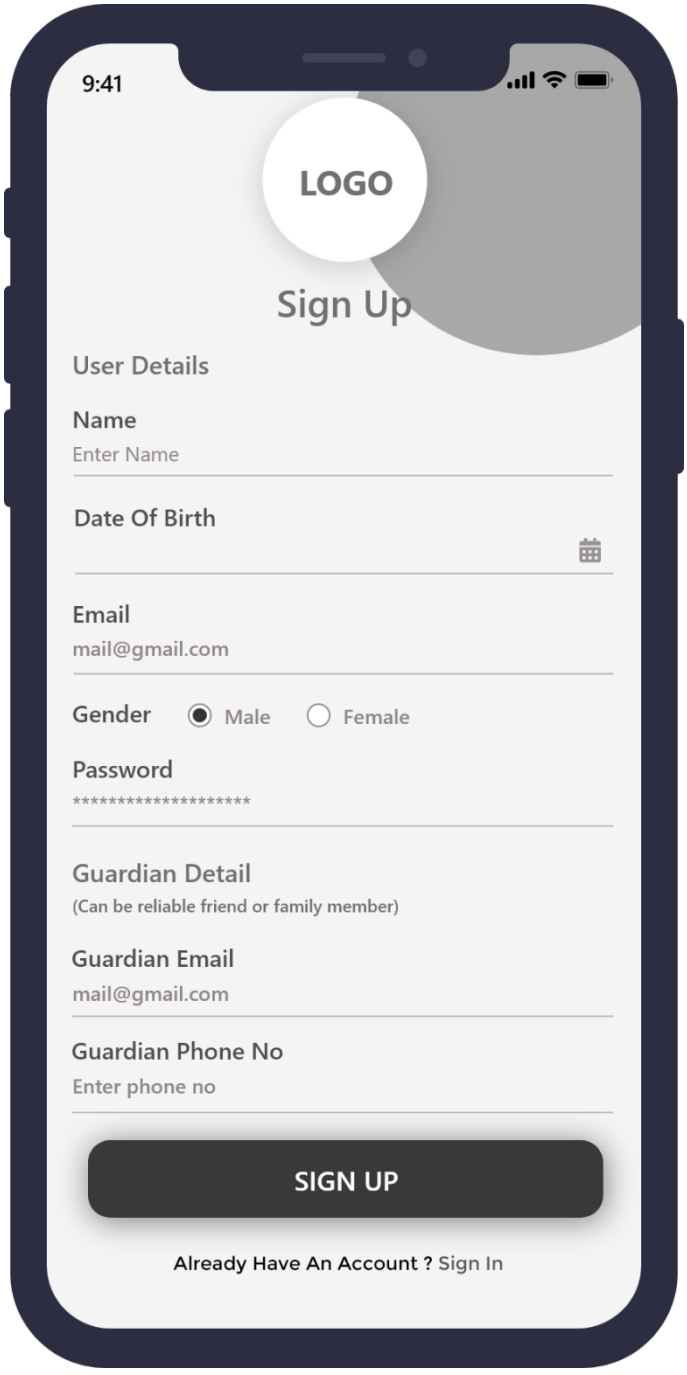
Appendix A- Final Application Design Flow

Application Initialization

Login to the Application

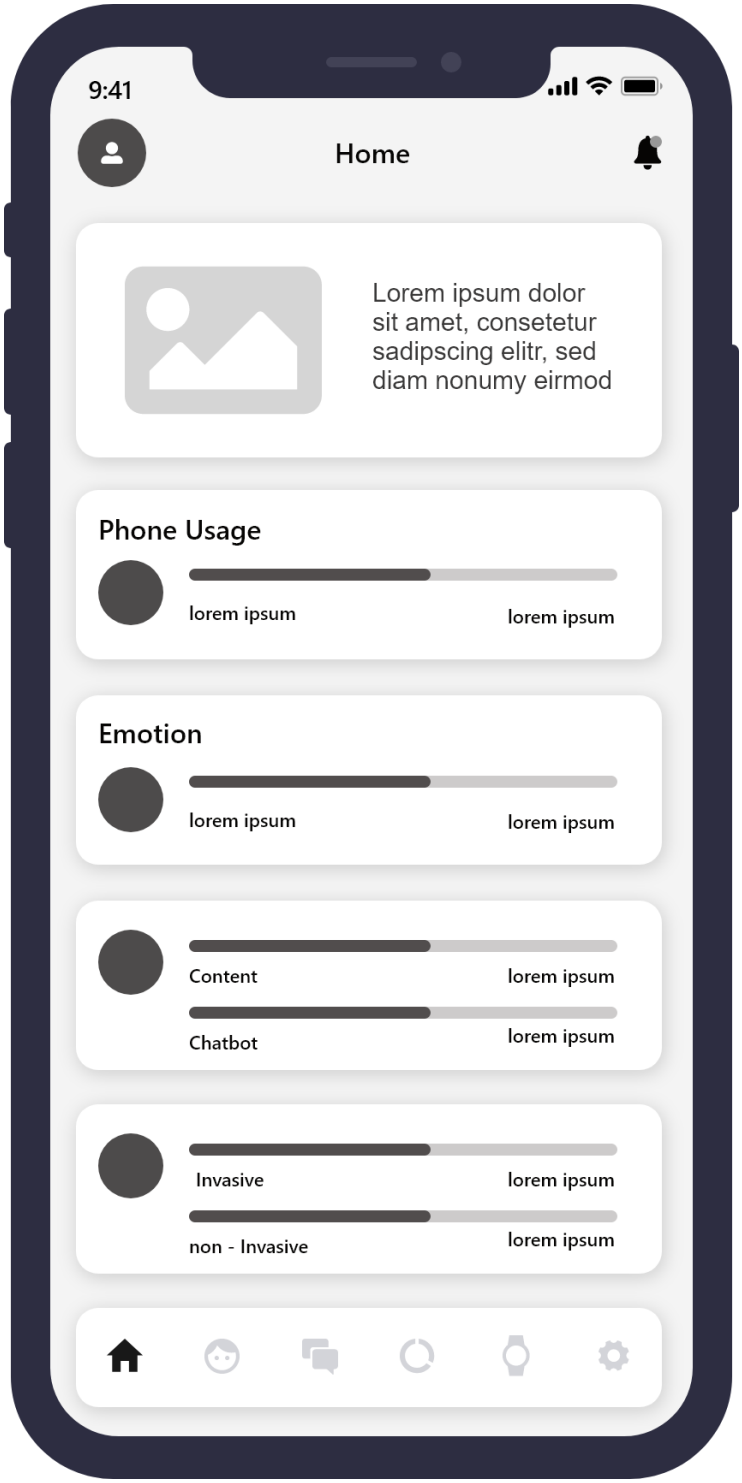


Sign up to the Application

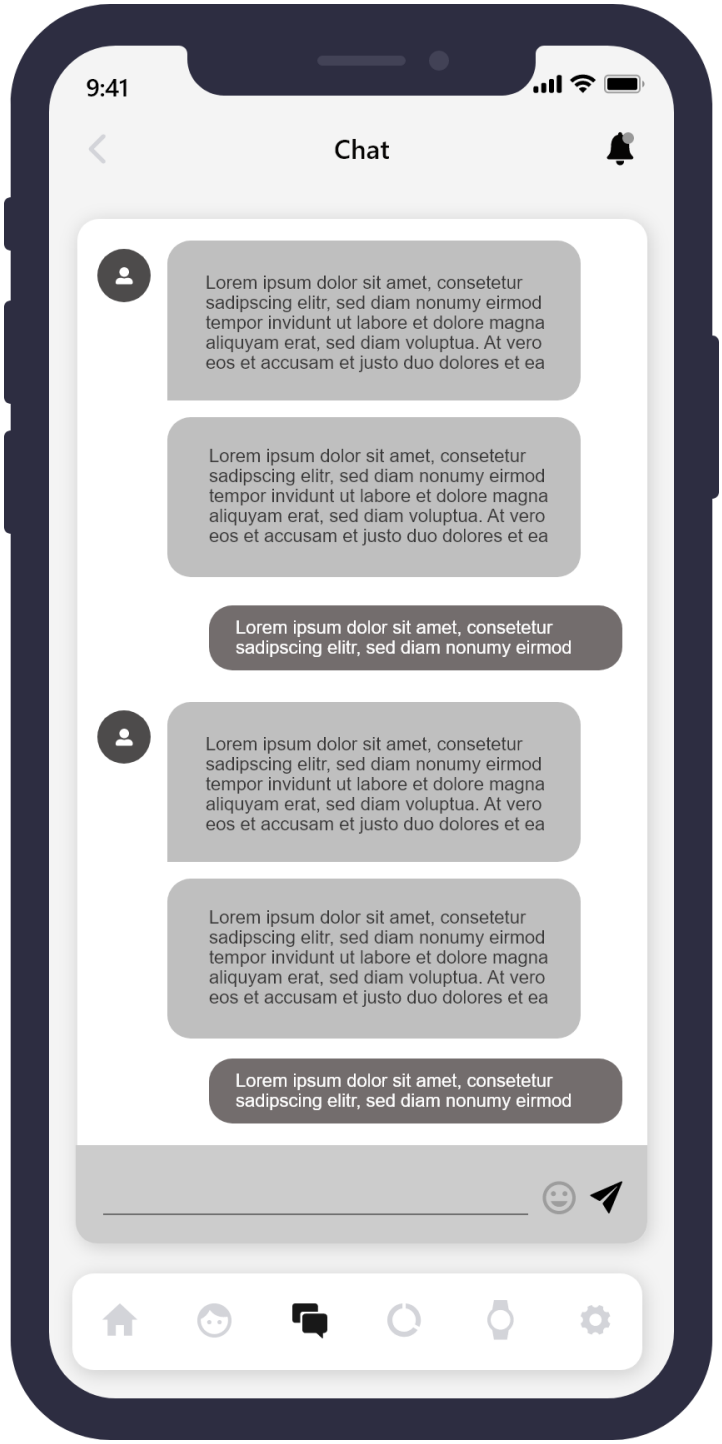


User Agreement

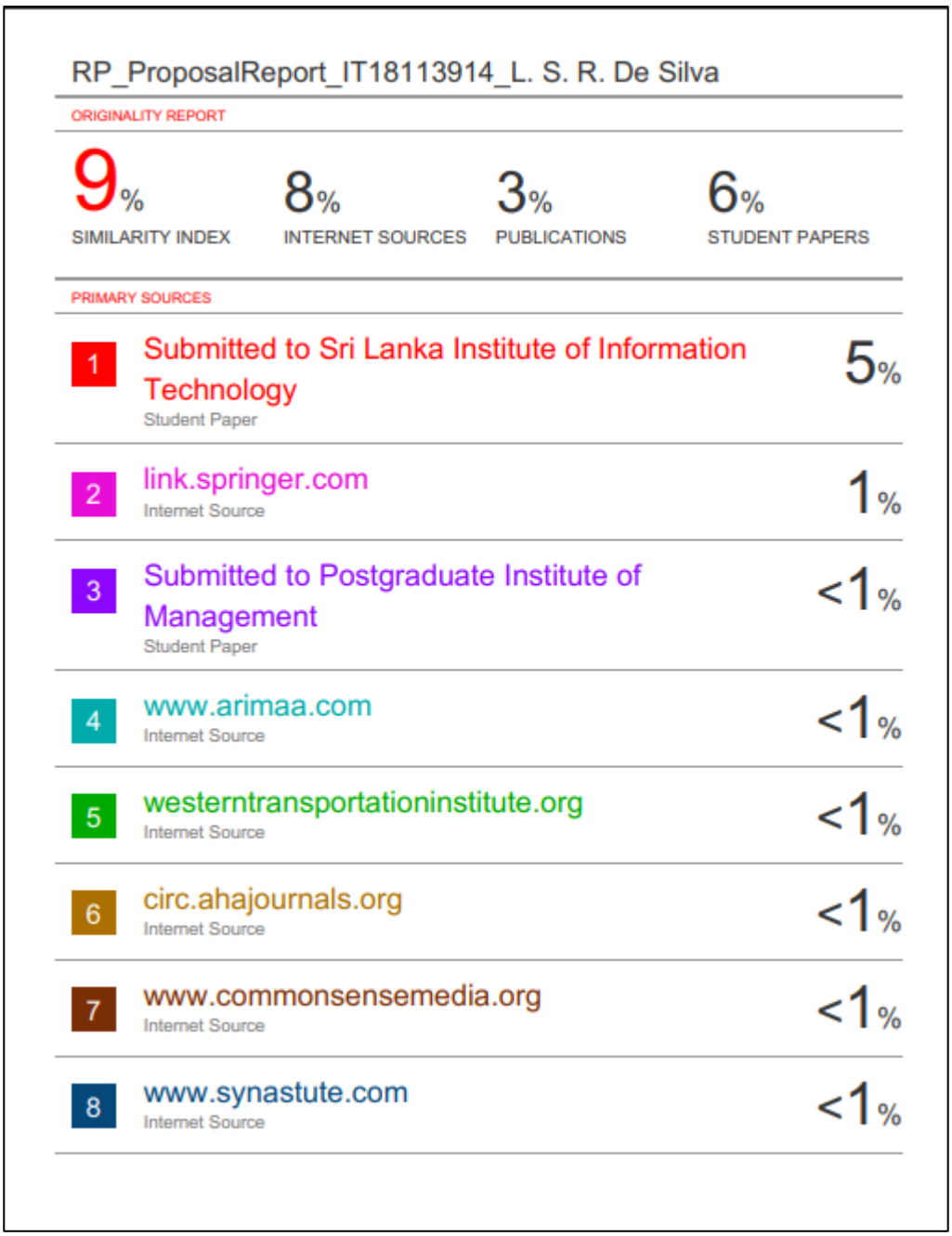
Application Dashboard



ChatBot Initialization through the Application.



Appendix B– Turnitin Plagiarism Report



RP_ProposalReport_IT18113914_L. S. R. De Silva

ORIGINALITY REPORT

9%	8%	3%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Sri Lanka Institute of Information Technology Student Paper	5%
2	link.springer.com Internet Source	1%
3	Submitted to Postgraduate Institute of Management Student Paper	<1%
4	www.arimaa.com Internet Source	<1%
5	westerntransportationinstitute.org Internet Source	<1%
6	circ.ahajournals.org Internet Source	<1%
7	www.commonsensemedia.org Internet Source	<1%
8	www.synastute.com Internet Source	<1%

9	nam.edu Internet Source	<1%
10	onemindpsyberguide.org Internet Source	<1%
11	Shoumen Datta. "Aptamers for Detection and Diagnostics (ADD) is a proposed mobile app acquiring optical data from conjugated quantum nanodots to identify molecules indicating presence of SARS-CoV-2 virus: Why public health and healthcare need smartphone sensors as a platform for early detection and prevention", American Chemical Society (ACS), 2021 Publication	<1%
12	export.arxiv.org Internet Source	<1%
13	www.isca-speech.org Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		