

# **Software Requirements Specification**

**for**

## **AI-based Social Media Data Analysis for Mental Health Evaluation**

**Version 1.0 approved**

**Prepared by  
Mohit Kunder  
Benita Rego  
Nolita Rego**

**Fr. Conceicao Rodrigues College of Engineering**

**December 2020**

# Table of Contents

<b>Revision History</b>	<b>3</b>
<b>1. Introduction</b>	<b>4</b>
1.1 Purpose	4
1.2 Document Conventions	4
1.3 Intended Audience and Reading Suggestions	5
1.4 Product Scope	5
1.5 References	6
<b>2. Overall Description</b>	<b>7</b>
2.1 Product Perspective	7
2.2 Product Functions	7
2.3 User Classes and Characteristics	7
2.4 Operating Environment	7
2.5 Design and Implementation Constraints	8
2.6 Assumptions and Dependencies	8
<b>3. External Interface Requirements</b>	<b>9</b>
3.1 User Interfaces	9
3.2 Hardware Interfaces	9
3.3 Software Interfaces	9
3.4 Communications Interfaces	9
<b>4. System Features</b>	<b>10</b>
4.1 User Login	10
4.2 Evaluation of the Twitter data	10
4.3 Feature Extraction	10
4.4 Result Visualization	10
4.5 Consulting recommendations	10
4.6 Online Consultation	11
<b>5. Other Nonfunctional Requirements</b>	<b>12</b>
5.1 Performance Requirements	12
5.2 Safety Requirements	12
5.3 Security Requirements	12
5.4 Software Quality Attributes	12
5.5 Business Rules	12
<b>Appendix A: Glossary</b>	<b>13</b>

## Revision History

Name	Date	Reason for Changes	Version

# 1. Introduction

## 1.1 Purpose

The aim of this project is brought in light of the growing mental health issues among people due to coronavirus outbreak and other issues. In recent years, there has been a continuous increase in popularity of social media platforms such as Twitter, Facebook, Instagram, etc that became an integral part of people's life. As these social media data timely reflect users' real-life states and emotions in a timely manner, it offers new opportunities for representing, measuring, modeling, and mining users behavior patterns through the large-scale social networks, and such social information can find its theoretical basis in psychology research. For example, it is found that stressed users are more likely to be socially less active, and more recently, there have been research efforts on harnessing social media data for developing mental and physical healthcare tools. This close relationship between social media platforms and their users has made this proposed system to reflect the users' personal life on many levels. The system provides a background on depression, use of social media platforms for prediction and machine learning algorithms. The system monitors the social media activities of each person and predicts their mental health factors such as depression, anxiety, stress, etc. Time and frequency of tweet is analyzed for irregularities and opinion polarity analytics is done to find inconsistencies in posting behaviour. This system will use real time online social media data by investigating the correlations between users' mental health and the content they post on social media. This system could be used by mentors such as teachers, doctors, etc to acquire a weekly analysis of their person's stress levels and thereafter help in providing consultation accordingly.

## 1.2 Document Conventions

This document uses Font style Times New Roman and Font size 12.

Other document conventions used in the SRS :

Admin	The person who is responsible for the upkeep, configuration and reliable operation of the system.
CNN	Convolutional neural network is a class of deep neural networks, most commonly applied to analyzing visual imagery.
Tweepy	Tweepy is an open source Python package that gives you a very convenient way to access the Twitter API with Python.
TensorFlow	An end-to-end open-source platform for machine learning.
Keras	A high-level neural networks API, written in Python and capable of running on top of TensorFlow

NumPy	Library for adding support for large, multi-dimensional arrays and matrices, along with high-level mathematical functions to operate on these arrays
Pandas	Pandas is a software library written for the Python programming language for data manipulation and analysis
Scikit-learn	Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library for the Python programming language.
NLP	Natural language processing is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data.
Flask	Flask is a popular Python web framework, meaning it is a third-party Python library used for developing web applications.

### 1.3 Intended Audience and Reading Suggestions

This project is a prototype for detecting deteriorating mental health of an individual using his social media feed. It is beneficial for users living away from their family to keep a track of their mental health. It has been implemented under the guidance of college professors. It can also be used as a reference by students for research purposes.

### 1.4 Product Scope

Statistics show that 1 in every 5 individuals suffers from some form of mental health illness symptoms. 50% of mental health conditions begin by age 14 and 75% of mental health conditions develop by age 24. Mental illness can be triggered by multiple factors. Complex interactions between mind, body, and environment result in psychological disorders. Some of the factors are long term acute stress, biological factors, drug abuse and overdose, cognitive behavior like constant negative thoughts, low on energy, etc, social problems like financial problems, breakdowns, isolation, etc. While it isn't possible to accurately pinpoint the exact cause of a person's mental illness without a professional psychologist, it is still possible to detect decline in a person's mental state using the data from his social media usage. Detecting the first stages of depression is equally important when compared to detecting the exact causes of depression and that's what the system does by giving the user an overview of

his own mental early detection can enable the individual to become aware of his own mental state and take the necessary steps in order to improve his condition.

## 1.5 References

- "Convolutional Neural Network - Wikipedia." [wikipedia.org/wiki/Convolutional\\_neural\\_network](https://wikipedia.org/wiki/Convolutional_neural_network). (Accessed 17 July. 2020)
- "Social Computing - Wikipedia." [wikipedia.org/wiki/Social\\_computing](https://wikipedia.org/wiki/Social_computing). (Accessed 17 July. 2020)
- "Anxious Depression Prediction in Real-time Social Data" - International Conference on Advanced Engineering, Science, Management and Technology -2019 (ICAESNT19), March 2019
- "Sentiment Analysis/Text Classification Using CNN(Convolutional Neural Network) - Towards Data Science (Accessed 30 July. 2020)
- "Deep Learning for Depression Detection of Twitter Users" - University of Ottawa, ON K1N 6N5 Canada, January 2018
- "Detecting Stress Based on Social Interactions in Social Networks" - Journal of Latex Class Files, Volume-13, No-9, September 2014
- "Machine Learning for Depression Diagnosis using Twitter data" - International Journal of Computer Engineering in Research Trends: Volume-5, Issue-2, 2018 Regular Edition, E-ISSN: 2349-7084, March 2017
- "Stress Detection Methodology based on Social Media Network: A Proposed Design" - International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-3, January 2020

## **2. Overall Description**

### **2.1 Product Perspective**

Using a trained ML model the tweets are classified into positive, negative and neutral and accordingly. The proposed app with the Twitter API integrated into it can access the user's public tweets. Once the tweets are classified and an appropriate analysis is made based on them and sent to the user in a readable format. In case the user's mental state is consistently declining then the system will recommend nearby counsellors and mental health professionals to the user.

### **2.2 Product Functions**

User:

- A sign-in screen is displayed. The user should have a Twitter handle. Details such as the Twitter username is used to login to his account for further needs.
- After the user enters his Twitter handle (unique username) he is redirected to the authorization page where Twitter API authorizes the user's data and matches the username that was entered by the user. If the user has a Twitter account, it satisfies the authentication and Twitter gives appropriate permissions and the user reaches the home page.
- The user will receive periodic reports of what the system perceives to be his current mental state.
- If the mental state is consistently declining then the user receives recommendations of mental health professionals based on his current location.

### **2.3 User Classes and Characteristics**

As the model is made solely for the user, it is a single party app where the login using his Twitter handle is completed by that user using the Twitter OAuth 2.0 feature and the system's Login feature. Rest of the functionalities are done by the system i.e weekly analysis. The model will display analysis based on the data present in the user's Twitter handle. The user's data should be stored in the database for evaluating the mental state for further weeks.

### **2.4 Operating Environment**

The software will operate with the following software components and applications:

The software being developed will be running under Windows 7, 8 and 10 and/or Android operating systems. The hardware that will be running these programs will follow the specifications that appear in this document in section 3. To briefly state, Android Smartphone with Android version 6+, GPU, High Speed connection to database, and a strong Internet access.

## **2.5 Design and Implementation Constraints**

Memory: For designing purposes, Google drive with a memory of about 15GB will be used for storing and loading the dataset. Google Colaboratory provides connection to Python 3 Google Compute Engine Backend with 12 GB RAM and disk space of about 360 GB.

Language requirements: Software will be available only in the English language.

Programming Constraints : Python 3 programming language, TensorFlow 1.15.0, Flask backend

## **2.6 Assumptions and Dependencies**

Since, Google Colaboratory has been used, all the required dependencies like GPU, CNN, TensorFlow, Keras library, Sci-kit learn library, Numpy, Pandas are installed. Full working of the system will depend on the network. So the status of the internet connection must be maintained at all times.

It is assumed that the app interface will be compatible with every android mobile. It is also assumed that the dataset obtained is 100% accurately labelled and the information in the dataset is correct.



## **3. External Interface Requirements**

### **3.1 User Interfaces**

- Home Page : This is the initial page of the app. This page has option of login to be done with Twitter credentials and about us displaying details of the app, etc
- Login Page: The user must enter his Twitter username that matches with the Twitter database during sign in.
- AboutUs Page: This page consists of various details of the system.
- Dashboard Page: Once the user has logged in, he reaches his dashboard page which consists of analysis of the user's data.

### **3.2 Hardware Interfaces**

- A phone: Android OS
- High Speed connection to Database
- GPU (Google Colab)
- Strong Internet Access

### **3.3 Software Interfaces**

- CNN
- Tweepy
- NLP
- Tensorflow
- Keras
- Numpy
- Pandas
- Scikit-learn
- REST API
- Python
- Flask
- Postman

### **3.4 Communications Interfaces**

The system requires active internet connection to fetch data from the user's Twitter credentials using Twitter OAuth 2.0 and run the neural network model so as to obtain results. The Internet is also required to install plugins and update components like API's, modules, etc.

## **4. System Features**

This section demonstrates the system's most prominent features and explains how they can be used and the results they will give back to the user.

### **4.1 User Login**

- This feature allows the Twitter user to login using their credentials.
- Once login is successful, the user is directed to a page where they can check for the analysis generated by the model.
- Login is only successful on giving proper credentials.

### **4.2 Evaluation of the Twitter data**

- This is a major feature of this application where the user gets the result of the analysis done using their Twitter data and whether they are in a bad mental state or not.
- After the user data is fetched from the Twitter API, the application sends it to the saved model of the convolutional neural network and the model processes this data.
- The end result of this functionality being the data fetched from the database is classified into three different classes of factors.

### **4.3 Feature Extraction**

- This is a feature of the application where the user gets to know what state of mind he is in after classification of the data done by evaluation.
- Once the data is fetched and the result is displayed whether the user is in any mental state or not and the class of it, the user is then provided with an option to consult a medical expert by searching assistance from his nearby location if he is detected with negative tweets.
- In the medical consultation page, there will be a list of medical experts near him which he can contact if he wishes to.
- Feature extraction is helpful for immediate consulting recommendation.

### **4.4 Result Visualization**

- This feature allows the user to check the analysis in a graphical format which is done on a weekly basis for easy understanding and quick analysis.
- After the negative tweets are detected and evaluated, a graphical representation of the user's tweets is displayed, depending on the polarity given to each class by the model.

### **4.5 Consulting recommendations**

- This feature is basically a static page that recommends the medical experts to be consulted depending on his will.

- This is just a recommendation feature and not necessarily be taken for assistance or medications.

#### **4.6 Online Consultation**

- This is an additional feature of this application which allows online consultations of the medical experts located nearby.
- This feature is enabled using an API.
- The online consultation though entirely depends on the availability of the doctor.

## **5. Other Nonfunctional Requirements**

### **5.1 Performance Requirements**

- The application will run if the system has reliable internet connection. The application can run on mobile phones.
- The results are generated within minutes and does not take a lot of time thus making response time appropriate and quick for the users.
- Response to user errors and undesirable situations have been taken care of to ensure that the application runs without any uncertainty.
- Performance of the app also depends on the kind of data fetched from the user's Twitter account.

### **5.2 Safety Requirements**

The analytical application is not a substitute to any kind of consultation further provided by the app on negative results. To ensure that the product delivers satisfactory results, the developer team updates the database regularly to avoid incorrect results.

### **5.3 Security Requirements**

The application uses Twitter OAuth 2.0 Authentication which is accessed to get the user data, the user should have a Twitter account and can use it without any additional privileges.

### **5.4 Software Quality Attributes**

The application provides users with both simple and complex features. Due to its well designed and easy to use interface, it can be used by many Twitter users. However the users must have a Twitter account to further know how the application functions and interpret the analysis although sufficient help will be provided by the app.

### **5.5 Business Rules**

For any user to use this application they require to have a Twitter account with them. In case, the user does not have a Twitter account, they need to create one with sufficient tweets in them for better analysis although the results will not be as good as the results obtained from a Twitter account with adequate tweets.

## **Appendix A: Glossary**

OAuth - Open Authorization

CNN - Convolutional Neural Network

NLP - Natural Language Processing

GPU - Graphics Processing Unit

REST API - Representational State Transfer Application Programming Interface