AI-Based Social Media Data Analysis for Mental Health Evaluation

A project report submitted in partial fulfilment of
the requirements for the degree of
Bachelor of Engineering

By
Mohit Kunder (8344)
Benita Rego (8362)
Nolita Rego (8363)

Under the guidance of
Prof. Sunil Chaudhari

DEPARTMENT OF COMPUTER ENGINEERING
FR. CONCEICAO RODRIGUES COLLEGE OF ENGINEERING
FR. AGNEL ASHRAM, BANDRA (W),
MUMBAI - 400 050.

UNIVERSITY OF MUMBAI
(2020 – 2021)

FR. CONCEICAO RODRIGUES COLLEGE OF ENGINEERING

FR. AGNEL ASHRAM, BANDRA (W),
MUMBAI - 400 050.

<u>CERTIFICATE</u>

This is to certify that the following students working on the project "**AI-based Social Media Data Analysis for Mental Health Evaluation**" have satisfactorily completed the requirements of the project in partial fulfillment of the course B.E in Computer Engineering of the University of Mumbai during academic year 2020-2021 under the guidance of Prof. Sunil Chaudhari.

Submitted By: **Mohit Kunder (8344)**
Benita Rego (8362)
Nolita Rego (8363)

Prof. Sunil Chaudhari	**Dr. B.S. Daga**
Guide	**Head of the Department**

Principal

CERTIFICATE

This is to certify that the project synopsis entitled **"AI-based Social Media Data Analysis for Mental Health Evaluation"** submitted by the following students is found to be satisfactory and the report has been approved as it satisfies the academic requirements in respect of Major Project - I work prescribed for the course.

Mohit Kunder (8344)
Benita Rego (8362)
Nolita Rego (8363)

Internal Examiner	**External Examiner**
(Signature)	(Signature)
Name:	Name:
Date:	Date:

Seal of the Institute

DECLARATION OF THE STUDENT

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources.

We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea / data / fact / source in my submission.

We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Signature of the student **Date:**

(Mohit Kunder)
(8344)

Signature of the student **Date:**

(Benita Rego)
(8362)

Signature of the student **Date:**

(Nolita Rego)
(8363)

Abstract

The abstract is proposed in light of the growing mental health issues among people due to COVID-19 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. 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. 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.

Table of Contents

1. INTRODUCTION

Mental well-being and social media have been closely related domains of study. Psychological stress is becoming a threat to people's health nowadays. With the rapid pace of life, more and more people are feeling stressed. Mental illnesses, such as depression, are highly prevalent and have been shown to impact an individual's physical health. Thus, there is significant importance to detect stress before it turns into severe problems. Traditional psychological stress detection is mainly based on face-to face interviews, self-report questionnaires or wearable sensors. However, traditional methods are actually reactive, which are usually labor-consuming, time-costing and hysteretic. The rise of social media is changing people's lives, as well as research in healthcare and wellness.With the popularity of social media, people are used to sharing their daily activities and interacting with friends on social media platforms, making it feasible to leverage online social network data for stress detection. We find that a user's stress state is closely related to that of his/her friends in social media, and we employ a large-scale dataset from real-world social platforms to systematically study the correlation of users' stress states and social interactions.

Recently, artificial intelligence methods have been introduced to assist mental health providers, including psychiatrists and psychologists, for decision-making based on patients' historical data (eg. medical records, behavioral data, social media usage, etc) Deep learning as one of the most recent generations of AI technologies, has demonstrated superior performance in many real-world applications ranging from computer vision to healthcare. With the development of social networks like Twitter, Instagram, Facebook more and more people are willing to share their daily events and moods, and interact with friends through the social networks. 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. For example, proposed to leverage Twitter data for real-time disease surveillance, while trying to bridge the vocabulary gaps between health seekers and providers using the community generated health data. Another reason for considering social interactions in stress detection is based on our empirical findings on a large-scale dataset crawled from Sina Weibo that the social structures of stressed users are less connected and thus less complicated than those of non-stressed users.

Just like emotional instabilities, social media is omnipresent, making it easier for people to share their lives on a platform where everyone can view it. Although it is easier to reach out to people, it is difficult to accurately analyse a person's state of mind. But, it may be possible to detect their mental health using their social network data. For this, real-world social media data should be analyzed. Along with detection, the person will also require appropriate medical assistance to help him get through their declining mental state.

1.1 AIM

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 SCOPE

- Early stages detection of negative factors like depression and stress through live streaming data.
- Social media content assessment would help the user to get a clear picture, creating awareness within him about his psychological changes.
- Providing medical assistance by recommending health experts would help the user overcome from the negative changes he is suffering from.

2. LITERATURE REVIEW

Researchers have been probing to find novel methods for mental health analysis. Surveys have found that social computing is one of the relevant ways to examine a person's mental state.

Various models are available for analysis of social media contents that depend on certain variables like posts, tweets, likes, comments on several social media platforms like Instagram, Twitter, etc. These models predict based on the live data that is collected, providing appropriate results. Furthermore, there are many algorithms used for this purpose:

- CNN (84.2%)
- Support vector machines (81.71%)
- Naïve Bayes (77.42%)
- Random Forest (73.8%)
- K-Nearest Neighbours (55.55%)

Numerous studies explored means to make predictions out of Twitter data. But, only a few successfully implemented machine learning algorithms to detect depression in Twitter users.

A. Data Collection and Sampling

Different studies adopted different strategies for acquiring the necessary Twitter data. Most of them used Twitter API to fetch public tweets. However, attributes used differed as per strategy adopted. Some studies directly utilised the dataset generated for CLPsych 2015 Shared Tasks. The rest either conducted surveys to find suitably depressed as well as non-depressed individuals and fetched their public Twitter activity for specified time duration with their permission; or directly mined all public tweets in English language containing either word "depression" or some suitable conducting surveys possess another challenge finding subjects. Finding prospecting subjects directly or online forums may be suitable only for smaller studies. De Choudhary et al. used Amazon's Machine Turk interface to conduct large-scale studies on crowd workers. They asked the crowd workers to take standardised clinical depression surveys along with sharing their depression history, demographics and username of their public Twitter profile. The similar approach was later adopted by Reece et al. They discarded data samples where the crowd worker spent insufficient time in completing the survey. Also, auxiliary screening tests were included in the review to filter out those entries whose depression scores didn't correlate much across the scales.

B. Choice of Depression Scale

Most of the studies used the CES-D Scale for estimating the degree of depression in the subjects. CES-D is a 20-questions long questionnaire designed to measure the extent of depression in the general population. The scale ranges from zero to sixty. Depending upon the CES-D score, the likelihood of depression may be: low (0-15), mild to moderate (16-22) or high (23-60). Park et al. chose 22 as the threshold CES-D score to improve specificity and false-positive diagnosis. The same threshold was followed by the consecutive studies. Studies even used CES-D along with additional auxiliary screening like Beck Depression Inventory

(BDI) and Kellner Symptom Questionnaire. The individuals whose depression scores didn't correlate across the main CES-D scale and its auxiliary scales were filtered out.

C. Feature Extraction

Most of the studies came up with their unique approach to feature extraction. Here, Park et al. in 2012 performed sentiment analysis on tweet text using Linguistic Inquiry Word Count (LIWC). LIWC is a text analysis program that categorises words into multiple psychologically meaningful categories and sub-categories. It returns scores based on the number of the word belonging to its categories and subcategories. Out of the obtained scores, they eliminated categories with very high multicollinearity by examining the bivariate correlation between independent variables and condition numbers.

D. Machine Learning Models for Depression Identification

Among all studies, use of classifiers to categorise the user as either "depressed" or "not depressed" seems to be a common practice. However, some early studies attempted to predict depression scores using multiple regression.

For Supervised learning-based classifiers, the specified number of tweets selected randomly formed the training and testing datasets. Training dataset was mostly annotated manually. To avoid overfitting, De Choudhary et al. implemented Principal Component Analysis (PCA). In their study, Nadeem et al. evaluated the performance of four classifiers by implementing a Decision Tree, a Linear Support Vector Classifier, a Logistic Regression based classifier, Convolutional Neural Networks and a Naïve Bayes classifier. Reeve et al. observed that Convolutional Neural Networks worked the best with their data.

E. Direction for future studies

The major scope for future development lies in finding the novel technique for automated large-scale implementation of machine learning in the diagnosis of Depression. It shall be great to automate this and generate reports regularly. That may help to check the spread of Depression. As techniques of Natural Language Processing are new and still in the development stage, we may expect the algorithms to get better in the future. We may also want the accuracy of these algorithms to improve further to be reliably used with large-scale data. As against the two categories of classification in the current algorithms, it shall be good to have an additional category for those vulnerable to the onset of depression shortly, or for example those with a CES-D score between 16-22. The sentiment analytics programs used currently and in previous studies face issues due to the informal language used in social media posts (example: "tireddddd"). These issues significantly lower the performance of NLP tools and techniques developed for use in formal language. We may hope to see new NLP algorithms with the ability to make sense out of such informal text being used for depression diagnosis on Twitter data.

3. PROPOSED SYSTEM

3.1 Problem Statement Analysis

The motive behind this project is to build a system for analysis and detection of the mental state of a person. The system contains:

- Analysis of a person's state of mind through their social media contents and providing resources to help said person find appropriate assistance.
- Periodic Statistics and Analysis
- Alert the user if the stress level exceeds a certain limit and recommending appropriate assistance
- Displaying locations of nearby counselling firms when required

3.2 Design and Methodology of Proposed System

3.2.1 Architecture Diagram of Complete System

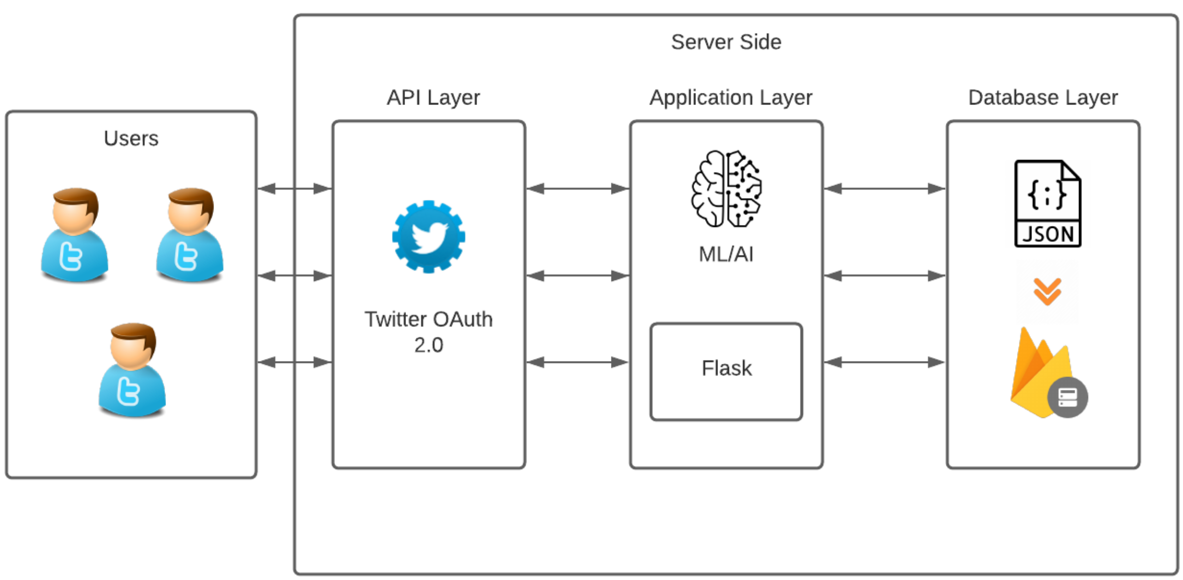

Figure 3.2.1. Architecture Diagram of Complete System

3.2.2 Architecture Diagram for Tweets Classification

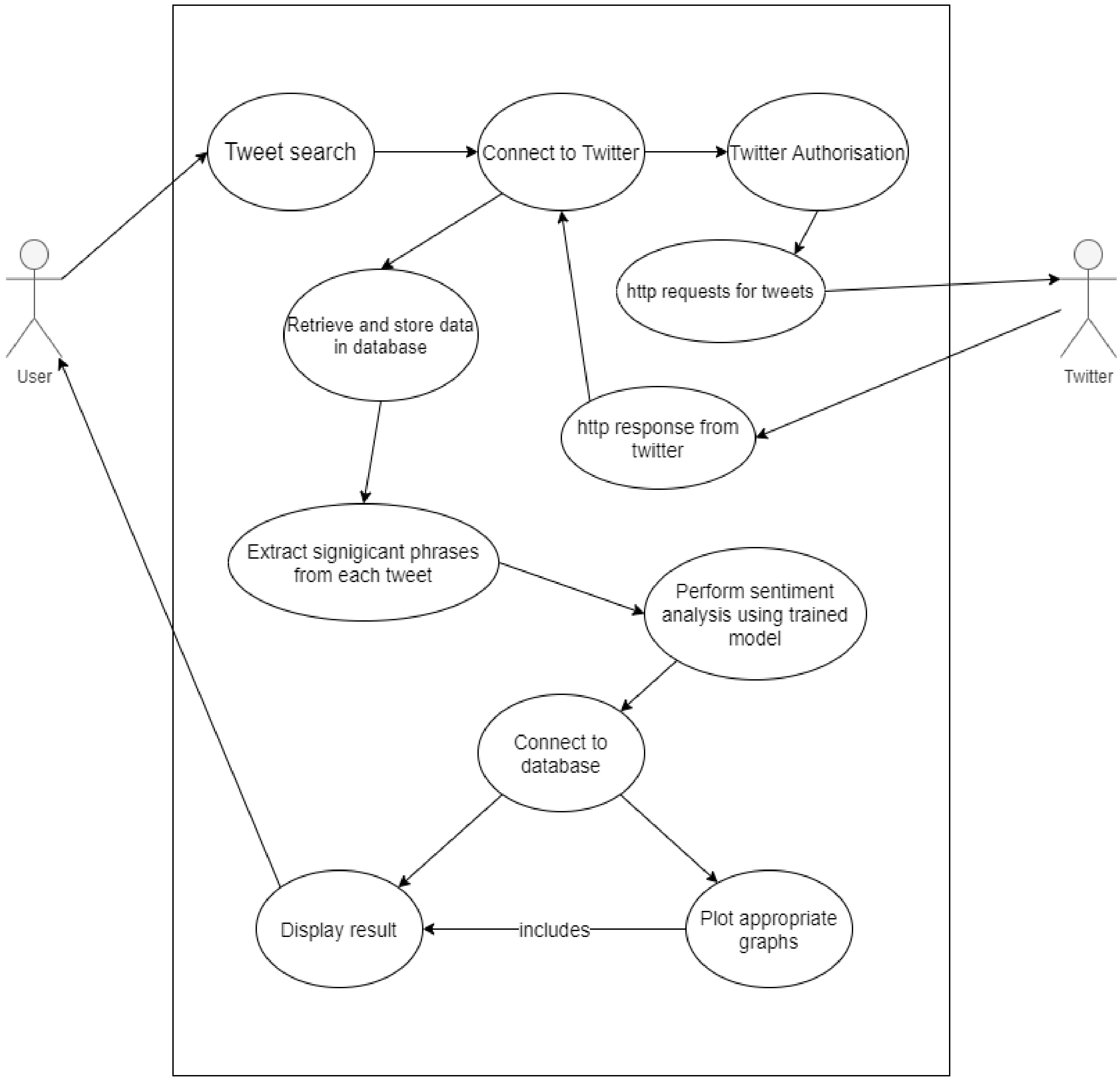

Figure 3.2.2(a) Architecture Diagram of Tweets Extraction and Classification

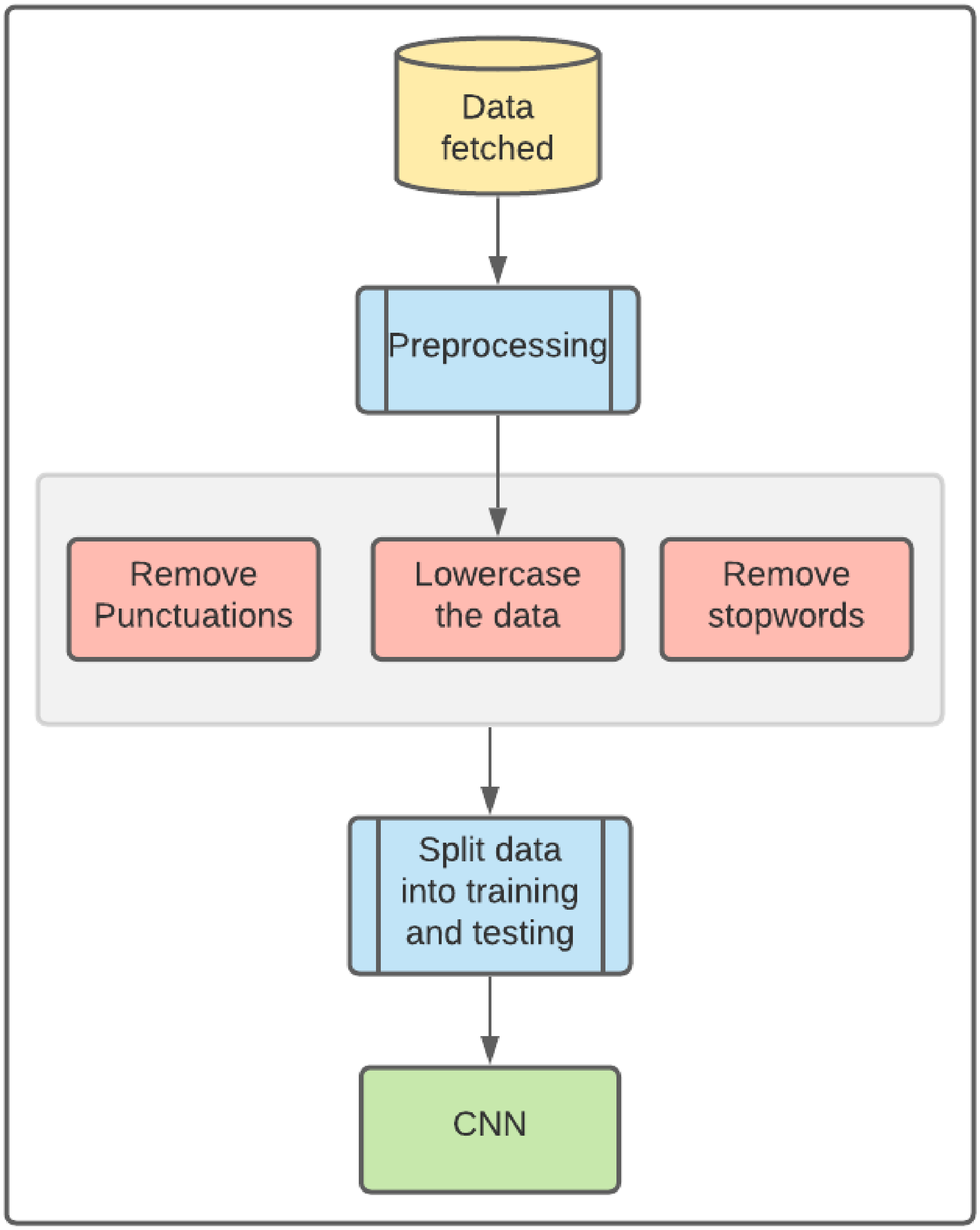

Figure 3.2.2(b) Architecture Diagram of Tweets Extraction and Classification

3.2.3 Flow Diagram of the Mobile Application

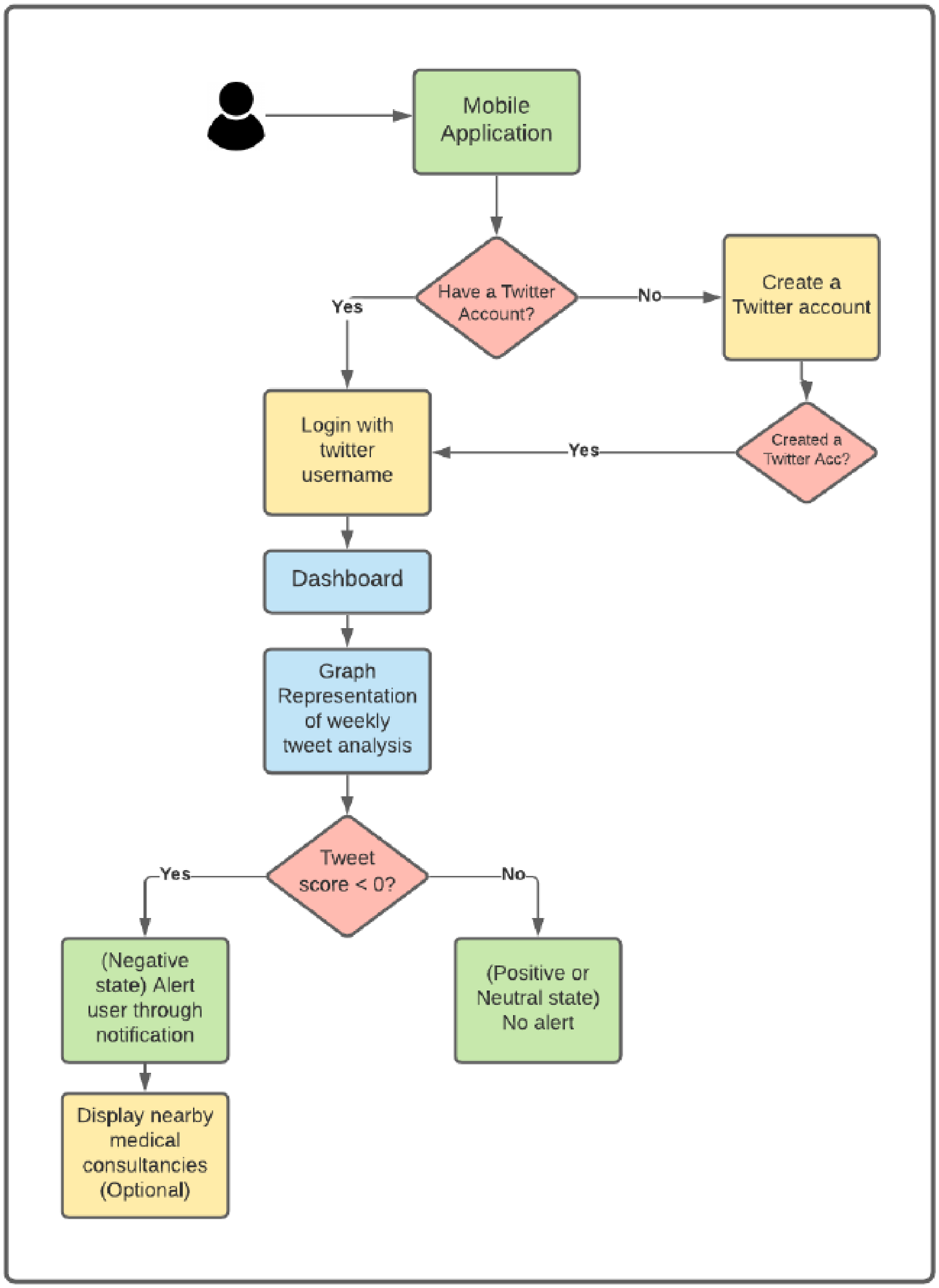

Figure 3.2.3 Architecture Diagram of the Mobile Application

3.3 Methodology

- In the frontend, a Flutter-based mobile application will be used by the user.
- In order to have a proper scaled and distributed architecture, we have split the backend into two, a Flask based server connected to our Realtime Firestore database in order to keep track of data and to to handle tweets classification requests powered by our CNN model.
- The system works with the Twitter account of the data from where the data is fetched. Using the users' data fetched, the user gets access to the analysis of the tweets from where they can keep track of their data going towards extreme sides like negative or positive and take medical assistance appropriately.

3.3.1 Methodology of the Complete System

- Initially the user will have to login into the system and provide it permission to connect to their Twitter account.
- From there the system will fetch the user's Twitter data like tweets, date and time.
- Using the trained model the tweets will be assigned a score and classify it to positive, negative or neutral.
- The score will be displayed in the form of a graphical representation as per the weekly analysis of the tweets of the user.
- If there is more negative display on the graph, then the user is alerted via a notification which will display nearby medical consultation.

3.4 Algorithm

The model used in this project for feature extraction is Twitter OAuth 2.0 API and for classification is Convolutional Neural Network (CNN or ConvNet) . They are also known as shift invariant or space invariant artificial neural networks (SIANN). This model has been used to perform categorical cross entropy classification which classifies into three different factors and further achieves 84.2% test accuracy, but can be improved by training more data. After increasing the data size and further training gave us a test accuracy of 89.16%. This is the best algorithm for our project as compared to other algorithms that gives accuracy less than what CNN gives.

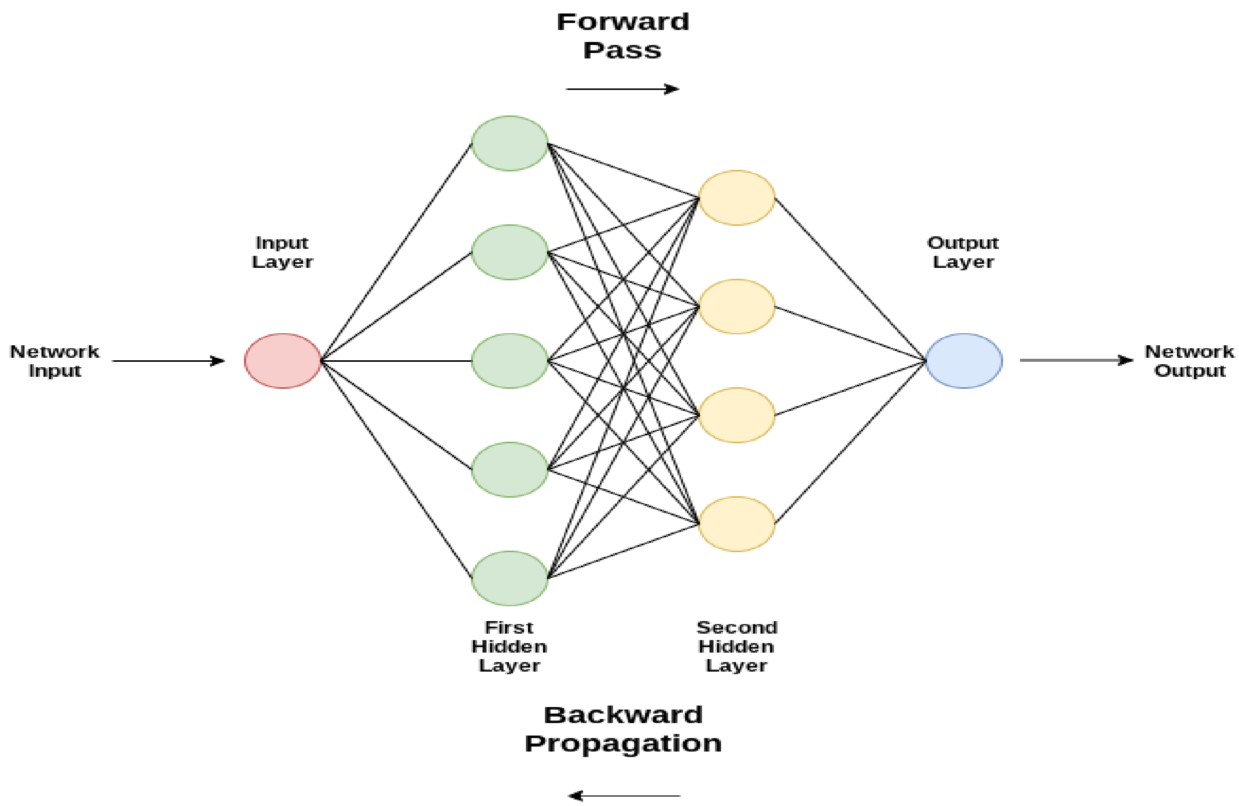

Figure 3.4 Convolutional Neural Network

1. **Data Exploration and Fetching**: Fetch data from users' accounts using REST API (Twitter OAuth 2.0) along with Tweepy by Python and store the data fetched in tab-separated file or csv format.
2. **Data Loading and Cleaning:** Since the data contains stop words and punctuations, NLP with RegEx is used to remove them, further tokenizing them with NLTK's word_tokenize.
3. Using TextBlob, we extract the subjectivity and polarity from the cleaned texts and apply them on the cleaned data, further computing the negativity, positivity and neutrality in the tweets.
4. **Splitting into Train and Test:** Split data set into train and test. We will use 90% data for training and 10% for testing. Then, we build training and testing vocabulary and get maximum training and testing sentence length and total number of words training and testing data.
5. **Load Word2Vec:** Train the classifier using a pre-trained model (Google News Word2Vec).
6. **CNN:** Train the classifier on a basic CNN network and compare the results.
7. Use test set accuracy to select the best model for our use case.

8. Deploy the model to a cloud-based service.

4. HARDWARE SOFTWARE REQUIREMENT AND IMPLEMENTATION

4.1 Hardware Requirements

- A phone: Android OS
- Good Internet Access

4.2 Software Requirements

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

4.3 Implementation Plan for Next Semester

With the help of the user's Twitter data the system can detect early stages of depression. This detection during the early stages can go a long way to help a person avoid some serious and long lasting problem in the future. But only detection is not enough, since it may not be something that people can cure on their own. More often than not they require professional help as well.

1. Integrate the model with User Interface: A user-friendly application using Flutter will be built for easy access of tweet analysis in graphical format. Also, the UI will be more intuitive and usable so that the user can easily navigate through the app without any difficulties.
2. Integrate Google Maps API: Using Google maps API the system will recommend psychological medical experts to the user based on his current location once his sentiment score goes down below a certain threshold.

5. RESULTS AND CONCLUSION

5.1 Result

This report shows the progress of the project until the seventh semester. We have successfully completed stages of Definition of Problem, Analysis, Requirement Gathering, Procuring the Dataset, Comparing two different Deep Learning Models and determining the model for our system that gives better accuracy than the other and implementing data Extraction from Twitter with a trained model. We will now proceed with the integration of the model with the mobile application.

5.2 Conclusion

The analysis of Twitter data for the detection of negativity that has the possibility to cause mental disturbance has increased in the past few years and is only going to grow and advance as more research continues. This project is implementing a method of tweet extraction that is further used to analyze the state of the person with good accuracy. This project plans on using the CNN model for implementation. This was decided by comparing it against Naive Bayes, SVM, Random Forest and K-nearest neighbours models. A lot of planning and analysis has gone into locking down one model for execution. This has increased the clarity about the design and implementation procedure as a whole.

6. APPENDIX

OAuth - Open Authorization
CNN - Convolutional Neural Network
NLP - Natural Language Processing
REST API - Representational State Transfer Application Programming Interface

7. REFERENCES

[1] "Convolutional Neural Network - Wikipedia." wikipedia.org/ wiki/ Convolutional_neural_network. (Accessed 17 July. 2020)

[2] "Social Computing - Wikipedia." wikipedia.org/ wiki/ Social_computing. (Accessed 17 July. 2020)

[3] "Anxious Depression Prediction in Real-time Social Data" - International Conference on Advanced Engineering, Science, Management and Technology -2019 (ICAESNT19), March 2019

[4] "Sentiment Analysis/Text Classification Using CNN (Convolutional Neural Network) - Towards Data Science (Accessed 30 July. 2020)

[5] "Deep Learning for Depression Detection of Twitter Users" - University of Ottawa, ON KIN 6N5 Canada, January 2018
[6] "Detecting Stress Based on Social Interactions in Social Networks" - Journal of Latex Class Files, Volume-13, No-9, September 2014
[7] "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
[8] "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