Predicting depression levels using social media posts

W.T.M.Samaranayaka

2021

Predicting depression levels using social media posts

A dissertation submitted for the Degree of Master of Computer Science

W.T.M.Samaranayaka

University of Colombo School of Computing

2021

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

Student Name: W.T.M.Samaranayaka

Registration Number: 2018/MCS/076

Index Number:18440768

29/11/2021

Signature of the Student & Date

This is to certify that this thesis is based on the work of Ms. W.T.M.Samaranayaka under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by,

Supervisor Name: Dr. Ajantha Atukorale

29/11/2021

Signature of the Supervisor & Date

I would like to dedicate my dissertation work to my beloved parents as gratitude for all the support they gave me and for the words of encouragement and push for completing the research.

Also, I would like to dedicate this dissertation work to my supervisor Dr. Ajantha Atukorale who provided guidance and feedback throughout the research.

ACKNOWLEDGEMENTS

I would like to thank Dr. Ajantha Atukorale for his guidance and valuable advice even in the pandemic situation through phone calls/ emails/ zoom meetings. Also, I would like to thank the respondents who took the time to complete my questionnaire. Furthermore, I would like to thank my parents for their patience and encouragement.

ABSTRACT

Depression is a frequent and dangerous medical ailment that affects how you feel, think, and act. It has a negative effect on people's feelings, thoughts, and actions. Severe depression expresses itself in a variety of ways, including insomnia, anger, hopelessness, and even suicide attempts. There will be many reasons for depression like abuse, certain treatments, genes, deaths, or losses etc. The COVID-19 pandemic is one of the major health crises that has changed the life of millions of people globally. The COVID – 19 pandemics has affected 220 countries around the world and have reported 190,347,496 confirmed cases. Also, due to COVID – 19, people's mental health was reduced significantly, and this was found by one of the research groups at Boston University. Loneliness, worry, economic instability, and the daily hearing of bad news caused by the coronavirus epidemic are all taking a toll on people's mental health and may be feeling depressed or anxious. Also, in the last few decades, social media usage was increased drastically, and there tends to share people's emotions in public through social media.

Many types of research revealed that people's mental health can be measured using social media data. This research aims to prove that we can use social media data to predict depression in a pandemic situation like COVID – 19. As well as to prove that there is an increment in the number of depressed people or there is reduction in mental health in a COVID-19 like situation compared to the normal period using the social media data. For this purpose, I have used a combination of social media posts on Facebook and Vkontakte social media. Vkontakte's data was taken from the available online dataset. The Facebook dataset was developed by scraping the Facebook posts from publicly available Facebook pages. In order to train and test the model I have used that dataset. Also, different classification methods like trees, naïve Bayes, neural nets, rules, and logistic regression are used to train the dataset. Moreover, for each classification method has many sub-classifications. Emotion, sentiment, linguistic style, depression language and combination of all features (emotion, sentiment, linguistic style, depression language) are used as the features. The results were interpreted that the higher the number of features used, the higher the F-measure scores in detecting depression users. The highest f – measure was acquired by all features. When considering each individual features linguistic style feature obtains the highest f- measure.

Then I have collected another separate Facebook dataset to show that there is a reduction in people's mental health before COVID and in the COVID period. That dataset was scraped by the Facebook user profiles and that user group was selected based on the people who work in Information Technology (IT) industry. The main reason to select people working in the IT industry was due to the new normal of work from home and the working stress they were facing. Moreover, due to those reasons, there is a probability of falling into depression easily. Also, before accessing their Facebook profiles, their consent was taken through a Google Form, and a questionnaire was provided to them to answer. The questionnaire contains screening test questions for the diagnosis of depression. After collecting the dataset, the dataset was used to predict the depression of each respondent using the previously developed model. Based on the results of this dataset it shows a significant difference in mental health before the COVID period and after the COVID period. Result of this dataset reveals that there is reduction in mental health. Also, it reveals that social media data can be used to predict depression in a pandemic situation like COVID – 19. Furthermore, results obtained by the questionnaire and prediction model are similar for 66.67%. That reveals the prediction model gives moreover correct result.

This is a data-driven method and predictive approach for the early detection of depression. The main contribution of this research is the explore that the impact of some features to predict depression and to prove that can be used the social media data to predict people's mental health during pandemic situations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1Motivation

The motivation behind this research is the quote mentioned below, which was found through an article.

> "COVID-19 has tripled the rate of depression in US adults in all demographic groups—especially in those with financial worries—and the rise is much higher than after previous major traumatic events, according to a study published in JAMA Network Open" (Sep 03 and 2020, n.d.).

According to the findings of that study, 27.8 out of 100 adults experienced depressive symptoms, up from 8.5 percent prior to the outbreak (Sep 03 and 2020, n.d.). Increases were more significant across the depression severity spectrum, from mild (24.6 percent versus 16.2 percent before the pandemic) to severe (5.1 percent vs 0.7 percent) (Sep 03 and 2020, n.d.). Also, females were more likely than males to experience depressive symptoms prior to and throughout the outbreak (10.1 percent of females and 6.9 percent of males before the pandemic, vs. 22.2 percent of females and 21.9 percent of males during the outbreak) (Sep 03 and 2020, n.d.). They discovered that Asians had an 18.7 percent greater prevalence of depressive symptoms during the outbreak than before it (8 individuals [23.1 percent] versus 26 participants [4.4 percent]), although this was a small subgroup (Sep 03 and 2020, n.d.). Married people had an 18.3 percent rate of depressive symptoms, compared to 31.5 percent for widowed, divorced, or separated people, 39.8 percent for those who had never married, and 37.7 percent for those who lived with a partner(Sep 03 and 2020, n.d.).

This study was undertaken to see whether there is any change in people's mental health in a pandemic situation like COVID-19 using social media posts. For that I have used machine learning based approach and used social media posts to predict whether the person has depression or not.

1.2 Statement of the problem

This study mainly focused on solving the issue of whether we can use Facebook data to predict depression in a pandemic situation like COVID – 19 and prove that there is mental health reduction in situations like COVID-19.

Depression is one of the serious illnesses that negatively affect how people think, act, and feel. Globally, it has affected to 264 million people in the world. This makes people persistent feelings of sadness or lack of interest in a normal lifestyle. This is also known as major depressive disorder. Depression symptoms can be as follows, and it could be mild to severe. Feeling sadness, loss of interest in activities in once enjoyed, abnormal weight loss or gain, fatigue, slowed movement or speech, feeling worthless or guilty, difficulty in concentrate, difficulty in making decisions, thoughts of suicide are some of the symptoms of depression and in depression, those symptoms are lasting at least two weeks. In any given year, depression affects around one in every 15 individual adults (6.7 percent). One out of every six (16.6

percent) will suffer from depression at some stage of life. Also, while depression can hit at any age, most commonly, it starts in the mid-20s and late teens. Also, depressed females are more likely than depressed males. According to some studies, one-third of females will have a significant depressive episode at some point during their lives. When first-degree relatives (parents, children, or siblings) suffer from depression, there is a substantial degree of inheritance (about 40%) ("What Is Depression?," n.d.).

Job losses, relationship breakups, and the death of loved ones can be caused to sadness and those are typical situations in our lives. Hence, it is normal to feel sadness in such situations. However, sadness is not the same as depression. Also, grief and sadness are not the same. In grief, painful emotions occur one after another and are often mixed with positive memories; in depression, emotions and interest are weakened for almost the entire two weeks. Self-esteem is maintained throughout the grief, and low self-esteem and worthlessness are common during the depression ("Sadness, Grief, Depression," n.d.). When grief and depression occur at the same time, sadness is more intense and lasts longer than grief alone. Depression can be occurred due to many reasons and some of the most common reasons are environmental factors, genetic features, changes in the brain, psychological and social factors, additional conditions like bipolar disorder. There are also several forms of depression. Some of the most frequent depression kinds are major depression, persistent depression, bipolar disorder, depressive psychosis, perinatal depression, Premenstrual dysphoric depression, seasonal depression, situational depression, and atypical depression (Charlesthelindenmethod, 2019). Most of the people don't know they were suffering from depression, and they were seeing their family doctors for the problems such as fatigue, sleeping problems. Also, sometimes people think they don't need to get antidepressants for their lifetime and due to that they were avoiding treating depression. But if the depression affects badly to their life if it's got severed. Since the medication is most important. Normally, depression is diagnosed by medical practitioners through face-to-face clinical depression criteria. There are many reasons that can be caused for depression. Abuse, conflicts, loss or death, genetics, certain medications, major events of life, personal problems, serious illnesses, and substance abuse etc.

Depression is not normally going away naturally, and it may get severe without treatment. In depression detection, the most essential diagnostic tool is talking to the patient. To effectively treat or effectively diagnose, the doctor should be aware of specific symptoms of patients that reflect depression. To fulfill that, they were using a set of standard questions to screen for depression. The doctor will learn about other crucial details to making a depression diagnosis by conversing with the patient. Clinical depression can manifest itself in a variety of ways, making it challenging to identify. Doctors will diagnose the depression through the cause of depression with a physical examination, face to face interviews with patients and some lab tests. Doctors will identify the symptoms, which will include how long the patient has been experiencing them when they began and how the patient handled them. The following are some of the signs and symptoms of depression:

1. A sad mood that lasts much of the day or virtually every day
2. A decrease in appreciation of before pleasurable items
3. A significant weight change (a loss or increase of more than 5 percent of body weight in a month)
4. Sleeping incessantly or experiencing insomnia - this occurs practically every day.
5. Sense of being reduction of energy that others can notice or physical restlessness
6. Loss of energy - almost every day
7. Feeling's worthlessness or despondences or unnecessary guilt - almost every day

8. Difficulties with decision-making or focus - this occurs practically every day
9. Suicidal ideas, plans, or attempts that are repeated regularly

A patient must exhibit at least five of the symptoms listed above to be diagnosed with severe depression, and at least one of the first two symptoms must occur every day for at least two weeks to be considered depressed. Also, some physical issues may occur if anyone is falling with depression such as back pain, headaches, joint pain, limb pain, gut problems (digestion issues and belly pain), constant tiredness, sleep problems, slowing of physical movement and thinking (Bruce and PhD, n.d.). After discussing patients' emotional state and how it affects their lives, the doctor may also ask some questions used mainly to screen for depression.

In the screening test initially, doctors will give the two questions for the patients as below mentioned:

1. Is it possible that you've been troubled by feelings of sadness, depression, or hopelessness throughout the last month?
2. Have you been troubled by a lack of interest or enjoyment in the activities you have been doing over the last month?

Based on the answer to the above questions next step will be decided. If the patient's answer indicates that they were not felt with depression, the psychologist or doctor may review the patient's symptoms from the beginning to continue to find the cause. Studies have shown that these two questions, especially when combined with another test as part of the diagnosis process, are more effective tools for identifying most cases of depression.

1. The Patient Health Questionnaire-9 - (PHQ-9)
   Diagnostic screening severity test for self-management of major depression. This includes nine items based on the current diagnostic criteria for the disease.
2. Beck Depression Inventory (BDI)
   This tool assesses the severity of depression symptoms and emotions. It contains the 21 multiple choice questions in the questionnaire.
3. Zung Self-Rating Depression Scale
   A brief assessment that assesses the severity level of depression, with scores ranging from normal to severely depressed.
4. Center for Epidemiologic Studies-Depression Scale (CES-D)
   Patients to assess their emotions, behavior, and perspective from the preceding week using a structured questionnaire
5. Hamilton Rating Scale for Depression (HRSD)/Hamilton Depression Rating Scale (HDRS)/ HAM-D
   This tool is a multiple-choice questionnaire that physicians may use to determine the degree of a patient's depression. (Bruce and PhD, n.d.)

Using above screening tools and other test results doctors will identify whether the patient has depression or not.

Symptoms reflected by major depression can be different person-wisely. To identify the type of depression the patient hold, the doctor may add one or more specifiers. A specifier means that people have depression with distinct features, as mentioned below:

1. Anxious distress - worry about possible events or loss of control or depression with unusual restlessness
2. Mixed features - simultaneous depression and mania, which includes elevated self-

esteem, talking too much and increased energy
3. Melancholic features - severe depression with lack of response that used to bring pleasure and associated with early morning awakening, worsened mood in the morning, major changes in appetite, and feelings of guilt, agitation or sluggishness
4. Atypical features - depression that includes the ability to increased appetite, excessive need for sleep, temporarily be cheered by happy events, sensitivity to rejection, and a heavy feeling in the arms or legs
5. Psychotic features - depression accompanied by hallucinations or delusions, which may involve personal inadequacy or other negative topics
6. Catatonia - depression that includes motor activity that involves either purposeless and uncontrollable movement or fixed and inflexible posture
7. Peripartum onset - depression that occurs during pregnancy and in the weeks or months after delivery of baby (postpartum)
8. Seasonal pattern - depression relates to changes in seasons and reduce exposure to sunlight

Other disorders that cause depression symptoms

Several other disorders, such as those below, include depression as a symptom. It's important to get an accurate diagnosis, so you can get appropriate treatment.

1. Bipolar disorders - These mood disorders include mood swings that range from highs (mania) to lows (depression). It's sometimes difficult to distinguish between bipolar disorder and depression.
2. Cyclothymic disorder - Cyclothymic disorder involves highs and lows that are milder than those of bipolar disorder.
3. Disruptive mood dysregulation disorder - This mood disorder in children includes chronic/severe irritability and anger with frequent extreme temper outbursts. This disorder typically develops into depressive disorder or anxiety disorder during the adulthood or teen years.
4. Persistent depressive disorder - Sometimes called dysthymia, this is a less severe but more chronic form of depression. While it's usually not disabling, persistent depressive disorder can prevent you from functioning normally in your daily routine and from living life to its fullest.
5. Premenstrual dysphoric disorder - This involves depression symptoms associated with hormone changes that begin a week before and improve within a few days after the onset of your period and are minimal or gone after completion of your period.
6. Other depression disorders - This includes depression caused using recreational drugs, some prescribed medications or another medical condition ("Depression (major depressive disorder) - Diagnosis and treatment - Mayo Clinic," n.d.).

COVID-19 is an infectious disease caused by severe respiratory syndrome coronavirus 2. The first COVID positive case was identified in Wuhan, in China, in the last month of 2019("Coronavirus disease 2019," 2021). It has since spread all around the world, leading to a continuing pandemic. On December 31, the World Health Organization's national office in China stated that pneumonia with an unknown etiology had been reported in Wuhan. A novel type of coronavirus, COVID-19, was discovered by relevant experts and named due to their discovery. After a Public Health Emergency of International Concern (PHEIC) was proclaimed by the Director-General of the World Health Organization on January 30, 2020, and a pandemic was declared on March 11, 2020, the outbreak of COVID-19 was formally classified as a pandemic. Currently, nearly 161,206,768 cases were confirmed, and 3,345,317 deaths have occurred. This current pandemic situation also could be one primary reason to fall into the depression easily due to unpredictable mutations of the virus, the loss of manageability and freedoms, the contrary messages from authorities, unexpected changes in plans for the period ahead, the emotional toll of uncertainty, grief, fear, worrying and dramatic changes in habits and lives like new realities of working from home, temporary unemployment. Also, due to some public health actions such as social distancing, people feel lonely, which will cause anxiety and depression.

In the past few years, people addicted to social media such as Twitter, Instagram, Facebook. They used those platforms to share their thoughts, feelings, and experiences. So, it's become natural in posting social media when something happens in day-to-day life. Hence, social media posts provide a way for obtaining behavioral characteristics such as thoughts, moods, communication, socialization, and activities. The language and emotion used in social media posts may represent emotions of worthlessness, guilt, helplessness that characterize major depression. Due to that, social media can be used to predict the level of depression. That will be helpful for individuals who are experiencing depression to pay more attention to their mental health. Also, that might be helpful for the people who are in the health care industry.

Previous researchers have demonstrated that social media can be utilized to assess depression (De Choudhury et al., 2013). Also few research studies were done on the depression trend based on Twitter social media during the COVID-19 (Zhang et al., 2020). Munmun De Choudhury has demonstrated the ability to use Twitter social media to measure and predict major depression in people (Choudhury et al., n.d.). Hua Wang has proved the ability to use Facebook to measure and identify major depression (Islam et al., 2018). Hatoon AlSagri has defined a binary classification problem. This method uses the tweets and activities in the Twitter profile to detect whether the person is suffering from depression (AlSagri and Ykhlef, n.d.). They have used various machine learning algorithms and various feature datasets to prove that Twitter can be used as a resource to predict depression.

1.3 Research aims and objectives

1.3.1 Aims

The main aim of this study is to determine whether we can use Facebook posts and machine learning approaches to predict depression in a pandemic situation like COVID – 19.

1.3.2 Objectives

The main objectives of this study are as below:

1. Critical study about the problem domain – depression and it's behavior
2. Identification of a technology
3. Design a model
4. Evaluate the solution

In this study, I have used Facebook posts as the data source and the machine learning approach to design a model.

1.4 Scope

Based on the survey done by researchers in the Boston University they have identified there was a significant difference in mental health during pre-COVID and COVID time. This will be more useful to the people who are using social media and health care industry due to people will be able to identify the depression at the initial stage without waiting it's getting severed. This might be helped to reduce the risk of falling into depression.
Main contribution in this research are as follows:

1. I have used part of online available benchmark dataset and another part is collected from publicly available posts in the Facebook. Totally there were 37736 data rows. From that dataset 19226 data rows are labeled as True (depression indicative posts) and the rest of the data rows (18510) are labeled as false (not a depression indicative posts). By using that dataset, I have trained and test the model.
2. Based on the data set, I have used several features to label an individual's social media content
    a. Emotion
    b. Sentiment
    c. Depression language
    d. Linguistic style
    e. All (Emotion, Sentiment, Depression language, Linguistic Style)

    I have analyzed the data using Linguistic Inquiry Word Count(LIWC) software and get the values for the emotion feature and the linguistic feature. For the depression language feature was analyzed by using a python code using Google Colab. Also, for the sentimental analysis feature I have used the python code and retrieve the polarity and subjectivity through the code.

3. Then I have trained and test the data set using different classifiers and get the accuracy, precision, f – measure. The highest f- measure was gained for the 'All' feature for the Deep learning sub classifier as 92.36. From individual features the highest f – measure was obtained by the linguistic feature for the gradient boosted tree classifier. Lowest obtained by the language feature.
4. Then validate the other dataset using the trained model to prove that there is mental health reduction in pandemic situations and Facebook posts can be used to predict depression in a pandemic situation like COVID -19. That dataset was based on Facebook posts of people who engaged in work from home in the IT industry. For that, I have used 36 respondents for the initial screening and acquire a consent form. From that, I have removed 6 respondents due to different reasons and 30 respondents' data was acquired from Facebook profiles. There are two cohorts in the dataset. One is Facebook posts acquired based on the period before COVID-19(pre-COVID period) and another one is in the COVID period. Here I have identified there is a significant difference in people's mental health before the COVID period and in the COVID period. Also, I have proved that Facebook posts can be used to predict depression in pandemic situations like COVID - 19.

This project will be valuable for every person to measure their mental health condition and the people who are in the health care industry

The main limitation of the project is the time. For the data collection it was took a considerable amount of time and with the limited remaining time must complete the rest of the project. Also, since this is highly ambitious project it could be go wrong if unable to collect a proper data set. Also, since the posts are directly taken from the Facebook it took considerable time to preprocess the data. As well since this online available dataset is in Russian language and it needs to translate into English. To fulfill that I have used Google Translate. But since it does not translate the text 100% accurately there might be errors in the result.

1.5 Structure of the thesis

The remainder of this thesis is organized as follows:
“Literature Review” presents the related work of predicting depression using machine learning methods. Methodology is explained in the third chapter and fourth chapter will explain the results and evaluation. Last chapter will present the conclusion and future work.

Next chapter will present the critical review analysis about the predicting depression using social media posts.

CHAPTER 2

LITERATURE REVIEW

This section will present the different methods used to predict depression and there is broad area of literature that analyses the properties of depression.

2.1 Literature review

Munmun De Choudhury and his team demonstrated the ability to use Twitter social media to measure and predict depression in people(Choudhury et al., n.d.). They have used crowdsourcing to collect the data since it is economical and less time-consuming. Initially, they have done a screening test to identify the actual depression patients. For that, they have used the Center of Epidemiologic Studies Depression Scale (CES-D) screening test (Choudhury et al., n.d.). Based on that group, they have introduced several measures used to identify social media users' behavior in social media. They measured user engagement and emotion, linguistic style, egocentric social graph, antidepressant medications, and depressive language. They have measured depressed and non-depressed user classes separately through the measures as mentioned above. They have found that users who have depression interpret less social activity, more prominent negative emotion, increased relational and medicinal concerns, high self-attentional focus, and increased the expression of religious thoughts. Since they have used crowdsourcing, it may affect by the noisy responses, and to avoid that, they were used Beck Depression Inventory (BDI) as an additional screening test and discard the data points that took less than 2 minutes.

Initially, they proposed a collection of characteristics that can be used to characterize the variations in behavior between the two groups of Twitter users. One class is individuals exhibiting clinical depression and the other class does not show clinical depression. When considering engagement, they have defined 5 measures as volume, reply, retweets, links, question-centric posts. In the Egocentric social graph structure, they have taken measures as node properties, dyadic properties, and Network properties. They have taken another 4 measures for emotion. Those are negative effect (NA), positive effect (PA), activation, dominance. In linguistic style they have determined 22 specific linguistic styles articles, auxiliary verbs, functional words, assent, conjunctions, adverbs, personal pronouns, prepositions, negation, certainty, quantifiers (Choudhury et al., n.d.). Lastly, they discovered two characteristics that define the language of people who have been diagnosed with depression. The 2 classes are depression lexicon and antidepressant usage. The depression vocabulary measures the use of depression-related phrases represented broadly in Tweets. They have developed a vocabulary of terms possible to use in Tweets of individual users considering depression or depression symptoms in online environments to achieve that goal. The other feature, “Antidepressant usage” considers the degree of use of antidepressants recommended in the practice of clinical depression diagnosis.

They have explored the behavior of depressed and non-depressed user classes in light of the above measures. In the non-depression class, most users are not active on the nightstand early morning with activity generally increased throughout the day. The depression group shows a rise late in the night, with lower activity during the day. Two groups of users have seen decreased numbers of followers based on egocentric network measurements, and these users have a diminished desire to socialize or a willingness to ingest external information and stay linked with people (Choudhury et al., n.d.).

They've shown that utilizing Twitter as a source for assessing and forecasting severe depressive illness in individuals is a viable option (Choudhury et al., n.d.). According to their conclusions, people with depression have less social engagement, high self-attentional focus, enhanced medical and relational concerns, more negative feelings, and more tendency to religious expression thoughts. Also, they have found that there is a tendency to highly clustered closed networks and highly correlated to their audience. Finally, they have leveraged all the attributes mentioned above to build a support vector machine classifier that can predict ahead of the reported onset of a user. Furthermore, that classifier obtained results with 70% classification accuracy.

Also, Munmun De Choudhury and his team have demonstrated the feasibility of using social media to track depression tendencies at a community level (De Choudhury et al., 2013).

They have researched the possibility of using social media data as the new way to understand depression in communities. They also employed crowdsourcing to acquire data on a group of social media users' depression levels. Also, they have used the CES-D questionnaire to determine the depression levels of the employed workers in the study. Additionally, they have collected information about the employed workers' depression history.

- Are they have ever been diagnosed with clinical depression? When?
- If yes, when did the depression start?
- How many depression waves had they experienced since the commencement of depression?

Initially, they have created a ground truth dataset with depression positive class and depression negative class. Then they have identified several features to characterize the posts. Categorized features are post-centric and user centric. Emotion, linguistic style, time are post features. Engagement and ego-networks are user features. Emotion is classified as another 4 features: positive affect, negative affect, activation, and dominance. Also, they have defined timestamps as daytime or nighttime posts. Then introduce another feature to distinguish social media posts based on the linguistic styles used. There they have used 22 specific linguistic styles. In user features. To define the normal behavior connected with the people, they have used a set of engagement measures. They have assumed that the clinically depressed will display significant behavior in their postings. The engagement features are like below:

1. The number of Tweets each user has twitted so far
2. Percentage of user replies
3. Author retweets
4. Number of links twitted by each user
5. The portion of question-based posts from each author

They defined two other elements of an author's egocentric network:

1. The count of followers or in links of the user
2. The number of followees or out links

They present a descriptive analysis of differences in two types of postings based on their findings. Posts are depression-indicative and standard.

In standard class most of the unigrams are related to commonplace details of daily life and ranged from work to entertainment. Ex: Life, work, friends, tomorrow, movie. Also, in depression positive class most of the words are emotional. Ex: hurts, hate, hope.

They employed supervised learning to determine whether a given post is depression-indicative. They depicted Twitter posts as vectors of prior features (Ex: rime, style, n-grams, and engagement features). Hence, they have used principal component analysis. The classification algorithm is a standard SVM classifier with an RBF kernel. Furthermore, they have used 5 -fold cross-validation and conduct 100 randomized experimental runs.

Using the prior mentioned model, they observed the ability to predict the class of the posts. They used n-grams, emotions and time features, engagement and ego network features, linguistic style, all features together, and dimensionality reduced features to study the value of each feature.

Their top model has an average accuracy of roughly 73% and a high precision of 0.82. That result was obtained compared to the depression-indicative posts. They found more incredible performance for models that only incorporate linguistic style features and emotion and time features. According to previous research studies, language styles reveal how people react to psychological triggers. Hence style traits can be used to reflect depression.

Secondly, the model performed better by combining emotion and time parameters that reflect negative, positive, activation, and dominance. Depression can be characterized by disrupted processing of emotional information and a decreased feeling of daily emotional activities.

According to psychiatric research, 8 out of 10 depressed patients experience worsening symptoms at night. That was the main reason to get the timestamp as one feature to predict depression.

To measure the real-world depression rate in a large group of people, they have employed a predictive model. It classifies whether a given post is depression reflecting or not and it can automatically specify a large set of data shared on Twitter on any particular date. Then they were developed a metric called the “Social Media Depression Index”. Twitter users' daily posts are used to calculate this index.

$$SMDI(t) = \frac{n_d(t) - \mu_d}{\sigma_d} - \frac{n_s(t) - \mu_s}{\sigma_s},$$

t = given day

$n_d(t)$ = Standardized difference between the frequencies of depression-indicative posts

$n_s(t)$ = Standardized difference between the frequencies of non-depression-indicative posts

μ_d = Mean of the number of depression indicative posts shared in a fixed time period before t

μ_s = Mean of the number of non - depression indicative posts shared in a fixed time period before t

$\boldsymbol{\sigma}_d$ = Standard deviation of the number of depression indicative posts shared in a fixed time period before t

$\boldsymbol{\sigma}_s$ = Standard deviation of the number of non-depression indicative posts shared in a fixed time period before t

To find out the population characteristics of the depression they have identified few cities in the US which reported in 2011 as 'unhappiest US cities. Then they have escaped 30% random tweets between Jan 2011 and December 2011. The city was identified via the user post authors' self-reported location on the Twitter profile. Then they have used their prediction model to label the post as to whether they are depression indicative or not. Then they calculated SMDI for each day and the mean SMDI for each city in 2011. Then they plotted their data against the reported depression rates in each city. They found that they overestimated despair in more cities than they thought. Then they have expanded their geographical analysis to 50 states. Then they have constructed a heat map of actual and forecasted depression rates in various states. By using that, they were able to show that their metric captures the content to some extent.

Also, they have performed a demographic analysis to find gender differences in depression. They have the same set of Tweets used in the analysis reported in the geographical study. Furthermore, since gender is not a character in Twitter, they have used gender classifiers. This gender classifier is known to be between 85% - 90% accurate. Then they have calculated the individual-centric measure of SMDI. They discovered that females suffer from depression 1.5 times more than males. Finally, they analyzed diurnal and yearly trends of depression. They have found that the diurnal SMDI value for females is higher than that of men. Also, they have found that SMDI is higher at night than in the daytime for both men and women. The lowest SMDI value is appearing at noon and peaking at around midnight. Also, they have identified a seasonal pattern in depression throughout the year and the maximum depression rate recognized during the wintertime in the US while minimum during summertime.

In conclusion, they have identified a pattern in the expression of depression on social media throughout the whole day across males and females and seasonal levels across various locations.

Munmun De Choudhury, Scott Counts, Aaron Hoff and Eric Horvitz demonstrated the ability to use Facebook as a medium to detect, characterize, and predict postpartum depression in new mothers(De Choudhury et al., 2014). Childbirth is a major incident in the life of every parent. Also, a considerable number of new mothers are experienced with postpartum depression They have surveyed to collect data about new mother's postpartum depression experiences and their Facebook data. From the survey, they have gathered some demographic data related to the childbearing experience. Also, it includes the patient health questionnaire to detect whether the mother is suffering from depression.

To measure the behavioral characteristics, they have used seven user characteristics that measure a user's behavior of activity on Facebook. The measures are;

1. The number of status updates made by a mother
2. The number of media items uploaded by a mother

3. The number of wall posts made to specific friends on Facebook
4. The rate of change of posting activity over time
5. Captures the degree to which a mothers Facebook activity shows a negative trend
6. Variation in the number of posts per week (Entropy)
7. Mean power of the number of posts per week

Previous research has demonstrated an adverse link between depression and cognitive, social capital ("Fujiwara and Kawachi - 2008 - A prospective study of individual-level social cap.pdf," n.d.). They used one-on-one social engagement as the social capital of moms. Also, they have considered positive affect (PA) and negative affect (NA) measures of mothers' emotional state. Also, they have defined a content characterization measure to determine whether they are looking for advice or sharing the information. For that, they have checked whether the posts contain a question or not. They assessed prenatal and postnatal behavioral changes based on mothers' linguistic style in their posts.

Initially, they have used the antenatal period data to detect whether that mother will have postpartum depression throughout the postnatal period. Considering different behavioral measures, they have fitted several regression models to understand relative values: user characteristics, linguistic styles, social capital, and content features, and ultimately used all the above measures. Also, they have fitted a null model that uses all the self-reported attributes and demographic attributes related to childbirth. For all the models, they have used stepwise logistic regression models.

Independent variables - different behavioral measures

Response variable - whether mother reported the onset of postpartum depression following both

The first model they have used is the demographic model and it contains age, ethnicity, occupation, income and whether the childbirth in context of their survey was the first child of the mother and if it was a premature child (De Choudhury et al., 2014). Then, they test the performance of numerous prototypes that incrementally add to the demographic model's variables: characteristics of the user, social capital, properties of the content, and linguistic style (De Choudhury et al., 2014). Their third model, which adds content features to the prior model, only slightly improved performance, explaining approximately 27% of the data variation (deviance decreased to 79.64) (De Choudhury et al., 2014). The fourth model uses the different markers of linguistic style as new variables in the regression model (De Choudhury et al., 2014).

At the end of the study, they discovered that postpartum depression increased social isolation, as seen by lower social activity and usage of Facebook and reduced access to social capital. Because Facebook friend networks often contain offline social links spanning friends and coworkers, and because mental illness is stigmatized, emotional measurements were founded as less efficient predictors (Zhang et al., 2020).

Maryam Mohammed and Hafiz Farooq have demonstrated an association between social network sites user behaviors and mental health disorders (Aldarwish and Ahmad, 2017). They have assumed that social network activities can reveal mental disorders at the primary stage (Aldarwish and Ahmad, 2017). They were trying to overcome the problem of self-reporting. They have proposed a web application that can classify users into a particular depression level.

First, they have collected the user-generated content from the patient's Facebook and/or Twitter accounts. Then collect the answers to the questions based on the Baker Depression Index-II (Aldarwish and Ahmad, 2017). Then, it analyses the user-generated content using various text analysis API's. Then, it classifies the user into one level out of four levels as severe, moderate, mild, and minimal (Aldarwish and Ahmad, 2017). Then they used RapidMiner to evaluate Support Vector Machine and Naive Bayes Classifiers on a depression model that they developed. The model contains two data sets and seven different operators (Aldarwish and Ahmad, 2017). The first training data set includes 2073 posts from each group. The posts are manually trained as depressed and non-depressed. It has three columns. The first column is the binary sentiment (depressed or not), the second is the category, and the third is the trained post (Aldarwish and Ahmad, 2017). The second data set contains the patient's social media network posts. The first operator is a select attribute and it decides which operators are removing and which remain. The Nominal to Text operator comes next. This operation converts the chosen nominal attribute to text and maps all nominal attributes to text. Also, it was used in the training and test dataset. The fourth and fifth operators are employing in the training and test data sets, which construct word vectors from string properties, and it has four operators. The four operators of the process document are stemming, conversion cases, filtering stop words, and tokenization (Aldarwish and Ahmad, 2017). The sixth operator is the validation operator. It is applicable to the training data set, which is divided into two sections: training and testing. The training portion contains the classifier operator, and each time patients are tested, the classifier model is switched from SVM(Linear) to Naive Bayes. The testing part includes two operators: Apply Model, which applies the trained model to the supervised data set, and performance, which is used to evaluate performance. The final operator applies the model that connects the test and training data sets, providing us with the final result of the prediction using a single classifier in patients. The classification accuracy is determined by the training set on which the classifier was trained. To do this, sample training nodes representing edge situations that fall within or outside of a class must be chosen. They accomplished this by amassing data from Facebook, Twitter, and LiveJournal. They measured accuracy, precision, and recall in order to assess the suggested model's performance.

Yipeng Zhang, Hanjia Lyu, Yubao Liu, Xiyang Zhang, Yu Wang, Jiebo Luo have developed transformer-based models and they have trained with the biggest depression data set so far (Zhang et al., 2020). They have compared their models' performance against other existing models, and they have found that the largest data set increases the performance of the model. Also, they have shown that the models can be used to monitor the depression and stress trend against geographical entities such as states. They have found effective methods to find depression users on Twitter (Zhang et al., 2020). They built a tool combining deep learning models and psychological text analysis to improve the classification (Zhang et al., 2020).

Hatoon AlSagri and Mourad Ykhlef have explained that determining whether the person is depressed or not as a binary classification (AlSagri and Ykhlef, n.d.). They have exploited various machine learning algorithms and various feature sets. Also, they have performed many preprocessing steps such as data preparation and aligning, data labeling, feature extraction, feature selection (AlSagri and Ykhlef, n.d.). Support vector machine classifier achieved optimal accuracy by converting an extremely nonlinear classification problem to linear separable classification. Although the decision tree model is comprehensive, there is a tendency to fail it is using for a brand new dataset (AlSagri and Ykhlef, n.d.).

Hayda Almeida, Antonic Briand, and Merie-Jean Meurs have demonstrated that depression predicting systems performs well when using the multipronged approach, which combines predictions from information retrieval systems and supervised learning methods. Supervised learning-based systems were made by using the feature types, classification algorithms, logistic model algorithm, ensemble sequential minimal optimization and ensemble random forests. As features they have selected user posting frequency, n-grams, selected part of speech, dictionary words (Almeida et al., n.d.). They have merged the predictions obtained from supervised learning and information retrieval approaches by using a decision algorithm (Almeida et al., n.d.). Those results demonstrated that the combination of supervised learning and information retrieval methods outperform the results obtained by each approach applied individually (Almeida et al., n.d.).

Rafiq Islam, Muhammed Ashad Kabir, Ashir Ahmed, Abu Raihan, Hua Wand, and Anwaar Ulhaq set out to analyze Facebook data to identify any indicators associated with depression among relevant Facebook users (Islam et al., 2018). In this study they have used the publicly available data containing user comments. Initially they have cleaned the data and then analyze the data by using the software application which is called LIWC. LIWC is a text analysis strategy and can process the text line by line. Then they have done the feature extraction. Finally, they have suggested machine learning techniques. They have applied decision tree, k-nearest neighbor, support vector machine and ensemble classifier techniques to detect emotional terms (Islam et al., 2018). They also shown that all classification algorithms based on linguistic style, emotional process, temporal process, and all aspects can successfully extract the depressing emotional result (Islam et al., 2018). They also show that the decision tree classification method outperforms the others. (Islam et al., 2018).

S.K.Schafer and team have examine the mental health before and after the COVID-19 outbreak and potential modulatory effects of sense of coherence(SOC) (Schäfer et al., 2020). They have found there was significant clinically symptom changes in 18% of respondents. In addition, they also used a bivariate latent change scoring model and determined that the high-stress group had a higher incidence of psychopathological symptoms and a lower incidence of sensory consistency, while the low-stress group showed the opposite of this style (Schäfer et al., 2020). They have found small group which has low levels of sense of coherence is experiencing increased if psychopathological symptoms from pre-outbreak(Schäfer et al., 2020).

Guangyao Sheng and the team has demonstrated that how to detect a depression using multimodal dictionary learning solution with the use of social media data (Shen et al., 2017). They suggested a multimodal depressive dictionary learning approach to detect depressed people in Twitter using benchmark depression and non-depression datasets and well-defined discriminative depression-oriented feature groups. The researchers next examined the contribution of the feature modalities and identified depressed users on a large-scale depression-candidate dataset to uncover some underlying online behavior differences between depressed and non-depressed users on social media. There are three main contributions in this paper: First one is they have created and released well labelled (depressed/ non-depressed) benchmark dataset released that to the online. They extracted six sets of distinct depression-related characteristics to characterize users from various perspectives. As only a few of the users' behaviors are symptoms of depression, they present a multimodal depressive dictionary learning model to learn the sparse representation of users (Shen et al., 2017). The techniques they have developed may be used to identify depression in real time and to take proactive measures to prevent the depressed state from worsening. Then they have analyzed feature contributions and online behaviors of depression. They have defined 6 depression-oriented features: Social network feature, user profile feature, visual feature, emotional feature, topic level feature, domain specific (Shen et al., 2017).

They have presented a multimodal depressive dictionary learning model (MDL) to identify depressed users, with the idea of (Shen et al., 2017):

1) Learn the latent and sparse representation of users by dictionary learning
2) Jointly model cross modality relatedness to capture the common patterns and learn the joint sparse representations
3) Train a classifier to detect depressed users with the learned features specifically

They have used Uni-modal Dictionary Learning, Multi-modal joint sparse representation, depression classification methods (Shen et al., 2017). Whether they extract a set of features from each modality, they may not all be related to the depressed group (Shen et al., 2017). In addition, since the content of Twitter posts is usually in free form, some noise is added to the modal, which may affect the accuracy of detection (Shen et al., 2017). By using dictionary learning, they learned the latent and sparse representation of the user. Because the various modalities are not independent of each other and share some common patterns, monomodal dictionary learning cannot capture them. Therefore, dictionary learning is extended to multiple modalities to link features across modalities and learn to combine sparse representations to obtain latent features.

They have found some interesting features about depression users related to their posting time, emotion and self-awareness. Related to the posting time they have identified most of the depression users are posting in Twitter between 23:00 to 6:00. Also, they have revealed most of the depression users are willing to express about their emotions, mostly negative emotions (Shen et al., 2017). Depressed patients use 200% more first-person pronouns in tweets than ordinary users, representing their self-awareness. Depressed individuals post about an antidepressant and depression symptom phrases 165 percent more than non-depressed users on average, demonstrating that they are open to sharing what they face in real life. The main goal of this study is to use social media posts to diagnose depression in a timely manner. They have analyzed the feature modalities and detect depressed users (Shen et al., 2017).

Keumhee and the team was able to propose new method to identifying the users which has depressive moods by analyzing their social media data for a considerable time of period (Keumhee Kang et al., 2016). For more precise understanding about the users, they have used all types of media such as text, emoticons and images and developed a multi modal system to analyze twitter posts. Initially they have retrieved the user's hidden moods by analyzing text, emoticons and images using three single analyses. The three models are learning based text analysis, word-based emoticon analysis and SVM based image classifier (Keumhee Kang et al., 2016). Then they have integrated those data into a mood and again aggregated per a day, which allows for continuous monitoring of user's mood fluctuations. To validate the proposed method, two types of tests were performed (Keumhee Kang et al., 2016):
1) the proposed multimodal analysis was tested with a number of tweets, and its performance was compared to SentiStrength
2) it was applied to classify 45 users' mental states as depressive and non-depressive ones

Then, the results demonstrated that the proposed method outperforms the baseline, and it is effective in finding depressive moods for users (Keumhee Kang et al., 2016).

That proposed system contains with four modules. Those are crawling, sentence segmentation and classification, single model analysis and classification. Initially they have performed crawling and they have used open API. They have used keywords for collecting ground-truth dataset and depressive users. The crawled tweets were filtered by using queries that include in tweet information such as user ID, date, and keyword. After filtering the data, they have split those into sentences using conjunctions and punctuation marks. After that, those split sentences were divided into text, emoticons and images. To precisely predict users' moods from the content of tweets, three analyzers are developed. For the texts, a learning-based analysis is conducted, which considers the forms and structures of a phrase and the terms linked to the human emotions contained in the sentence and learns the connections between them using a support vector machine classifier. For the emoticon, they have built a new lexicon that includes 136 negatives and 66 positives and perform the word-based analysis (Keumhee Kang et al., 2016). The images belong to a tweet is analyzed by a SVM-based classifier (Keumhee Kang et al., 2016). To observe the daily moods of a user, the results of three single-modal analyses were aggregated for a tweet, for a day and further for a month. Finally, they have analyzed the pattern of the user's mood trends and predict the depressive state.

An anxious depression model has been introduced by the Kumar and authors (Kumar et al., 2019). The feature set was defined using a five-tuple vector based on language cues and the user posting patterns: word, timing, frequency, sentiment, and contrast. An anxiety-related vocabulary was developed to identify the presence of anxiety indicators (Kumar et al., 2019). Social media and mental health of users can be related in three methods: Social media anxiety disorder, Anxious depression social media verbalization and social anxiety (Kumar et al., 2019). Initially they have scraped the data from Twitter using an API. They have selected a group of people who can be fallen to depression easily. They have assumed that younger generation who are lived away from their home may have more ability to be victims of depression. After the data collection they have pre-processed that collected data. As pre-processing they have removed numbers, empty texts, URLs, mentions, hashtags, stop words, non-ASCII characters, and punctuations. Then they have tokenized the cleared data and convert the dataset to normalize. Then they have replaced emojis and slangs by descriptive texts using SMS Dictionary("Vodacom Bulk SMS Messaging - SMS Dictionary," n.d.) and Emojipedia("📙 Emojipedia — 😃 Home of Emoji Meanings 💁👌🎍😍," n.d.). Then they have used stemming to reduce the words to their roots. In feature engineering, they have trained a model using a vector with five - tuples: word, timing, frequency, sentiment, contrast. The feature word is based on the anxiety lexicon. That anxiety lexicon is based on the keywords that express anxious depression in the texts (Kumar et al., 2019). The seed list is ultimately grown with the WordNet (Kumar et al., 2019). The timing feature was based on whether the user is more active on night hours 12am to 6am. Sleeplessness is a frequent sign of anxious sadness. The next feature, frequency also based on number of posts shared within 24 hours. They have chosen the values in below manner: If number of Tweets per hour is equal or greater than three they have set the feature value to true '1', else set to false '0'. Next feature is the sentiment. To determine the sentiment of each tweet they have used SentiWordNet. That SentiWordNet is a lexical database that categorizes words according to their polarity (Kumar et al., 2019). Since same word can have different sentiment in different scenarios it could have positive sentiment, negative sentiment and neutral sentiment. A cumulative number representing the average of negative polarity tweets sent throughout the period is computed. If 25% or more of the tweets have negative polarity, the feature is set to 1; otherwise, it is 0. The polarity shift of postings from positive to negative or negative to positive reveals an unstable mental state and restlessness. Flip-flop behavior is a term used to describe someone who suffers from anxious depression and has muddled thinking. Thus, to determine the difference in opinion polarity of tweets, they utilized the following equation (Kumar et al., 2019):

$$C = \{(\delta.PP + pw) - (\delta.NP + nw)\} / \{(\delta.PP + pw) + (\delta.NP + nw)\}$$

pw = count of words with positive opinion polarity
nw = count of words with negative opinion polarity
PP = count of positive post
NP = count of negative post
δ = post co-efficient, value is set to 3

Then they have used the supervised learning methods to train the features. Mainly they have used 3 machine learning classifiers Multinomial Naïve Bayes, Gradient Boosting and Random Forest (Kumar et al., 2019). In addition, they generated the final prediction using an ensemble vote classifier with a majority voting process (Kumar et al., 2019). For data splitting they have used 80 and 20 and 10-fold cross-validation was done.

To test the performance of the ensemble vote classifier, they used accuracy and F-score measurements (Kumar et al., 2019). They have achieved 85.09% accuracy and 79.68% for F-score for the proposed AD prediction model.

Lang He and Cao et al., has introduced a new method to overcome few problems in detecting depression efficiently (He and Cao, 2018). They have proposed a combination of deep-learned and hand-crafted features which can measure the severity of depression from speech (He and Cao, 2018). In proposed method, Deep Convolutional Neural Network are initially built to learn deep-learned features from spectrograms and raw speech waveforms (He and Cao, 2018). Then they have manually extracted the state-of-the-art texture descriptors named Median Roust Extended Local Binary Patterns from spectrograms (He and Cao, 2018). To detect the complementary information within the hand-crafted features and deep-learned features, they have proposed joint fine-tuning layers to combine the row and spectrogram Deep Convolutional Neural Network (DCNN) to boost the depression recognition performance (He and Cao, 2018). They've also presented a data augmentation strategy to deal with the problem of tiny samples.

As the initial steps of methodology for hand crafted features they have extract the Low-Level Descriptors (LLD) from raw audio clips and Median Robust Extended Local Binary Patterns features from the spectrograms of audio (He and Cao, 2018). Then they have used DCNN to directly learn the deep-learned features from the raw audio and spectrogram images (He and Cao, 2018). At last, they have described the proposed joint fine-tuning method to combine the four streams for the last depression prediction (He and Cao, 2018). For hand crafted features they have used two different kinds of descriptors were used (He and Cao, 2018). The two descriptors are Median Robust Extended Local Binary Patterns (MRELBP) and audio features extracted by openSMILE toolkit (He and Cao, 2018). MRELBP is a novel descriptor for texture classification (He and Cao, 2018). At the end of that study, they were able to introduce joint tuning layers, to combine the raw and spectrogram DCNN, which able to improve the performance of depression recognition (He and Cao, 2018).

2.2 Research gap identified

Considering about findings of the previous studies done by researchers they were observed that there is a significant difference in mental health when people have a low sense of coherence. Also, there was several studies that proved that social media data like posts, comments, social media activities can be used to predict depression. Also, some of the researchers have revealed that what are the features that can be used to predict depression using social media data. And the other important fact is the demonstration of using machine learning techniques to detect depression. They have used different supervised and unsupervised learning methods. Also, there is another new finding that mental health is reduced with the COVID situation. With the consideration of all the above facts, there was a gap between the previously identified depression detection methods and the new findings of depression during the COVID period. That was to predict the depression using social media data and machine learning techniques to identify whether there is any impact on depression on COVID situation. This gap was identified due to there was no research related to the pandemic period impact on depression identification using machine learning methods.

CHAPTER 3

METHODOLOGY

3.1 Problem analysis

This section will present the analysis about the problem going to be solved.

The main aim behind this study is to detecting depression by using social media data and identify whether there is any impact from COVID to the people's mental health. Most of the previous researchers have used machine learning methods to identify depression and they have found that social media presents a persons' mentality. In this research I'm planning to be used machine learning techniques to predict whether that person has depression or not.

Research questions:

1. Was mental health reduced and getting more depressed during a pandemic situation like COVID-19?
2. What is the impact on non – depressed and depressed people in a pandemic situation like COVID-19?
3. Which features are most appropriate to predict depression when using social media posts?

Once the prediction model was prepared, with the help of that model, will be able to predict the mental health of people. Whether they are in good mental health or not and that will present if there is any difference in mental health in pre-COVID and COVID time.

The main objectives of this study are as below:

1. Critical study about the problem domain – depression and it's behavior
2. Identification of a technology
3. Design a model
4. Evaluate the solution

3.2 Proposing model or design

The main objectives in this research study are to be doing critical analysis about the depression, depression detection, machine learning, identifying the technology, design and develop the prototype and evaluate the solution.

Research questions:

1. 1. Was mental health reduced and getting more depressed during a pandemic situation like COVID-19?
2. What is the impact on non – depressed and depressed people in a pandemic situation like COVID-19?
3. Which features are most appropriate to predict depression when using social media posts?

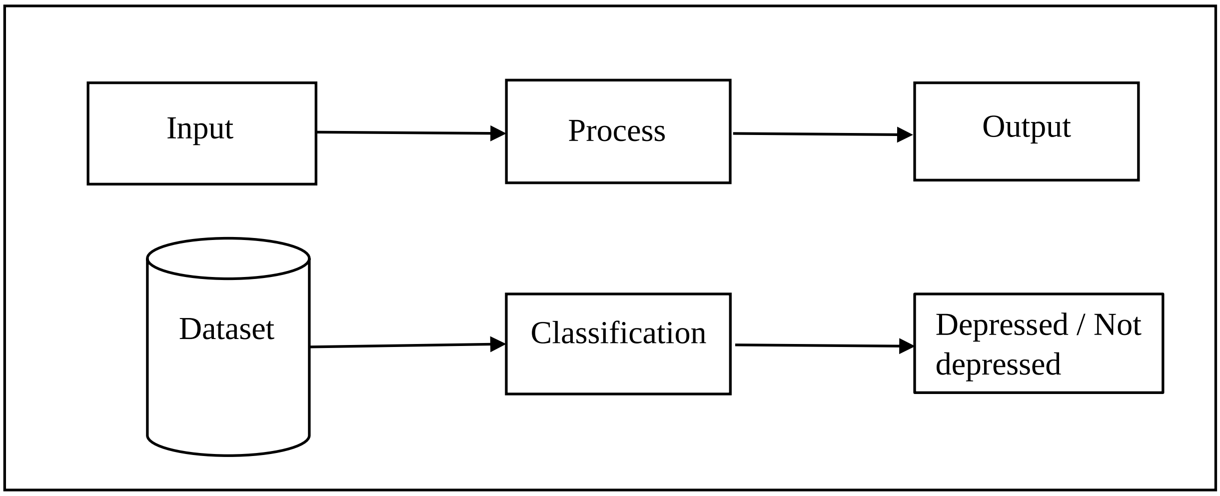


Figure 1 : Basic Flow

For this study I have used machine learning techniques to predict the model. Depression language, emotion and linguistic style are used as the features. Inputs for the system will be data acquired through the Facebook. Mainly there are two datasets. Facebook will be the resource to gather test data. One dataset will be used to train and test the model. And the other data set will be used to test the model and prove the research questions. The first dataset is acquired from online available dataset and that dataset is in Russian language. Due to that I have used Google translate to translate that dataset into English. In that dataset the posts were labelled by a psychiatrist. Then another few posts were acquired from the Facebook which are posted in the public pages reserved mainly to the people who felt in depression. That dataset was labelled as depression indicative posts. Non-depression indicative posts will be taken from the randomly selected publicly shared pages. Those are mainly the pages which shared in public and most of them are about positive thoughts, news of day-to-day life etc. The other dataset will be collected from the participants who are in IT industry and doing work from home after the COVID. For that corpus, screening test was taken place to measure whether they are fell in depression or not. The score of that questionnaire will reveal the mental health of that person.

Features:

1. Depression language – Depression symptoms, words used in depression medical field
2. Emotion – Positive affect, Negative affect, sadness affect, anger affect, anxiety affect
3. Linguistic style – I, prepositions, adverbs, pronouns, conjunctions, articles, auxiliary verbs, verbs and negations, personal pronouns, impersonal pronouns
4. Sentiment

As an initial step data will be normalized by using python program and then proceed with the feature extraction. For the depression language feature, the data will be taken through simple python program, and it will identify whether the text contains one of the words define in the depression language features. If text contains that word, those were labelled as 'True' (depression indicative post) and else, text is labelled with the value 'False' (Not a depression indicative post). For the emotion feature and the linguistic style values will be acquired from the Linguistic Inquiry and Word Count (LIWC) software. It will give the percentage values and the basic behind that is, it reads the text and calculates the percentage of total words that match each of the built-in dictionary categories. Once the feature extraction was finished, the model will be trained and using the test data set model will be validated. Then evaluate the model performance by using confusion matrix. Then the model will be test with the second data set. Then will be able to compare the results of classification whether there is any impact on mental health by COVID pandemic situation.

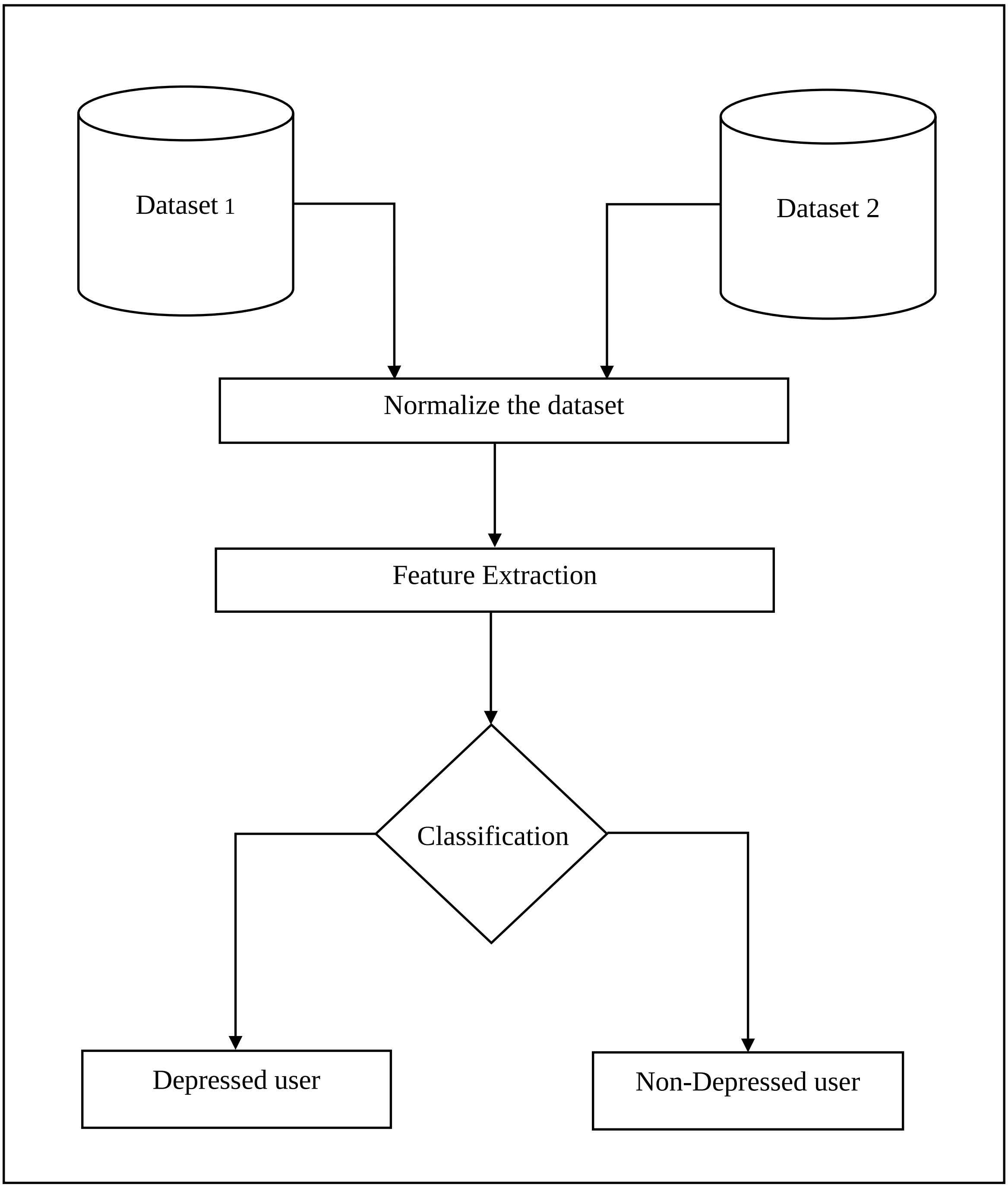


Figure 2: High Level Architecture

3.3 Methodology

This section will present the methodology used to solve the problem.

In this study I first focused on the four types of factors such as depression language, emotion and linguistic style, language and all the four features (depression language, emotion, linguistic style, language) together to detect the depressive data received as Facebook posts. Then I have applied supervised machine learning approaches to study about each attribute independently.

3.3.1 Data set exploration

I have used an online benchmark dataset compiled mainly from social networks used by young people in the Commonwealth of Independent States (CIS) countries (Narynov, 2020). Psychologists categorized that dataset into two categories: depression and depression. That dataset was in the Russian language. Since I need the data in English, I have used Google Translate to translate the database into English. I have excerpted the remaining posts (part of a dataset of non-depression indicated posts and depression indicated posts) from Facebook pages shared publicly. As there are several new words in the language dictionary, such as COVID, pandemic, vaccine, and variants, the excerpted section was integrated with the benchmark dataset.

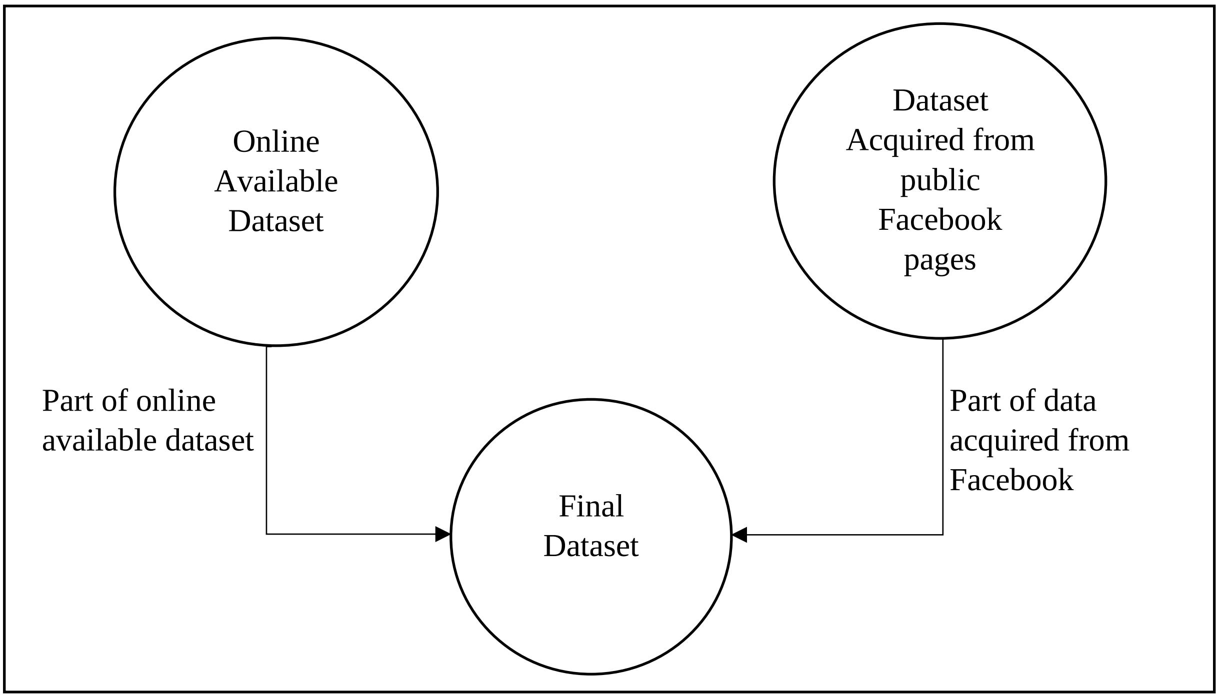


Figure 3: Main Dataset

As the second dataset I have collected a dataset which contains Facebook posts of the people who are in IT related fields and who are doing work from home. The main reason behind that selection is due to the work from home those people are work till late nights without any limitation and some of the people were experienced salary cut downs as well job losses. Due to those reasons, there is a tendency for that user group to fall into depression. Also, I have sent a google form to that user group to get consent before accessing their Facebook posts. Once I received their consent, I have accessed their Facebook accounts and get the Facebook posts in between time October - 2019 to July – 2020. That period was chosen to cover the period before COVID (October – 2019 to February - 2020) and COVID period (March – 2020 to July - 2020).

3.3.2 Data set preparation

After collecting the data from Facebook, it was analyzed using 'Linguistic Inquiry and Word Count (LIWC)' software. Before analyzing the data, I have cleaned and pre-process the dataset. Initially, I have removed the memes and some misleading posts from the acquired Facebook posts due to those can mislead the model. For pre-process I have used simple python program written in Google Colaboratory. Google Colaboratory, often referred to as Colab, is a product of Google Research that allows people to collaborate on projects ("Colaboratory – Google," n.d.). It enables users to create and run arbitrary Python code directly from their web browser, and it is particularly well suited for machine learning, data analysis, and educational purposes ("Colaboratory – Google," n.d.). In addition, Colab is a hosted Jupyter notebook service that does not require any installation and provides free access to computing resources (including GPUs). ("Colaboratory – Google," n.d.).

Post	
1	My name is Fox (I don't want to write my real name).
2	Good Afternoon…!!
3	My name is N, I'm 27.
4	Facebook is like the Fridge. If you are bored, you keep opening and closing it every few minutes to see if there is anything good in it.
5	Hi, my name is let's say oh.

Table 1: Posts removed while cleaning and pre-processing dataset

After cleaning and pre-processing the dataset, it was analyzed by using LIWC. It gets the text and counts the percentage of words that reflect emotions, thinking styles and parts of speech. In LIWC text analysis module contains built-in dictionaries. Once user enters the text it transcribes the text to computer readable form. Then it compares each word in the text with the built-in dictionary. This process was done by the text analysis module. Once the reading was finished by the processing module for the user given text it calculates the percentage of total words that match each of the dictionary categories.

My primary dataset contains totally 17 columns as below. For linguistic style I have used 12 columns (I, prepositions, adverbs, pronouns, conjunctions, articles, auxiliary verbs, verbs and negations, personal pronouns, impersonal pronouns, pronouns, adjectives). For emotion I have used 5 columns (negative effect, positive effect, sadness affect, anger and anxiety affect).

Then for the Depression language feature I have used set of words such as different depression symptoms like helplessness, hopelessness, lack of sleep, anger, insomnia, irritate, pain, fatigue, anxious, bad mood, concentrate. Also, I have used negative words people used when in pressurized situations like abuse, cry, anxiety, blues, broke, die, dead, depression, dilemma, disappoint, feel bad, hate, mental disorder, sad, stress, suicide, terrible, tired, unfair. If those words were identified in the text those were labelled with True (Depressed) and if not labelled those with False (Not Depressed).

For the sentiment analysis I have used python program to find the sentiment of each text. Sentiment was analyzed by using the polarity of the text and other than that I have taken subjectivity of that text. Subjectivity measures the how subjective the text and polarity measures the how negative or positive the text is. Here the negative posts were labelled as True (Depressed). Neutral and positive posts were labelled as False (Not Depressed).

3.3.3 Building ground truth dataset

I have taken the depressed patients social media content as depression indicative posts. Then those posts were labelled as True (depression indicative) and non-depressive posts were labelled as False (non-depression indicative). Out of 37725 posts 19226 posts obtained True (Depressed) and rest of the corpus (18499) obtained False (Not Depressed).

3.3.4 Depression screening test

This screening test was done on the set of people who were doing work from home and jobs are related to the IT field. The purpose of this screening is to validate the results obtained for the second dataset. I have used the Center for Epidemiologic Studies Depression Scale (CES - D) questionnaire as the primary tool to determine whether the participant has depression or not. That questionnaire contains 20 questions and was designed to measure depression symptoms in the general population. This was invented by a person named Radloff in 1977. Also, this questionnaire is one of the most used screening tests by clinicians and psychiatrists. It quantifies the depressive feelings and behaviors during the past week. For example, the test seeks responses to questions such as ‘I felt fearful’, ‘My sleep was restless’ and participants were asked to choose one of the following responses to each question; (1) Rarely or none of the time, (2) Some or little of the time, (3) Occasionally or a moderately amount of the time, (4) Most or all of the time. The maximum score is 60 and minimum is 0 and a score of 16 or more is considered depression.

3.3.5 Tackling noisy responses

Since there could be noisy responses in the screening test, I have deployed an auxiliary screening test additional to the CES-D. This can be happened due to intentionally or unintentionally by the respondent. I have used Patient Health Questionnaire – 9 (PHQ-9) for this purpose. PHQ is a question instrument, and it has modules on mood, eating, anxiety, etc. The PHQ-9 contains questions about mood. It was invented by Dr. Kurt Kroenke and the team in 1990. This questionnaire also used in Sri Lanka as a screening tool by psychiatrists. The aim behind this is to get high-quality responses and the scores in PHQ – 9 and CES – D would correlate.

Individuals who are scored high in both tests (CES-D and PHQ-9) are considered truly depressed. Individuals who get totally different results for two tests are removed from the research respondents' list.

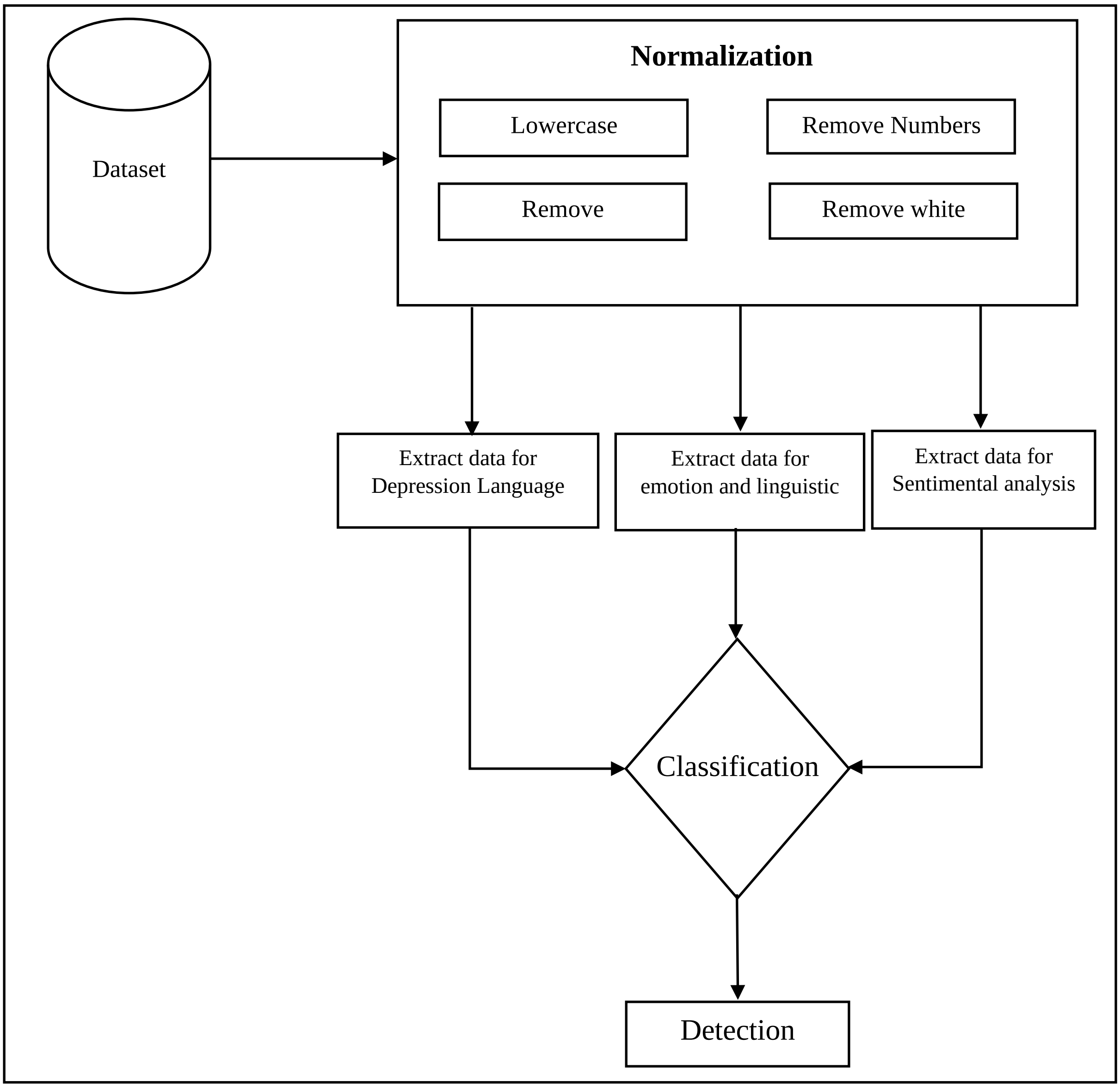


Figure 4: Flow of the model generation

3.3.6 Feature extraction

To describe these Facebook posts (depressive and non-depressive), I have extracted the different features in view of psycholinguistic measurements. For that I have used Linguistic Inquiry and Word Count (LIWC) and from that I have acquired data which needs to verify the linguistic styles. LIWC returns different higher levels of psycholinguistic features:

- Psychological process – Social process, Cognitive process, biological process, relativity, personal concerns, affective process, perceptual process
- Linguistic Process – word count, pronoun, personal pronoun, articles, prepositions, adverbs, auxiliary verbs, conjunctions
- Other grammar – verbs, comparisons, adjectives

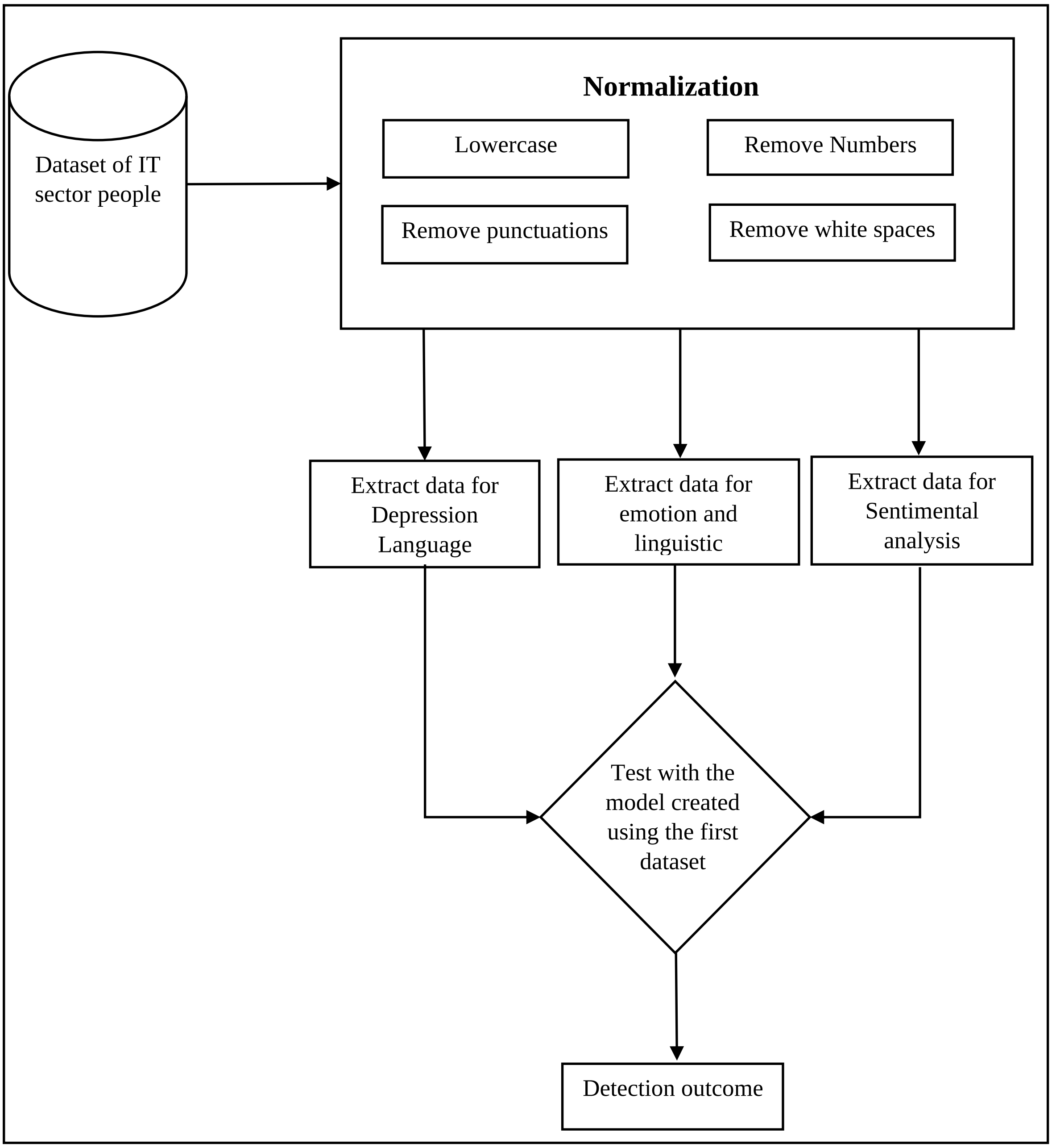


Figure 5: Flow of the model testing using dataset of the IT sector

Above levels divided into subcategories as below

- Biological process – Sexual, body, ingestion, health
- Affective processes – anger, sadness, anxiety, negative and positive emotions
- Time orientation – Future, Present, Past
- Perceptual processes – see, feel, hear
- Social processes – male, female, friends, family

For my research I have used 17 among 70 factors.

3.3.7 Sentimental analysis

Sentimental analysis is a method used to determine whether text is positive, negative or neutral. This is a natural language processing technique which is often performed on textual data. There are different types of sentimental analysis methods based on polarity, feelings and emotions, urgency and intension. Emotion detection systems are used lexicons or machine learning algorithms. Lexicons are lists of words and emotions they convey. In this study I have used sentimental analysis to analyze the Facebook posts and based on polarity and subjectivity. I have created a simple python program to obtain the sentiment of each text. It retrieves whether the text is negative or positive using the polarity value of each text.

3.3.8 Measure depressive behavior

I have used a set of attributes like depression language, emotion, sentiment, and linguistic style to measure the depressive behavior of users. Those attributes could be used to characterize the depressive behavior of users. As emotional variables I have used positive affect, negative affect, sadness affect, anger affect, anxiety affect. As linguistic styles, I have used I, prepositions, adverbs, pronouns, personal pronouns, impersonal pronouns, conjunctions, articles, auxiliary verbs, adjectives, verbs, and negations. For the depression language attribute, I have used words such as depression, anxiety, hate, depression symptoms (fatigue, lack of sleep, helpless, hopeless, etc.), bad mood, terrible mood, insomnia, broke, cry, pain, blues, anger, irritate, concentrate. For the attributes of emotion and linguistic style calculated the values using Linguistic Inquiry and Word Count (LIWC) scale. For the depression language attribute, I have used simple code written using python to identify whether the text contains that word. If the text contains that word it marked as True (depression indicative) and otherwise it marked as False (Not a depression indicative). Also, I have done a sentimental analysis using python. Those programming parts were done on the Google Colab platform which is especially suitable for data analysis and machine learning.

Emotion

A complicated interplay between cognitive awareness, body feeling, and conduct that reflects one's significance for a particular thing, event, or condition of circumstances are characteristics of the emotion process (Islam et al., 2018). The following are some of the most prevalent emotions people experience when depressed: sadness, guilt, irritation, and so on. It is possible to make reliable forecasts in some situations by analyzing the emotional remarks made on social network data (Shen et al., 2017).

Most of the time emotions controlled the people. Emotions have influenced every person's decision, every action, and every perception they are currently experiencing. In the 1970s, psychologist Paul Eckman developed a theory of six fundamental emotions universally experienced across all cultural boundaries(Facebook and Twitter, n.d.) . Happiness, sadness, disgust, fear, surprise, and fury were among the feelings experienced. Later, the list of basic emotions was broadened to include pride, humiliation, embarrassment, and exhilaration, among other things (Facebook and Twitter, n.d.).

Here I have used psycholinguistic dimensions for considering few features of the emotion state manifested in the posts as a negative effect, positive effect, sadness affect, anger affect, and anxiety affect. These values were computed with the help of Linguistic Inquiry and Word Count (LIWC) software which is the psycholinguistic resource used in text analysis (“LIWC,” n.d.). In LIWC, it reads a given text by the user and counts the percentage of words that reflect different thinking styles, emotions, parts of speech, etc (“LIWC,” n.d.).

Linguistic process

People suffering from depression are using more first-person singular pronouns – such as I, me, and myself – than the general population (Al-Mosaiwi, n.d.). In patients with depression, there is a tendency to use second and third person pronouns, such as her, him, and them. This pattern of pronoun usage suggests that people suffering from depression are more focused on themselves and less connected with others than the general population. A team of researchers found that pronouns are more trustworthy than negative emotion terms when it comes to identifying depression.

The LIWC psycholinguistic vocabulary package consists of several components, and the language process is one of the most important of these components. The main goal of this section is to measure the usage of words in cognitively important classification systems. Linguistic process has been effectively used to identify connections between people in social co-operations (relative status, trickiness) and the nature of close relationship. In this study I have used ten linguistic features (I, prepositions, adverbs, pronouns, conjunctions, articles, auxiliary verbs, verbs and negations, personal pronouns, impersonal pronouns) to characterize user Facebook posts.

Classification model

In this phase constructs prediction model for depression post recognition, by considering the psycholinguistic features as the input. Considering the training dataset each post is labelled with the class either as depressive or non-depressive.

In this study I have employed ten classifiers: Lazy, Bayesian, Trees, Rules, Neural Nets, Logistic Regression, Support Vector Machine, Ensembles.

Lazy classification is a method in which generalization of the training data is, delayed until a query is made to the system, where the system tries to generalize the training data before receiving queries. The primary motivation for employing lazy learning, as in the K-nearest neighbors' algorithm, used by online recommendation systems is that the data set is continuously updated with new entries. Because of the continuous update, the "training data" would be rendered obsolete in a relatively ("Lazy learning," 2020).

Naïve Bayes classifiers are based on the Bayes theorem. The fundamental of the Naïve Bayes theorem is that each feature gives independent and equal contribution for the end results. Bayes' theorem finds the probability of an event occurring given the probability of another event that has already occurred ("Naive Bayes Classifiers," 2017).

Tree classifier models give high accuracy due to those not like linear models. Decision tree is a simple and mostly used classification based systematic approach that makes the hierarchical tree using the training data set. The state of decision tree is to divide the data hierarchically according to the characteristics. In text documents classification, roots are commonly identified in terms and internal individual nodes may be sub divided to its child in view of the yes or no of a term in the document (Islam et al., 2018).

Rule – based classifiers are another type of classifier which makes the class decision depending by using 'IF ELSE' rules. These rule-based classifiers are generally used to generate descriptive models. The condition used with "IF" is called the antecedent and the predicted class of that rule is named as the consequent (“Rule-Based Classifier - Machine Learning,” 2020). The coverage of rule-based classifier is the percentage of data records which satisfy the antecedent conditions of a rule. The rules generated by this classifier type is not mutually exclusive. Also, there may be some instances where some of the records are not covered with the rules. The decision boundaries created by rule-based classifiers are linear. But this classifier model is complex than the decision tree model due to many rules are triggered for the same record.

The logistic regression model is a simply models probability of output in terms of input and does not perform statistical classification (it is not a classifier), though it can be used to make a classifier, for instance by choosing a cutoff value and classifying inputs with probability greater than the cutoff as one class, below the cutoff as the other; this is a common way to make a binary classifier.

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable. In regression analysis, logistic regression is estimating the parameters of a logistic model. Normally a binary logistic model has a dependent variable with two possible values, such as true/falls which is represented by an indicator variable, where the two values are labeled 0/1. In the logistic model, the log-odds for the value labeled "1" is a linear combination of one or more independent variables; the independent variables can each be a binary variable (or a continuous variable (“Logistic regression,” 2021).

Support Vector Machine is a supervised machine learning algorithm which can be used for both classification or regression. But, it is mostly used in classification problems. The main aim of the SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. SVM classifiers are efficiently perform for the non-linear classification problems. Naturally, a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class which is known as functional margin. The generalization error of the classifier is getting low when the margin gets larger (“Support-vector machine,” 2021).

Neural networks are a set of algorithms, modeled loosely and designed to recognize patterns. They interpret sensory data through a kind of machine perception, clustering or labelling raw input. Neural networks help to cluster and classify. Neural nets are helping to group unlabeled data according to similarities among the example inputs, and they classify data when they have a labeled dataset to train on. Neural networks can also extract features that are fed to other algorithms for clustering and classification; since can think of deep neural networks as components of larger machine-learning applications involving algorithms for reinforcement learning, classification and regression (“A Beginner’s Guide to Neural Networks and Deep Learning,” n.d.).

CHAPTER 4

EVALUATION AND RESULTS

4.1 Results

The analysis was conducted using the RapidMiner Studio Version 9.8. I have applied six major classifiers: Lazy, Bayesian, Trees, Rules, Neural Nets and Logistic Regression. Each classifier has sub-classifiers as mentioned below:

1) Lazy
   a) K-NN
2) Naïve Bayes
   a) Naïve Bayes
   b) Naïve Bayes (Kernel)
3) Trees
   a) Decision Tree
   b) Random Forest
   c) Gradient Boosted Tree
   d) Decision Stump
   e) Random Tree
4) Rules
   a) Rule Induction
   b) Single Rule Induction (Single Attribute)
5) Neural nets
   a) Deep Learning
6) Logistic Regression
   a) Logistic Regression

Using the above classification techniques, I have examined detection performance of social media posts. The results of analysis are reported in Table 2 to 31.

Precision, recall and F-measure have been used as evaluation matrices parameters and those are used to evaluate these classifiers. It has conducted four different ways.

1) TP or True Positive: the depression cases that are positive and predicted as positive
2) TN or True Negative: the depression cases that are negative and predicted as negative
3) FN or False Negative: the depression cases are positive but predicted as negative
4) FP or False Positive: the depression cases are negative but predicted as positive

All the evaluation matrices are defined as follows:

$$Precision(P) = \frac{True\ Positive\ (TP)}{True\ positive\ (TP) + False\ Positive\ (FP)}$$

Precision is the proportion of TP to the cases that are predicted as positive.

Recall is the proportion of true positives to the cases that are truly positive.

$$Recall(R) = \frac{True\ Positive\ (TP)}{True\ positive\ (TP) + False\ Negative\ (FN)}$$

F-measure is the mean of precision and Recall. It takes both false negatives and false positives into a record. F-measure is calculated as:

$$F - measure = 2\frac{PR}{(P + R)}$$

Also, in here accuracy is not considered due to the dataset is imbalanced.

Emotion feature

Classifier	Sub Classifier	Train Dataset			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	49.13	53.01	**96.46**	68.42
Bayesian	Naïve Bayes	47.40	62.23	93.05	74.59
	Naïve Bayes(kernel)	32.83	97.79	64.45	77.70
Trees	Decision Tree	39.67	90.72	77.88	83.81
	Random Forest	40.72	**90.14**	79.94	**84.73**
	Gradient Boosted Tree	39.77	92.25	78.09	84.58
	Decision Stump	31.11	90.69	61.07	72.99
	Random Tree	31.11	90.69	61.07	72.99
Rules	Rule Induction	40.41	86.13	79.34	82.60
	Single Rule Induction (Single attribute)	39.39	82.62	77.34	79.89
Neural Nets	Deep Learning	36.54	93.58	71.74	81.22
Logistic Regression	Logistic Regression	41.01	61.67	80.51	69.84

Table 2: Emotion – Train Dataset – Positive Class

Table 2 contains the results that occurred while testing the trained model for emotion feature. Moreover, that is for the positive class, which contains depression indicated posts. The highest F-measure, 84.73, was obtained for the Random Forest classification method, and it comes under the Trees classifier.

Table 3 contains the results that occurred while testing the trained model for emotion feature. Moreover, that is for the negative class, which contains not depression indicated posts. The highest F-measure, 86.31, was obtained for the Gradient Boosted Tree classification method, and it comes under the Trees classifier.

Table 4 contains the results that occurred while testing the second dataset using the developed model for emotion feature. Moreover, that is for the positive class that contains depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 45.13, was obtained for the Rule induction classification method under the Rules classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	5.51	75.36	11.24	19.56
Bayesian	Naïve Bayes	20.30	**85.16**	41.38	55.70
	Naïve Bayes(kernel)	48.32	72.75	**98.49**	83.68
Trees	Decision Tree	45.01	79.98	91.73	85.46
	Random Forest	44.61	81.36	90.92	85.88
	Gradient Boosted Tree	45.73	80.38	93.19	**86.31**
	Decision Stump	45.87	69.82	93.49	79.94
	Random Tree	45.87	69.82	93.49	79.94
Rules	Rule Induction	42.56	80.17	86.74	83.33
	Single Rule Induction (Single attribute)	40.78	77.94	83.12	80.44
Neural Nets	Deep Learning	46.56	76.39	94.89	84.64
Logistic Regression	Logistic Regression	23.58	70.37	48.06	57.11

Table 3: Emotion – Train Dataset – Negative Class

Classifier	Sub Classifier	Pre-COVID			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	8.13	9.38	**89.71**	16.99
Bayesian	Naïve Bayes	6.67	10.87	73.53	18.94
	Naïve Bayes(kernel)	1.60	23.08	17.65	20.00
Trees	Decision Tree	4.40	29.20	48.53	36.46
	Random Forest	5.20	31.45	57.35	40.63
	Gradient Boosted Tree	3.20	20.51	35.29	25.95
	Decision Stump	3.87	**40.28**	42.65	41.43
	Random Tree	3.87	**40.28**	42.65	41.43
Rules	Rule Induction	5.87	34.65	64.71	**45.13**
	Single Rule Induction (Single attribute)	4.93	22.16	54.41	31.49
Neural Nets	Deep Learning	3.87	23.39	42.65	30.21
Logistic Regression	Logistic Regression	7.47	20.22	82.35	32.46

Table 4: Emotion – Pre-COVID Dataset – Positive Class

Table 5 contains the results that occurred while testing the second dataset using the developed model for emotion feature. Moreover, that is for the negative class, which contains not depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers before the COVID period). The highest F-measure, 93.04, was obtained for the Naïve Baye (Kernel) classifier, which comes under the Bayesian classifier.

Table 6 contains the results occurred while testing the second dataset using developed model for emotion feature. And that is for the positive class which contains depression indicated posts and using while COVID period dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure 56.32 was obtained for the Rule induction classifier method, and it comes under the Rules classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	12.40	93.00	13.64	23.79
Bayesian	Naïve Bayes	36.27	93.79	39.88	55.97
	Naïve Bayes(kernel)	85.60	91.98	**94.13**	**93.04**
Trees	Decision Tree	80.27	94.51	88.27	91.28
	Random Forest	79.60	95.37	87.54	91.28
	Gradient Boosted Tree	78.53	93.05	86.36	89.58
	Decision Stump	85.20	94.25	93.70	93.97
	Random Tree	85.20	94.25	93.70	93.97
Rules	Rule Induction	79.87	96.15	87.83	91.80
	Single Rule Induction (Single attribute)	73.60	94.68	80.94	87.27
Neural Nets	Deep Learning	78.27	93.77	86.07	89.76
Logistic Regression	Logistic Regression	61.47	**97.46**	67.60	79.83

Table 5: Emotion – Pre - COVID Dataset – Negative Class

Classifier	Sub Classifier	IN-COVID			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	15.47	17.34	**91.34**	29.15
Bayesian	Naïve Bayes	12.40	18.86	73.23	30.00
	Naïve Bayes(kernel)	2.67	43.48	15.75	23.12
Trees	Decision Tree	8.80	51.16	51.97	51.56
	Random Forest	9.87	51.39	58.27	54.61
	Gradient Boosted Tree	6.93	37.68	40.94	39.25
	Decision Stump	7.73	**59.79**	45.67	51.79
	Random Tree	7.73	**59.79**	45.67	51.79
Rules	Rule Induction	10.40	52.00	61.42	**56.32**
	Single Rule Induction (Single attribute)	8.80	37.29	51.97	43.42
Neural Nets	Deep Learning	6.27	32.87	37.01	34.81
Logistic Regression	Logistic Regression	13.87	29.89	81.89	43.79

Table 6: Emotion – In-COVID Dataset – Positive Class

Table 6 contains the results that occurred while testing the second dataset using the developed model for emotion feature. Moreover, that is for the positive class, which contains depression indicated posts and using while COVID period dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 56.32, was obtained for the Rule induction classifier method, and it comes under the Rules classifier.

Table 7 contains the results that occurred while testing the second dataset using the developed model for emotion feature. Moreover, the negative class that contains not depression indicated posts and using while COVID period (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 91.54, was obtained for the Decision stump and random tree classifier methods, and it comes under the Trees classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	9.33	86.42	11.24	19.89
Bayesian	Naïve Bayes	29.73	86.77	35.79	50.68
	Naïve Bayes(kernel)	79.60	84.80	**95.83**	89.98
Trees	Decision Tree	74.67	90.18	89.89	90.03
	Random Forest	73.73	91.25	88.76	89.99
	Gradient Boosted Tree	71.60	87.75	86.20	86.96
	Decision Stump	77.87	89.43	93.74	**91.54**
	Random Tree	77.87	89.43	93.74	**91.54**
Rules	Rule Induction	73.47	91.83	88.44	90.11
	Single Rule Induction (Single attribute)	68.27	89.35	82.18	85.62
Neural Nets	Deep Learning	70.27	86.82	84.59	85.69
Logistic Regression	Logistic Regression	50.53	**94.28**	60.83	73.95

Table 7: Emotion – In-COVID Dataset –Negative Class

Sentiment feature

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	49.22	51.58	**96.64**	67.26
Bayesian	Naïve Bayes	45.06	65.20	88.47	75.07
	Naïve Bayes (Kernel)	44.37	66.23	87.12	75.25
Trees	Decision Tree	42.36	70.24	83.16	76.16
	Random Forest	43.14	68.82	84.70	75.94
	Gradient Boosted Tree	33.65	**79.64**	66.07	72.22
	Decision Stump	46.36	61.86	**91.02**	73.66
	Random Tree	46.36	61.86	**91.02**	73.66
Rules	Rule Induction	40.04	73.99	78.61	**76.23**
	Single Rule Induction (Single Attribute)	39.06	70.77	76.68	73.61
Neural Nets	Deep Learning	36.91	75.03	72.47	73.73
Logistic Regression	Logistic Regression	46.60	60.99	**91.49**	73.19

Table 8: Sentiment – Train Dataset – Positive Class

Table 8 contains the results that occurred while testing the trained model for sentiment feature. Moreover, that is for the positive class, which contains depression indicated posts. The highest F-measure, 76.23, was obtained for the Rule Induction classification method, and it comes under the Rules classifier.

Table 9 contains the results that occurred while testing the trained model for sentiment feature. Moreover, that is for the negative class, which contains not depression indicated posts. The highest F-measure, 75.77, was obtained for the Gradient Boosted Tree classification method, and it comes under the Trees classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	2.86	62.61	5.83	10.67
Bayesian	Naïve Bayes	25.01	80.99	50.97	62.57
	Naïve Bayes (Kernel)	26.44	80.12	53.89	64.44
Trees	Decision Tree	31.12	78.40	63.43	70.12
	Random Forest	29.52	79.11	60.16	68.34
	Gradient Boosted Tree	40.46	70.07	**82.47**	**75.77**
	Decision Stump	20.48	**81.75**	41.73	55.26
	Random Tree	20.48	**81.75**	41.73	55.26
Rules	Rule Induction	34.99	76.26	71.31	73.70
	Single Rule Induction (Single Attribute)	32.94	73.50	67.13	70.17
Neural Nets	Deep Learning	36.78	72.40	74.96	73.66
Logistic Regression	Logistic Regression	19.26	81.63	39.25	53.01

Table 9: Sentiment – Train Dataset – Negative Class

Table 10 contains the results that occurred while testing the second dataset using the developed model for sentiment feature. Moreover, that is for the positive class that contains depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 47.73, was obtained for the Rule induction classification method under the Rules classifier

Classifier	Sub Classifier	Pre-COVID			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	15.33	16.27	**95.04**	27.78
Bayesian	Naïve Bayes	15.73	23.79	**97.52**	38.25
	Naïve Bayes(kernel)	14.27	22.43	88.43	35.79
Trees	Decision Tree	12.13	22.92	75.21	35.14
	Random Forest	12.13	22.69	75.21	34.87
	Gradient Boosted Tree	8.93	30.73	55.37	39.53
	Decision Stump	15.07	20.77	**93.39**	33.98
	Random Tree	15.07	20.77	**93.39**	33.98
Rules	Rule Induction	10.53	**37.62**	65.29	**47.73**
	Single Rule Induction (Single attribute)	10.67	22.86	66.12	33.97
Neural Nets	Deep Learning	10.40	34.06	64.46	44.57
Logistic Regression	Logistic Regression	16.13	21.27	100.00	35.07

Table 10: Sentiment – Pre - COVID Dataset – Positive Class

Table 11 contains the results that occurred while testing the second dataset using the developed model for sentiment feature. Moreover, that is for the negative class, which contains not depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers before the COVID period). The highest F-measure, 85.20, was obtained for the Rule Induction classifier, which comes under the Rules classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	4.93	86.05	5.88	11.01
Bayesian	Naïve Bayes	33.47	98.82	39.90	56.85
	Naïve Bayes (Kernel)	34.53	94.87	41.18	57.43
Trees	Decision Tree	43.07	91.50	51.35	65.78
	Random Forest	42.53	91.40	50.72	65.24
	Gradient Boosted Tree	63.73	89.85	75.99	82.34
	Decision Stump	26.40	96.12	31.48	47.43
	Random Tree	26.40	96.12	31.48	47.43
Rules	Rule Induction	66.40	92.22	**79.17**	**85.20**
	Single Rule Induction (Single Attribute)	47.87	89.75	57.07	69.78
Neural Nets	Deep Learning	63.73	91.75	75.99	83.13
Logistic Regression	Logistic Regression	24.13	**100.00**	28.78	44.69

Table 11: Sentiment – Pre - COVID Dataset – Negative Class

Table 12 contains the results occurred while testing the second dataset using developed model for sentiment feature. And that is for the positive class which contains depression indicated posts and using while COVID period dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure 55.83 was obtained for the deep learning classifier method, and it comes under the neural nets classifier.

Classifier	Sub Classifier	IN-COVID			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	17.60	18.38	**94.96**	30.81
Bayesian	Naïve Bayes	17.73	28.73	**95.68**	44.19
	Naïve Bayes(kernel)	16.67	27.90	89.93	42.59
Trees	Decision Tree	14.67	28.35	79.14	41.75
	Random Forest	14.53	27.25	78.42	40.45
	Gradient Boosted Tree	10.67	38.10	57.55	45.85
	Decision Stump	17.07	24.29	92.09	38.44
	Random Tree	17.07	24.29	92.09	38.44
Rules	Rule Induction	12.93	**40.76**	69.78	51.46
	Single Rule Induction (Single attribute)	10.80	24.25	58.27	34.25
Neural Nets	Deep Learning	15.33	**42.12**	82.73	**55.83**
Logistic Regression	Logistic Regression	18.53	25.93	**100.00**	41.19

Table 12: Sentiment – In - COVID Dataset – Positive Class

Table 13 contains the results that occurred while testing the second dataset using the developed model for emotion feature. Moreover, the negative class that contains not depression indicated posts and using while COVID period (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 83.70, was obtained for the Rule induction classifier methods, and it comes under the Rules classifier.

Classifier	Sub Classifier	IN-COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	3.33	78.13	4.09	7.78
Bayesian	Naïve Bayes	37.47	97.91	45.99	62.58
	Naïve Bayes(kernel)	38.40	95.36	47.14	63.09
Trees	Decision Tree	44.40	91.99	54.50	68.45
	Random Forest	42.67	91.43	52.37	66.60
	Gradient Boosted Tree	64.13	89.07	**78.72**	83.58
	Decision Stump	28.27	95.07	34.70	50.84
	Random Tree	28.27	95.07	34.70	50.84
Rules	Rule Induction	62.67	91.80	76.92	**83.70**
	Single Rule Induction (Single attribute)	47.73	86.06	58.59	69.72
Neural Nets	Deep Learning	60.40	94.97	74.14	83.27
Logistic Regression	Logistic Regression	28.53	**100.00**	35.02	51.88

Table 13: Sentiment – In – COVID Dataset – Negative Class

Language feature

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	50.93	50.93	**100.00**	67.49
Bayesian	Naïve Bayes	30.62	95.06	60.11	**73.65**
	Naïve Bayes (Kernel)	30.62	95.06	60.11	**73.65**
Trees	Decision Tree	27.05	97.98	53.11	68.88
	Random Forest	29.66	96.09	58.24	72.52
	Gradient Boosted Tree	27.90	96.87	54.77	69.98
	Decision Stump	11.48	**99.54**	22.53	36.75
	Random Tree	11.48	**99.54**	22.53	36.75
Rules	Rule Induction	50.93	50.93	**100.00**	67.49
	Single Rule Induction (Single Attribute)	50.93	50.93	**100.00**	67.49
Neural Nets	Deep Learning	30.71	94.49	60.29	73.61
Logistic Regression	Logistic Regression	30.60	95.10	60.08	73.64

Table 14: Language – Train Dataset – Positive Class

Table 14 contains the results that occurred while testing the trained model for language feature. Moreover, that is for the positive class, which contains depression indicated posts. The highest F-measure, 73.65, was obtained for the Naïve Bayes and Naïve Bayes (Kernel) classification methods, and it comes under the Bayesian classifier.

Table 15 contains the results that occurred while testing the trained model for language feature. Moreover, that is for the negative class, which contains not depression indicated posts. The highest F-measure, 81.26, was obtained for the Logistic Regression classification method.

Table 16 contains the results that occurred while testing the second dataset using the developed model for language feature. Moreover, that is for the positive class that contains depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 100, was obtained by few classification methods which are Naïve Bayes, Naïve Bayes (Kernel) and Deep learning. Those sub classifiers are relevant to the Bayesian and Neural nets main classifiers.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	0.00	0.00	0.00	0.00
Bayesian	Naïve Bayes	47.48	70.03	96.76	81.25
	Naïve Bayes (Kernel)	47.48	70.03	96.76	81.25
Trees	Decision Tree	48.51	67.01	98.87	79.88
	Random Forest	47.86	69.23	97.54	80.98
	Gradient Boosted Tree	48.16	67.65	98.16	80.10
	Decision Stump	49.01	55.40	**99.89**	71.27
	Random Tree	49.01	55.40	**99.89**	71.27
Rules	Rule Induction	0.00	0.00	0.00	0.00
	Single Rule Induction (Single Attribute)	0.00	0.00	0.00	0.00
Neural Nets	Deep Learning	47.28	**70.04**	96.35	81.11
Logistic Regression	Logistic Regression	47.49	70.02	96.79	**81.26**

Table 15: Language – Train Dataset – Negative Class

Classifier	Sub Classifier	Pre-COVID			
		Accuracy	**Precision**	**Recall**	**F-measure**
Lazy	K-NN	4.13	4.13	100.00	7.94
Bayesian	Naïve Bayes	4.13	**100.00**	**100.00**	**100.00**
	Naïve Bayes (Kernel)	4.13	**100.00**	**100.00**	**100.00**
Trees	Decision Tree	2.13	100.00	51.61	68.09
	Random Forest	2.40	100.00	58.06	73.47
	Gradient Boosted Tree	4.13	4.13	100.00	7.94
	Decision Stump	0.13	**100.00**	3.23	6.25
	Random Tree	4.13	4.13	**100.00**	7.94
Rules	Rule Induction	4.13	4.13	**100.00**	7.94
	Single Rule Induction (Single Attribute)	4.13	4.13	**100.00**	7.94
Neural Nets	Deep Learning	4.13	**100.00**	**100.00**	**100.00**
Logistic Regression	Logistic Regression	4.00	100.00	96.77	98.36

Table 16: Language – Pre - COVID Dataset – Positive Class

Table 17 contains the results that occurred while testing the second dataset using the developed model for language feature. Moreover, that is for the negative class, which contains not depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers before the COVID period). The highest F-measure, 100, was obtained for the Naïve Baye (Kernel) classifier, Naïve Bayes classifier and deep learning. Those sub classifiers are under the Bayesian classifier and Neural nets.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	0.00	0.00	0.00	0.00
Bayesian	Naïve Bayes	95.87	**100.00**	**100.00**	**100.00**
	Naïve Bayes (Kernel)	95.87	**100.00**	**100.00**	**100.00**
Trees	Decision Tree	95.87	97.96	**100.00**	98.97
	Random Forest	95.87	98.22	**100.00**	99.10
	Gradient Boosted Tree	0.00	0.00	0.00	0.00
	Decision Stump	95.87	95.99	**100.00**	97.96
	Random Tree	0.00	0.00	0.00	0.00
Rules	Rule Induction	0.00	0.00	0.00	0.00
	Single Rule Induction (Single Attribute)	0.00	0.00	0.00	0.00
Neural Nets	Deep Learning	95.87	**100.00**	**100.00**	**100.00**
Logistic Regression	Logistic Regression	95.87	99.86	**100.00**	99.93

Table 17: Language – Pre - COVID Dataset – Negative Class

Table 18 contains the results occurred while testing the second dataset using developed model for language feature. And that is for the positive class which contains depression indicated posts and using while COVID period dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure 99.13 was obtained for the naïve bayes, naive bayes(kernel) and logistic regression classifier method, and it comes under the Bayesian and logistic regression classifiers.

Classifier	Sub Classifier	IN-COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	7.73	7.73	100.00	14.36
Bayesian	Naïve Bayes	7.60	**100.00**	98.28	**99.13**
	Naïve Bayes (Kernel)	7.60	**100.00**	98.28	**99.13**
Trees	Decision Tree	3.73	**100.00**	48.28	65.12
	Random Forest	5.20	**100.00**	67.24	80.41
	Gradient Boosted Tree	7.73	7.73	**100.00**	14.36
	Decision Stump	0.13	**100.00**	1.72	3.39
	Random Tree	7.73	7.73	**100.00**	14.36
Rules	Rule Induction	7.73	7.73	**100.00**	14.36
	Single Rule Induction (Single Attribute)	7.73	7.73	**100.00**	14.36
Neural Nets	Deep Learning	7.20	**100.00**	93.10	96.43
Logistic Regression	Logistic Regression	7.60	**100.00**	98.28	**99.13**

Table 18: Language – In - COVID Dataset – Positive Class

Table 19 contains the results that occurred while testing the second dataset using the developed model for language feature. Moreover, the negative class that contains not depression indicated posts and using while COVID period (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 99.93, was obtained for the naïve bayes, naive bayes(kernel) and logistic regression classifier method, and it comes under the Bayesian and logistic regression classifiers.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	0.00	0.00	0.00	0.00
Bayesian	Naïve Bayes	92.27	**99.86**	**100.00**	**99.93**
	Naïve Bayes (Kernel)	92.27	**99.86**	**100.00**	**99.93**
Trees	Decision Tree	92.27	95.84	**100.00**	97.88
	Random Forest	92.27	97.33	**100.00**	98.65
	Gradient Boosted Tree	0.00	0.00	0.00	0.00
	Decision Stump	92.27	92.39	**100.00**	96.04
	Random Tree	0.00	0.00	0.00	0.00
Rules	Rule Induction	0.00	0.00	0.00	0.00
	Single Rule Induction (Single Attribute)	0.00	0.00	0.00	0.00
Neural Nets	Deep Learning	92.27	99.43	**100.00**	99.71
Logistic Regression	Logistic Regression	92.27	**99.86**	**100.00**	**99.93**

Table 19: Language – In - COVID Dataset – Negative Class

Linguistic style feature

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	45.34	84.57	89.02	86.74
Bayesian	Naïve Bayes	44.16	85.70	86.70	86.20
	Naïve Bayes (Kernel)	41.86	91.59	82.18	86.63
Trees	Decision Tree	40.81	**92.21**	80.12	85.74
	Random Forest	45.00	87.01	88.34	87.67
	Gradient Boosted Tree	43.55	91.00	85.51	**88.17**
	Decision Stump	47.24	79.48	**92.74**	85.60
	Random Tree	40.89	84.22	80.28	82.20
Rules	Rule Induction	44.52	87.52	87.41	87.46
	Single Rule Induction (Single Attribute)	45.54	82.50	89.41	85.81
Neural Nets	Deep Learning	42.12	91.22	82.70	86.75
Logistic Regression	Logistic Regression	44.17	79.06	86.73	82.71

Table 20: Linguistic Style – Train Dataset – Positive Class

Table 20 contains the results that occurred while testing the trained model for linguistic style feature. Moreover, that is for the positive class, which contains depression indicated posts. The highest F-measure, 88.17, was obtained for the Gradient Boosted Tree classification method, and it comes under the Trees classifier.

Table 21 contains the results that occurred while testing the trained model for linguistic style feature. Moreover, that is for the negative class, which contains not depression indicated posts. The highest F-measure, 88.45, was obtained for the Gradient Boosted Tree classification method, and it comes under the Trees classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	40.80	87.94	83.14	85.48
Bayesian	Naïve Bayes	41.70	86.03	84.98	85.50
	Naïve Bayes (Kernel)	45.22	83.28	92.17	87.50
Trees	Decision Tree	45.62	81.84	**92.98**	87.05
	Random Forest	42.35	87.70	86.30	87.00
	Gradient Boosted Tree	44.76	85.84	91.22	**88.45**
	Decision Stump	36.87	**90.89**	75.15	82.27
	Random Tree	41.40	80.47	84.39	82.38
Rules	Rule Induction	42.72	86.94	87.06	87.00
	Single Rule Induction (Single Attribute)	39.40	87.96	80.31	83.96
Neural Nets	Deep Learning	45.01	83.62	91.73	87.49
Logistic Regression	Logistic Regression	37.36	84.68	76.15	80.19

Table 21: Linguistic Style – Train Dataset – Negative Class

Classifier	Sub Classifier	Pre - COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	4.93	20.56	32.74	25.26
Bayesian	Naïve Bayes	5.47	19.52	36.28	25.39
	Naïve Bayes (Kernel)	3.60	14.06	23.89	17.70
Trees	Decision Tree	1.47	21.15	9.73	13.33
	Random Forest	3.47	25.74	23.01	24.30
	Gradient Boosted Tree	2.53	22.89	16.81	19.39
	Decision Stump	6.93	31.33	46.02	37.28
	Random Tree	12.00	19.91	**79.65**	31.86
Rules	Rule Induction	4.40	30.00	29.20	29.60
	Single Rule Induction (Single Attribute)	5.60	27.27	37.17	31.46
Neural Nets	Deep Learning	2.40	20.45	15.93	17.91
Logistic Regression	Logistic Regression	8.27	**36.69**	54.87	**43.97**

Table 22: Linguistic Style – Pre - COVID Dataset – Positive Class

Table 22 contains the results that occurred while testing the second dataset using the developed model for linguistic style feature. Moreover, that is for the positive class that contains depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 89.29, was obtained for the Decision Tree classification method under the Trees classifier.

Table 23 contains the results that occurred while testing the second dataset using the developed model for linguistic style feature. Moreover, that is for the negative class, which contains not depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers before the COVID period). The highest F-measure, 89.29, was obtained for the Decision Tree classifier, which comes under the Trees classifier.

Classifier	Sub Classifier	Pre - COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	65.87	86.67	77.55	81.86
Bayesian	Naïve Bayes	62.40	86.67	73.47	79.52
	Naïve Bayes (Kernel)	62.93	84.59	74.10	79.00
Trees	Decision Tree	79.47	85.39	**93.56**	**89.29**
	Random Forest	74.93	86.59	88.23	87.40
	Gradient Boosted Tree	76.40	85.91	89.95	87.88
	Decision Stump	69.73	89.55	82.10	85.67
	Random Tree	36.67	**92.28**	43.17	58.82
Rules	Rule Induction	74.67	87.50	87.91	87.71
	Single Rule Induction (Single Attribute)	70.00	88.09	82.42	85.16
Neural Nets	Deep Learning	75.60	85.65	89.01	87.30
Logistic Regression	Logistic Regression	70.67	91.22	83.20	87.03

Table 23: Linguistic Style – Pre-COVID Dataset – Negative Class

Classifier	Sub Classifier	IN-COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	7.07	27.75	33.97	30.55
Bayesian	Naïve Bayes	6.93	23.01	33.33	27.23
	Naïve Bayes (Kernel)	4.27	16.58	20.51	18.34
Trees	Decision Tree	3.47	50.98	16.67	25.12
	Random Forest	5.73	44.79	27.56	34.13
	Gradient Boosted Tree	3.33	34.72	16.03	21.93
	Decision Stump	9.33	40.94	44.87	42.81
	Random Tree	14.40	24.11	69.23	35.76
Rules	Rule Induction	5.47	39.81	26.28	31.66
	Single Rule Induction (Single Attribute)	7.33	36.18	35.26	35.71
Neural Nets	Deep Learning	3.60	35.53	17.31	23.28
Logistic Regression	Logistic Regression	12.93	**51.60**	62.18	**56.40**

Table 24: Linguistic Style – In-COVID Dataset – Positive Class

Table 24 contains the results occurred while testing the second dataset using developed model for linguistic style feature. And that is for the positive class which contains depression indicated posts and using while COVID period dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure 56.40 was obtained for the Logistic regression classifier method, and it comes under the Logistic Regression classifier.

Table 25 contains the results that occurred while testing the second dataset using the developed model for linguistic style feature. Moreover, the negative class that contains not depression indicated posts and using while COVID period (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 88.01, was obtained for the Decision tree classifier methods, and it comes under the Trees classifier.

Classifier	Sub Classifier	IN-COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	60.80	81.57	76.77	79.10
Bayesian	Naïve Bayes	56.00	80.15	70.71	75.13
	Naïve Bayes (Kernel)	57.73	77.74	72.90	75.24
Trees	Decision Tree	75.87	81.40	**95.79**	**88.01**
	Random Forest	72.13	82.72	91.08	86.70
	Gradient Boosted Tree	72.93	80.68	92.09	86.01
	Decision Stump	65.73	85.15	83.00	84.06
	Random Tree	33.87	84.11	42.76	56.70
Rules	Rule Induction	70.93	82.23	89.56	85.74
	Single Rule Induction Single Attribute)	66.27	83.11	83.67	83.39
Neural Nets	Deep Learning	72.67	80.86	91.75	85.96
Logistic Regression	Logistic Regression	67.07	**89.50**	84.68	87.02

Table 25: Linguistic Style – In-COVID Dataset – Negative Class

All (Emotion, Sentiment, Language, Linguistic Style)

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	46.43	85.98	91.15	88.49
Bayesian	Naïve Bayes	30.54	**95.36**	59.95	73.62
	Naïve Bayes (Kernel)	39.99	**96.61**	78.51	86.62
Trees	Decision Tree	42.74	95.27	83.92	89.24
	Random Forest	44.24	95.13	86.86	90.81
	Gradient Boosted Tree	44.23	93.39	86.83	89.99
	Decision Stump	47.24	79.48	**92.74**	85.60
	Random Tree	47.24	79.48	**92.74**	85.60
Rules	Rule Induction	44.88	89.06	88.11	88.58
	Single Rule Induction (Single Attribute)	45.54	82.50	89.41	85.81
Neural Nets	Deep Learning	45.99	94.52	90.29	**92.36**
Logistic Regression	Logistic Regression	44.67	88.61	87.69	88.15

Table 26: All features – Train Dataset – Positive Class

Table 26 contains the results that occurred while testing the trained model for all features. Moreover, that is for the positive class, which contains depression indicated posts. The highest F-measure, 92.36, was obtained for the Deep Learning classification method, and it comes under the Neural Nets classifier.

Table 27 contains the results that occurred while testing the trained model for all features. Moreover, that is for the negative class, which contains not depression indicated posts. The highest F-measure, 92.42, was obtained for the Deep learning classification method, and it comes under the Neural Nets classifier.

Classifier	Sub Classifier	Train Dataset			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	41.50	90.20	84.58	87.30
Bayesian	Naïve Bayes	47.58	69.99	96.97	81.30
	Naïve Bayes (Kernel)	47.66	81.32	**97.14**	88.53
Trees	Decision Tree	46.94	85.14	95.68	90.10
	Random Forest	46.80	87.49	95.38	91.26
	Gradient Boosted Tree	45.94	87.26	93.63	90.33
	Decision Stump	36.87	**90.89**	75.15	82.27
	Random Tree	36.87	**90.89**	75.15	82.27
Rules	Rule Induction	43.55	87.79	88.76	88.27
	Single Rule Induction (Single Attribute)	39.40	87.96	80.31	83.96
Neural Nets	Deep Learning	46.40	90.37	94.57	**92.42**
Logistic Regression	Logistic Regression	43.33	87.36	88.30	87.83

Table 27: All features – Train Dataset – Negative Class

Table 28 contains the results that occurred while testing the second dataset using the developed model for all features. Moreover, that is for the positive class that contains depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 87.19, was obtained for the Random Forest classification method under the Trees classifier.

Classifier	Sub Classifier	Pre - COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	5.07	24.68	31.40	27.64
Bayesian	Naïve Bayes	0.93	22.58	5.79	9.21
	Naïve Bayes (Kernel)	3.20	26.67	19.83	22.75
Trees	Decision Tree	1.60	21.82	9.92	13.64
	Random Forest	75.33	**82.48**	**92.47**	**87.19**
	Gradient Boosted Tree	2.53	28.36	15.70	20.21
	Decision Stump	3.73	16.87	23.14	19.51
	Random Tree	2.40	25.00	14.88	18.65
Rules	Rule Induction	2.67	21.98	16.53	18.87
	Single Rule Induction (Single Attribute)	3.73	18.18	23.14	20.36
Neural Nets	Deep Learning	2.00	31.25	12.40	17.75
Logistic Regression	Logistic Regression	3.33	31.65	20.66	25.00

Table 28: All features – Pre - COVID Dataset – Positive Class

Table 29 contains the results that occurred while testing the second dataset using the developed model for all features. Moreover, that is for the negative class, which contains not depression indicated posts and using the pre – COVID dataset (dataset acquired from the IT field workers before the COVID period). The highest F-measure, 89.76, was obtained for the Naïve Bayes classifier, which comes under the Bayesian classifier.

Classifier	Sub Classifier	Pre - COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	68.40	**86.07**	81.56	83.76
Bayesian	Naïve Bayes	80.67	84.14	**96.18**	**89.76**
	Naïve Bayes (Kernel)	75.07	85.30	89.51	87.35
Trees	Decision Tree	78.13	84.32	93.16	88.52
	Random Forest	2.53	29.23	13.67	18.63
	Gradient Boosted Tree	77.47	85.07	92.37	88.57
	Decision Stump	65.47	84.08	78.06	80.96
	Random Tree	76.67	84.81	91.41	87.99
Rules	Rule Induction	74.40	84.67	88.71	86.65
	Single Rule Induction (Single Attribute)	67.07	84.40	79.97	82.12
Neural Nets	Deep Learning	79.47	84.90	94.75	89.56
Logistic Regression	Logistic Regression	76.67	85.69	91.41	88.46

Table 29: All features – Pre – COVID Dataset - Negative Class

Table 30 contains the results occurred while testing the second dataset using developed model for all features. And that is for the positive class which contains depression indicated posts and using while COVID period dataset (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure 35.40 was obtained for the K-NN classifier method, and it comes under the Lazy classifier.

Classifier	Sub Classifier	IN-COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	7.60	**31.15**	**41.01**	**35.40**
Bayesian	Naïve Bayes	2.00	27.27	10.79	15.46
	Naïve Bayes (Kernel)	3.07	30.26	16.55	21.40
Trees	Decision Tree	2.00	24.19	10.79	14.93
	Random Forest	2.53	29.23	13.67	18.63
	Gradient Boosted Tree	2.40	26.87	12.95	17.48
	Decision Stump	3.87	16.96	20.86	18.71
	Random Tree	3.33	25.77	17.99	21.19
Rules	Rule Induction	2.80	21.43	15.11	17.72
	Single Rule Induction (Single Attribute)	3.73	18.42	20.14	19.24
Neural Nets	Deep Learning	3.20	32.43	17.27	22.54
Logistic Regression	Logistic Regression	4.93	33.04	26.62	29.48

Table 30: All features – In - COVID Dataset – Positive Class

Table 31 contains the results that occurred while testing the second dataset using the developed model for all features. Moreover, the negative class that contains not depression indicated posts and using while COVID period (dataset acquired from the IT field workers during the before the COVID period). The highest F-measure, 87.44, was obtained for the Naïve Bayes classifier methods, and it comes under the Bayesian classifier.

Classifier	Sub Classifier	IN-COVID			
		Accuracy	Precision	Recall	F-measure
Lazy	K-NN	64.67	**85.54**	79.38	82.34
Bayesian	Naïve Bayes	76.13	82.16	**93.45**	**87.44**
	Naïve Bayes (Kernel)	74.40	82.79	91.33	86.85
Trees	Decision Tree	75.20	81.98	92.31	86.84
	Random Forest	75.33	82.48	92.47	87.19
	Gradient Boosted Tree	74.93	82.28	91.98	86.86
	Decision Stump	62.53	81.00	76.76	78.82
	Random Tree	71.87	82.54	88.22	85.28
Rules	Rule Induction	71.20	81.90	87.40	84.56
	Single Rule Induction (Single Attribute)	64.93	81.44	79.71	80.56
Neural Nets	Deep Learning	74.80	82.99	91.82	87.18
Logistic Regression	Logistic Regression	71.47	84.01	87.73	85.83

Table 31: All features – In - COVID Dataset – Negative Class

Feature	Classifier	Train			Pre COVID			In COVID		
		Precision	Recall	F-measure	Precision	Recall	F-measure	Precision	Recall	F-measure
Emotion	Random Forest	**90.14**	**79.94**	**84.73**	31.45	57.35	40.63	51.39	58.27	54.61
Sentiment	Rule Induction	73.99	78.61	**76.23**	37.62	65.29	**47.73**	40.76	69.78	51.46
Language	Naïve Bayes	95.06	60.11	**73.65**	**100**	**100**	**100**	**100**	98.28	**99.13**
	Naïve Bayes	95.06	60.11	**73.65**	**100**	**100**	**100**	**100**	98.28	**99.13**
Linguistic Style	Gradient Boosted	91.00	85.51	**88.17**	85.91	89.95	87.88	34.72	16.03	21.93
All	Deep Learning	94.52	90.29	**92.36**	31.25	12.40	17.75	32.43	17.27	22.54

Table 32: Highest F – measure of trained data set in each feature for Positive class

Table 32 contains the highest F – measure values obtained for each feature and relevant classifier for the train and test dataset. This dataset is for the positive class which is depression indicative. Also, this table represents the pre – COVID (before the COVID period) and in – COVID (during the COVD period) data values relevant for that classification method.

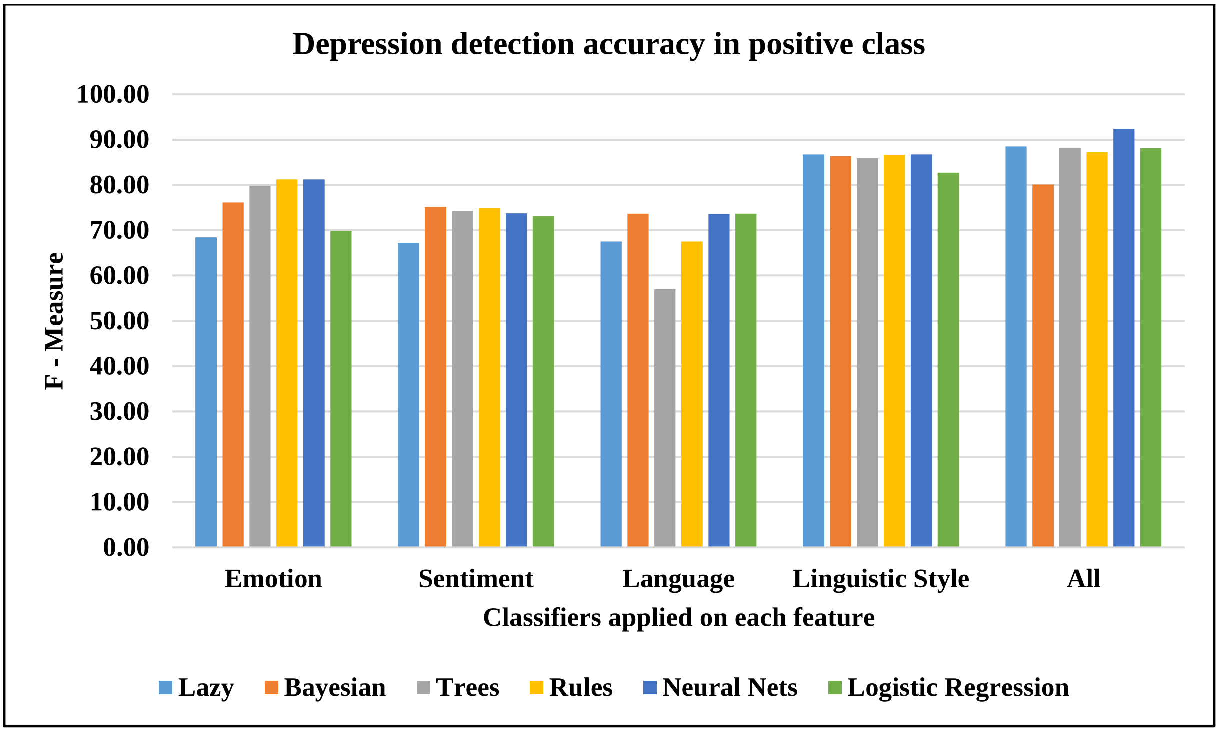


Figure 6: Depression detection accuracy of positive class for each feature

Figure 6 is representing the depression detection accuracy in each classification method for each feature. Each group represents different classification methods. This is representing depression indicative or positive class. Table 33 contains the highest F – measure values obtained for each feature and relevant classifier for the train and test dataset. This dataset is for the negative class which is not depression indicative. Also, this table represents the pre – COVID (before the COVID period) and in – COVID (during the COVD period) data values relevant for that classification method. Figure 7 is representing the depression detection accuracy in each classification method for each feature. Each group represents different classification methods. This is for the not depression indicative or negative class.

Table 34 represents the number of depressions indicated and not depression indicated posts count in each user during the COVID period and before the COVID period. That values were taken by the labeled data in the linguistic feature. That feature was selected due to it gave the highest f- measure among the other 3 features in the trained model.

Table 35 represents the percentages of posts which are predicted as depression indicated or not depression indicated by using the Table 34. The state was taken by considering and comparing the percentage of depression indicated posts and not depression indicated posts percentages. If the depression indicated posts percentage is increased in in – COVID period than the pre – COVID period those users were marked as Positive. If the pre – COVID and In - COVID percentages are equal those are marked as Neutral, and rest was marked as Negative.

Table 36 is represented the finalized result of the screening test of the survey respondents and the state acquired by the data of the linguistic style (based on table 35). The mismatching results were displayed in bold letter.

Feature	Classifier	Train			Pre COVID			In COVID		
		Precision	Recall	F-measure	Precision	Recall	F-measure	Precision	Recall	F-measure
Emotion	Gradient Boosted Tree	80.38	93.19	**86.31**	93.05	86.36	89.58	87.75	86.20	86.96
Sentiment	Gradient Boosted Tree	70.07	**82.47**	**75.77**	89.85	75.99	82.34	89.07	**78.72**	83.58
Language	Logistic Regression	70.02	96.79	**81.26**	99.86	**100.00**	99.93	**99.86**	**100.00**	**99.93**
Linguistic Style	Gradient Boosted Tree	85.84	91.22	**88.45**	85.91	89.95	87.88	80.68	92.09	86.01
All	Deep Learning	90.37	94.57	**92.42**	84.90	94.75	89.56	82.99	91.82	87.18

Table 33: Highest F – measure of trained data set in each feature for negative class

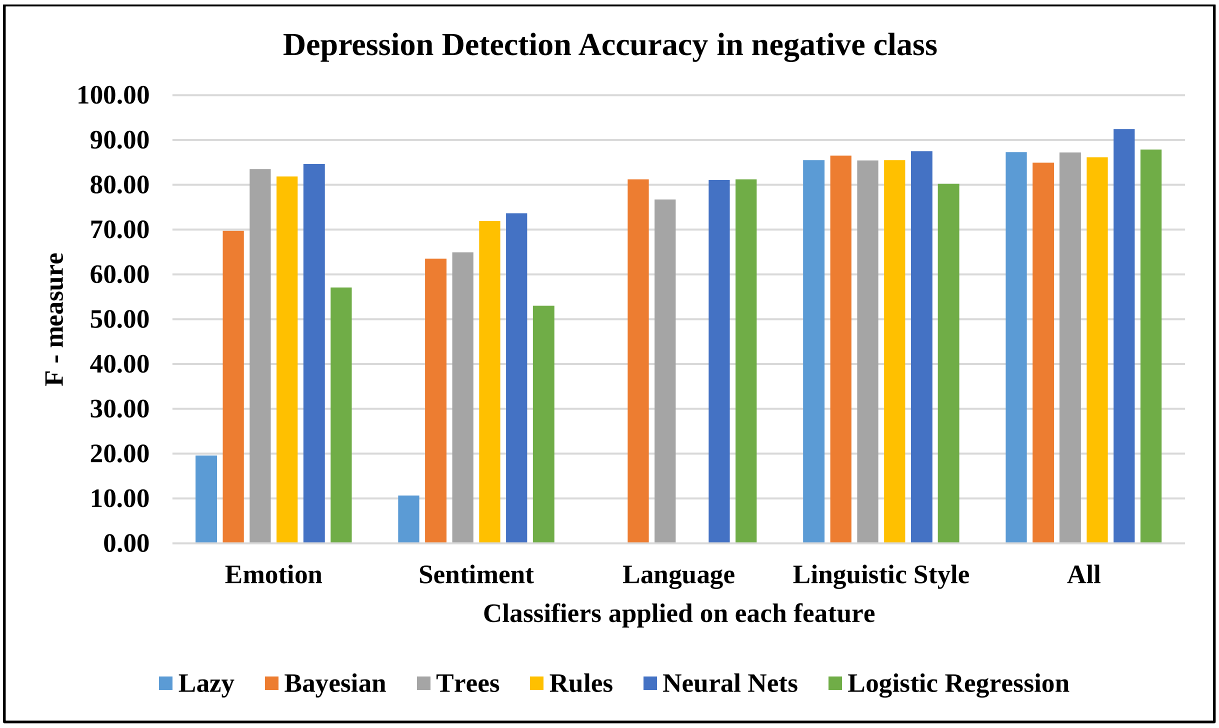


Figure 7: Depression detection accuracy of negative class for each feature

Respondent ID	Pre COVID				In COVID			
	Depression Indicated Post Count	Not a Depression indicated post Count	Depression indicated posts %	Not a Depression Indicated Posts %	Depression Indicated Post Count	Not a Depression indicated post Count	Depression indicated posts %	Not a Depression Indicated Posts %
1	1	24	4	96	3	22	12	88
2	1	24	4	96	1	24	4	96
3	3	22	12	88	1	24	4	96
4	1	24	4	96	5	20	20	80
5	1	24	4	96	2	23	8	92
6	6	19	24	76	8	17	32	68
7	2	23	8	92	2	23	8	92
8	2	23	8	92	0	25	0	100
9	1	24	4	96	5	20	20	80
10	2	23	8	92	1	24	4	96
11	2	23	8	92	7	18	28	72
12	1	24	4	96	3	22	12	88
13	3	22	12	88	7	18	28	72
14	6	19	24	76	6	19	24	76
15	4	21	16	84	8	17	32	68
16	3	22	12	88	5	20	20	80
17	8	17	32	68	8	17	32	68
18	1	24	4	96	1	24	4	96
19	2	23	8	92	9	16	36	64
20	1	24	4	96	5	20	20	80
21	0	25	0	100	1	24	4	96
22	5	20	20	80	3	22	12	88
23	8	17	32	68	8	17	32	68
24	8	17	32	68	6	19	24	76
25	4	21	16	84	9	16	36	64
26	10	15	40	60	10	15	40	60
27	7	18	28	72	7	18	28	72
28	7	18	28	72	5	20	20	80
29	5	20	20	80	6	19	24	76
30	8	17	32	68	14	11	56	44

Table 34: Each users depression indicated and not indicated post count in Pre- COVID and in COVID period

Respondent ID	Pre - COVID		In - COVID		State of mind is more towards to
	Depression indicated posts %	Not a Depression Indicated Posts %	Depression indicated posts %	Not a Depression Indicated Posts %	
1	4	96	12	88	Positive
2	4	96	4	96	Neutral
3	12	88	4	96	Negative
4	4	96	20	80	Positive
5	4	96	8	92	Positive
6	24	76	32	68	Positive
7	8	92	8	92	Neutral
8	8	92	0	100	Negative
9	4	96	20	80	Positive
10	8	92	4	96	Negative
11	8	92	28	72	Positive
12	4	96	12	88	Positive
13	12	88	28	72	Positive
14	24	76	24	76	Neutral
15	16	84	32	68	Positive
16	12	88	20	80	Positive
17	32	68	32	68	Neutral
18	4	96	4	96	Neutral
19	8	92	36	64	Positive
20	4	96	20	80	Positive
21	0	100	4	96	Positive
22	20	80	12	88	Negative
23	32	68	32	68	Neutral
24	32	68	24	76	Negative
25	16	84	36	64	Positive
26	40	60	40	60	Neutral
27	28	72	28	72	Neutral
28	28	72	20	80	Negative
29	20	80	24	76	Positive
30	32	68	56	44	Positive

Table 35: Based on the depression indicated and not indicated posts percentages

Respondent ID	Screening test result	State according to the Facebook posts
1	Depressed	Depressed
2	Not Depressed	Not Depressed
3	Not Depressed	Not Depressed
4	Depressed	Depressed
5	Depressed	Depressed
6	Depressed	Depressed
7	Not Depressed	Not Depressed
8	Not Depressed	Not Depressed
9	Depressed	Depressed
10	Not Depressed	Not Depressed
11	Not Depressed	Depressed
12	Not Depressed	Depressed
13	Depressed	Depressed
14	Not Depressed	Not Depressed
15	Not Depressed	Depressed
16	Not Depressed	Depressed
17	Depressed	Not Depressed
18	Not Depressed	Not Depressed
19	Depressed	Depressed
20	Depressed	Depressed
21	Not Depressed	Depressed
22	Not Depressed	Not Depressed
23	Not Depressed	Not Depressed
24	Not Depressed	Not Depressed
25	Not Depressed	Depressed
26	Not Depressed	Not Depressed
27	Not Depressed	Not Depressed
28	Depressed	Not Depressed
29	Not Depressed	Depressed
30	Not Depressed	Depressed

Table 36: Summary of screening test and state of mental health according to the Facebook posts

Screening result and labelled data matching count	20 Respondents
Screening result and labelled data mismatching count	10 Respondents

Table 37: Number of respondents which has matching results for screening and Facebook posts labelling

	Number of Respondents	Percentage
Depressed Respondents	8	26.67%
Non-depressed Respondents	12	40.00%
Removed respondents	10	33.33%

Table 38: Percentages of depressed and non-depressed respondents according to Screening Test and prediction of Facebook posts

Table 37 contains summarization of the number of respondents which has same result for the screening test and prediction of Facebook posts. Table 38 contains breakdown of the number of respondents which shows depression and number of respondents which does not show depression according to the screening test and predict the state (depressed / not depressed) using Facebook posts. Table 38 is a further breakdown of Table 37. Table 39 displays the number of respondents and percentages in depressed and non-depressed groups in COVID period and before the COVID period. The removed users are renamed as the not applicable respondents.

	Before COVID		**In - COVID**	
	Number of respondents	**Percentage**	**Number of respondents**	**Percentage**
Depressed Respondents	0	0	8	26.67%
Non-depressed Respondents	30	100%	12	40.00%
Not Applicable Respondents	0	0	10	33.33%

Table 39: Percentage of depressed and non-depressed respondents before and in-COVID period

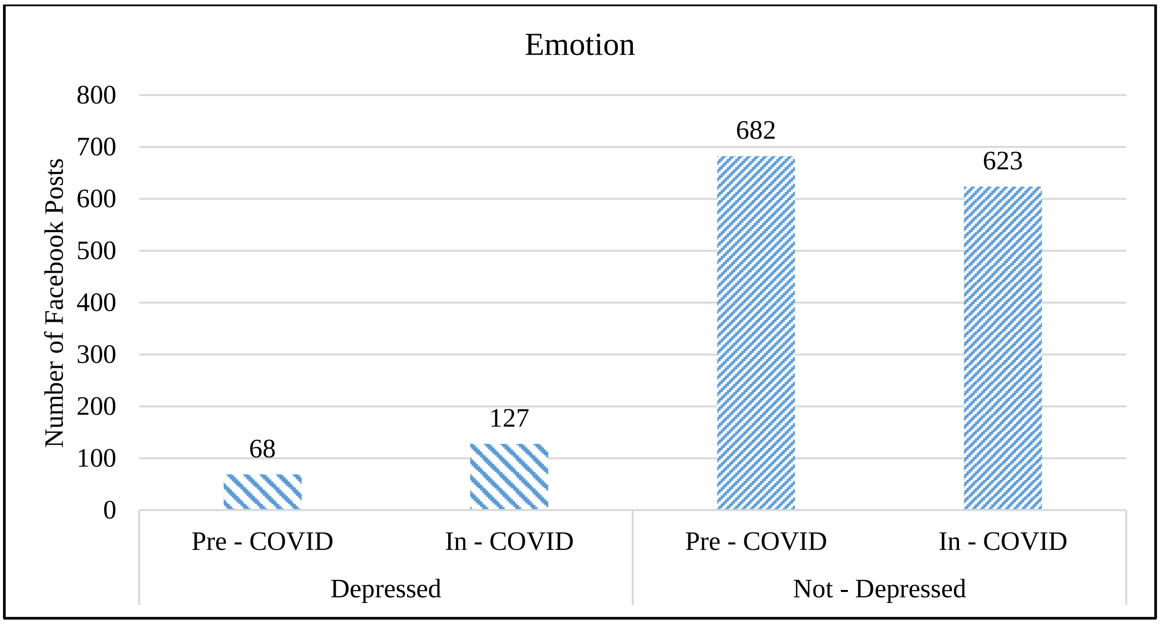


Figure 8**:** Depressed and Non-Depressed post counts of Emotion feature in pre – COVID and in – COVID period

Figure 8 is displayed the number of posts relevant to depressed and non-depressed groups before the COVID (pre-COVID) period and in the COVID(In-COVID) period. Total of 1500 posts, and out of that, 750 are related to the pre – COVID period, and the rest is relevant to in – COVID period. Before the COVID period, 68 posts were marked as depression indicated, and in the COVID period, it has increased to 127. Six hundred eighty-two posts were marked as the non – depressed before the COVID period, and 623 posts were marked as non – depressed in the COVID period.

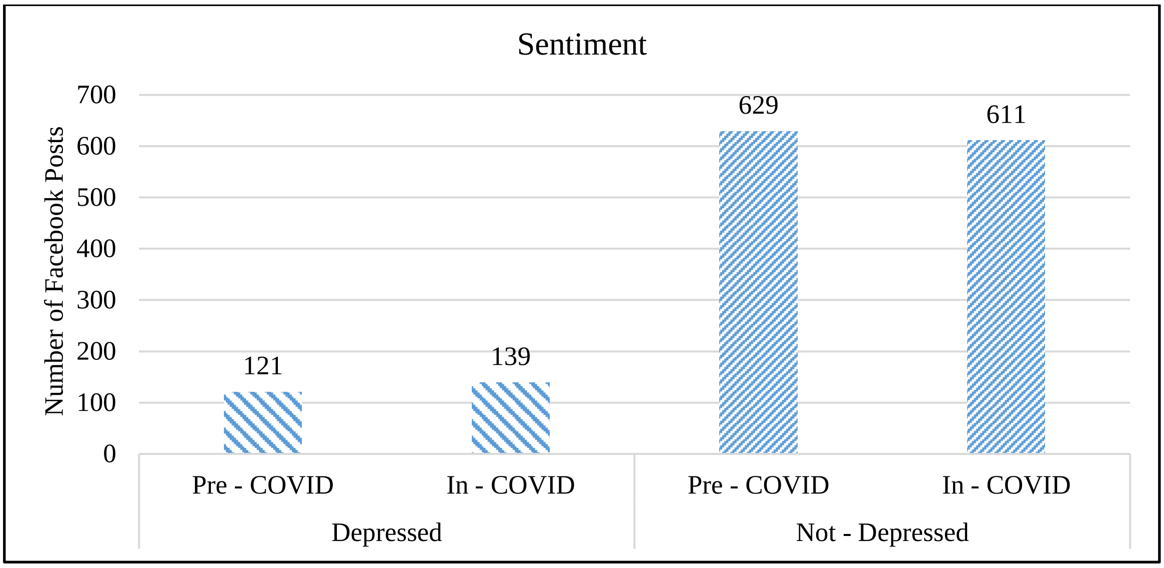


Figure 9: Depressed and Non-Depressed post counts of Sentiment feature in pre – COVID and in – COVID period

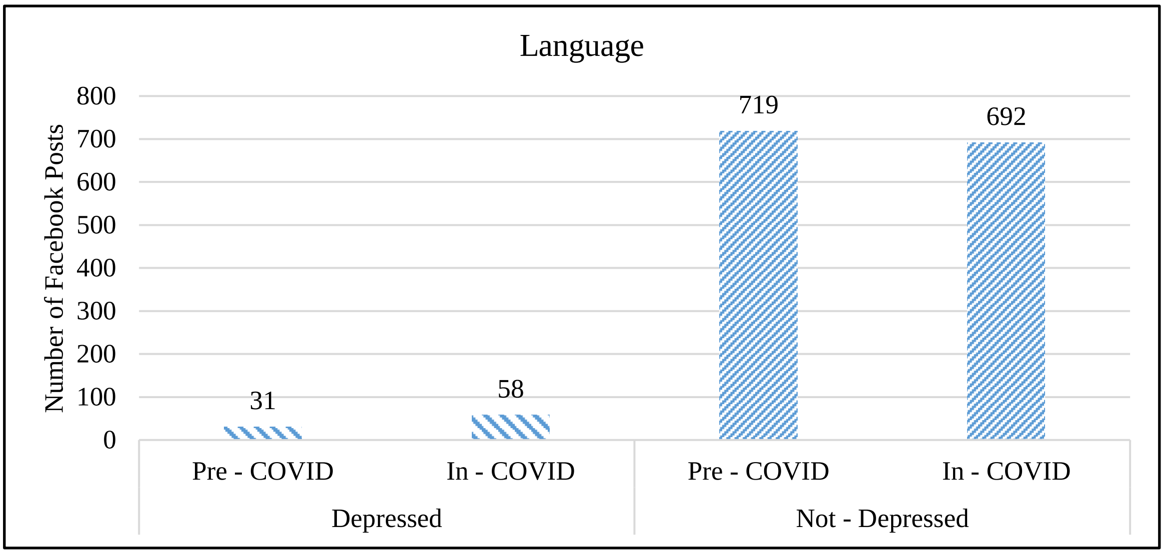


Figure 10: Depressed and Non-Depressed post counts of Language feature in pre – COVID and in – COVID period

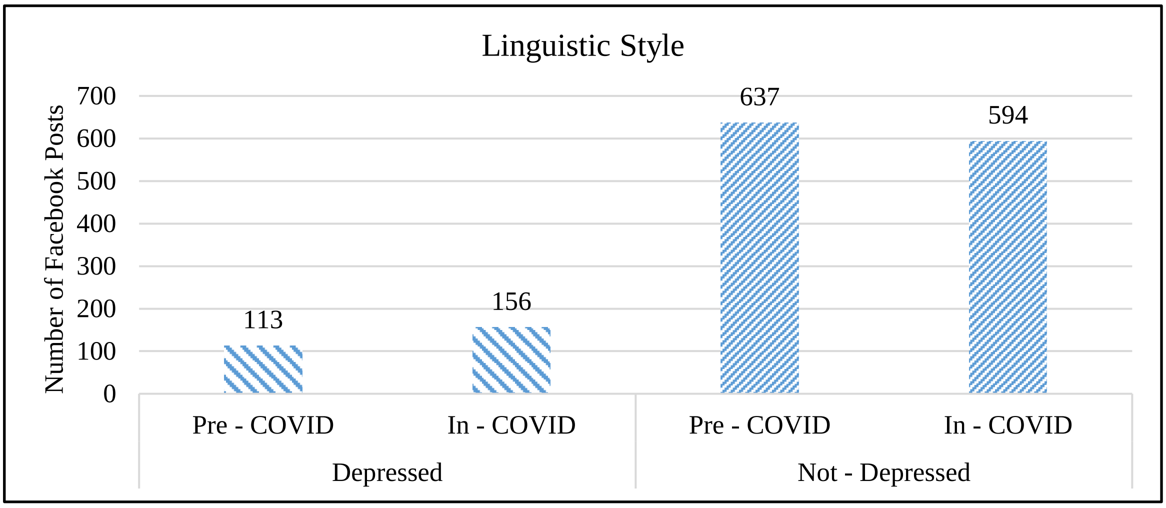


Figure 11**:** Depressed and Non-Depressed post counts of Linguistic Style feature in pre – COVID and in – COVID period

Figures 9,10, and 11 represent the number of posts marked as depressed and non-depressed during the COVID and before the COVID periods for each feature. Figure 9 shows the fluctuation of number of posts before the COVID period and after the COVID period for the sentiment feature. Figure 10 shows the fluctuation of number of posts before the COVID period and after the COVID period for the language feature. Figure 11 shows the fluctuation of number of posts before the COVID period and after the COVID period for the linguistic style feature.

4.2 Discussion

I have used f – measure value to prove that mental health can be predicted by using social media data in a pandemic situation like COVID-19. I have applied lazy, Bayesian, trees, rules, neural nets and logistic regression classification methods for depression detection. And I have shown that all the features emotion, sentiment, language, linguistic style, and combination of all features (emotional, sentiment, language, and linguistic style) were able to successfully predict the depression emotional result. F-measure is used to measure the accuracy of the classification and if it gives higher value that means it has the high accuracy compared to the other classification methods in the developed model. When considering the results obtained according to Table 32, combination of all features (emotion, sentiment, language, and language style) gives the highest f-measure as the 92.36. And if it considers only 4 features emotion, sentiment, language, and linguistic style, the linguistic style gives highest f-measure among the rest of the features. Similarly, for the gradient boosting tree sub-classifier, the next highest value of the f-measure of the language feature is 88.17. Among all the features, the language feature gives the lowest f-measure of 73.65. For emotional feature, the random forest sub-classifier gives the highest f-measure of 84.73. For the sentiment feature, rule induction classifier gives the highest result as 76.23. As a summary combination of all features gives the maximum result, it reveals that when the number of features is increased, it provides a more accurate result. These results reveal that it provides more accurate results when the number of features is increased. In addition, when considering each feature, the language style feature gives the highest results.

When considering the results before the COVID period and the COVID period, the results of all features except language give F-measures lower than the training data set. For language features, it provides 100 for the pre-COVID data set and 99.13 for the COVID time data set. That result shows that the data set is overfitting. Overfitting means that the pre-COVID data set and in-COVID data set are exactly fitted with their training data set. This might be due to the way I have used the language feature and due to the noise in the dataset. In the data set before the COVID period, the highest F-measure value achieved by the language style feature, for the gradient boosting tree, its value is 87.88. In addition, for the emotion, sentiment, and combination of all features show lower F-measure compared to the train dataset F-measure. There is a significant decrement in the F-measure values. That might be due to before the COVID period and in the COVID period dataset not fit to the training dataset. When considering the data set before COVID, the decision tree classifier gave the highest f-measure of language style features of 89.29. In the data set during the COVID period, the logistic regression classifier gave the highest F-measure of language style features of 56.40. When comparing the results, the language style gives the highest F-measure. Since the language feature shows overfitting, this feature is not suitable for predicting depression. For the model, I have obtained 92.36 F-measure and it's a one of the highest F-measure among the already available models developed by other researchers. The above results occurred in the positive class or depression indicator group. The same results were obtained for negative or non-depressive indicator categories. For the combination of all features, the highest F-measure was also obtained. This reveals that the model gives more precise results in predicting depression.

According to the screening test, the mental health of some interviewees has declined. Since they have marked themselves as not experiencing depression in their lifetime, the most relevant reason for reducing their mental health is this pandemic situation. In the questionnaire, I used two different screening test questions to obtain a more specific data set. By using two screening test questions, I deleted some noisy answers and filtered out more precise answers. When considering screening test results, it has marked some respondents as depressed, while others have not.

When considering the labeled data of linguistic style feature, Table 34 shows the number of posts indicating depression and the number of posts indicating non-depression for each respondent. According to the result of that table, I have marked each respondents' mental health state whether it's more towards to negative, positive, or neutral. This is based on an increase, decrease, or no change in the number of posts indicating depression. If count is increased those respondents were marked as more towards to positive, if count Is decreased those respondents were marked as negative and if counts were remaining unchanged those respondents were marked as neutral. When comparing the screening results with the labeled data results, a considerable amount of result matches. From the results of 30 respondents, the results of 20 respondents were matched. When taking it as a percentage value it's 66.67%. The results of the remaining respondents did not match. Of the 20 matched respondents, 8 respondents expressed depression, and the rest did not express depression. Since they all marked themselves in the survey as having no depression in their lifetime, the first research question proved to be a decline in mental health during pandemic situations like the COVID period. Also, this reveals that the people who are in good mental health get depressed in pandemic situations.

According to the collected dataset of pre – COVID and in – COVID period, it shows a significant difference in the number of posts that reflect depression. For all four features, it reflects an increase in the number of posts reflected by depression.

In the emotion feature, there is a significant difference in the number of posts that reflect the depression. As per figure 8, in the COVID period, 59 posts were reflecting depression emotion than before the COVID period. In the emotion feature, the main attributes which are considered are positive emotion, negative emotion, anger, anxiety, and sad. From those attributes negative emotion, sad, anxiety, and anger are some of the symptoms of depression. Therefore, the increment of this features' number of posts revealed that there is an increment in depression posts. This reflects that there is a reduction in mental health.

The sentiment feature analyzes the sentiment and subjectivity of each post. The sentiment is based on the polarity of each post. If the polarity is positive value sentiment is positive and if the polarity is negative value sentiment is negative. Subjectivity measures how subjective the text and polarity measures how negative or positive the text is. If some person has more negative thoughts all the time, it means that person has a risk of being depressed due to less emotional well-being. In the sentiment feature, there is a slight increase in the number of posts in the pre – COVID and in – COVID periods. This also reflects the decline in mental health during the COVID period. Figure 9 shows the slight increment in the depression.

In the language feature, the number of posts compared to before and during the COVID period has increased. Language is one of the main characteristics used to predict depression. People who have symptoms of depression use a high number of words conveying negative emotions, specifically negative adjectives, and adverbs such as sad, lonely. Also, there is a tendency to use their depression symptoms in their posts to indicate that they were felt with the depression. Here, I have used words such as abuse, anger, anxiety, anxious, bad mood, blues, broke, concentrate, cry, dead depression, die, dilemma, disappoint, fatigue, feel bad, hate, helpless, hopeless, insomnia, irritate, lack of sleep, mental disorder, pain, sad, stress, suicide, terrible, tired, unfair. Figure 10, which shows that quite a few posts contain a depression indicated language. This also stands as proof that there is a mental health reduction in the COVID period compared to the period before the COVID.

In terms of language style characteristics, there are also significant differences in the number of posts before and during the COVID period. The most interesting thing is depressed people are tending to use pronouns. I, me, myself etc. Also, there is less tendency to use second and third person pronouns such as she, them, they. This reveals that those depressed patients are more focused on themselves and have less connection with others. Researchers have found that using a linguistic style feature to predict depression is more accurate than using negative emotion words.

Considering the above, it shows that mental health has declined during the pandemic compared with the normal period. Also, there was no reduction in the number of posts which reflect depression during the COVID period. In all the features there is an increment in the number of posts that shows depression. This reflects that there are some portions of posts that remain unchanged during the COVID and before the COVID periods. Some portions of posts have been added newly to the depression category in the COVID period, so the number of posts reflecting depression has increased. Hence, that stands as a proof to there is less impact on the people who are felt with depression already and it reveals that there is a tendency to reduce the mental health of the people who are in good mental health during the pandemic.

4.3 Conclusion

In this research study, I have defined a model that can be used to predict depression in a pandemic situation like COVID 19. In addition, this research proves that in pandemic situations such as COVID-19, we can see a decline in mental health. Moreover, there is a reduction in the mental health of ordinary people during the COVID-like pandemic period, and the pandemic has less impact on depressed people. I have used different classifiers to train and test the datasets. Initially, I have used many pre-processing methods, and it includes cleaning and pre-processing data, labeling, and feature extractions. Among the features of emotion, sentiment, language, and linguistic style, the linguistic style feature gives the highest F-measure. When considering all 5 features the combination of all features (emotion, sentiment, language, and language style) provides the highest f-measure for the training data set. But the overall best result is given by the linguistic style feature and the lowest result is given by the language feature. In addition, the language feature is not suitable for use as a feature because it shows overfitting in the pre-COVID and COVID data sets.

In future work, can plan to use a dataset to train the model from the period before the pandemic and in the pandemic situation. It might give better results than this due to the trained data set which I have used is not totally the same as the period in which I have acquired the dataset in the pre-COVID and in-COVID period. Also, if we can acquire the dataset in which the original language is English, it might give better results. That’s due to the when translating the dataset to the English language from the Russian language there might be grammar mistakes.

APPENDICES

1. Survey Form - Predicting depression using social media posts - Test Data - Google Forms
2. RapidMiner Process flows
3. Source Code - Preprocessing. depression language and sentimental analysis
4. Survey Results
5. Datasets related to the research will be accessible in below mentioned URL - https://drive.google.com/drive/folders/18obI6KGY48DkeMsFRifOWb8lMPU5LSPj?usp=sharing

REFERENCES

📙 Emojipedia — 😃 Home of Emoji Meanings 💁👌🎍😍 [WWW Document], n.d. URL https://emojipedia.org/ (accessed 3.19.21).

A Beginner's Guide to Neural Networks and Deep Learning [WWW Document], n.d. . Pathmind. URL http://wiki.pathmind.com/neural-network (accessed 3.21.21).

Aldarwish, M.M., Ahmad, H.F., 2017. Predicting Depression Levels Using Social Media Posts, in: 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS). Presented at the 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS), IEEE, Bangkok, Thailand, pp. 277–280. https://doi.org/10.1109/ISADS.2017.41

Almeida, H., Briand, A., Meurs, M.-J., n.d. Detecting Early Risk of Depression from Social Media User-generated Content 12.

Al-Mosaiwi, M., n.d. People with depression use language differently – here's how to spot it [WWW Document]. The Conversation. URL http://theconversation.com/people-with-depression-use-language-differently-heres-how-to-spot-it-90877 (accessed 7.3.21).

AlSagri, H., Ykhlef, M., n.d. Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features 16.

Bruce, D.F., PhD, n.d. How Do Doctors Diagnose Depression? [WWW Document]. WebMD. URL https://www.webmd.com/depression/guide/depression-diagnosis (accessed 7.17.21a).

Bruce, D.F., PhD, n.d. Tests Used to Diagnose Depression [WWW Document]. WebMD. URL https://www.webmd.com/depression/guide/depression-tests (accessed 7.17.21b).

Charlesthelindenmethod, ~, 2019. What Is Clinical Depression? Charles Linden. URL https://charlesthelindenmethod.wordpress.com/2019/08/19/what-is-clinical-depression/ (accessed 8.31.21).

Choudhury, M.D., Gamon, M., Counts, S., Horvitz, E., n.d. Predicting Depression via Social Media 10.

Colaboratory – Google [WWW Document], n.d. URL https://research.google.com/colaboratory/faq.html (accessed 6.26.21).

Coronavirus disease 2019, 2021. . Wikipedia.

De Choudhury, M., Counts, S., Horvitz, E., 2013. Social media as a measurement tool of depression in populations, in: Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13. Presented at the the 5th Annual ACM Web Science Conference, ACM Press, Paris, France, pp. 47–56. https://doi.org/10.1145/2464464.2464480

De Choudhury, M., Counts, S., Horvitz, E.J., Hoff, A., 2014. Characterizing and predicting postpartum depression from shared facebook data, in: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '14. Presented at the the 17th ACM conference, ACM Press, Baltimore, Maryland, USA, pp. 626–638. https://doi.org/10.1145/2531602.2531675

Depression (major depressive disorder) - Diagnosis and treatment - Mayo Clinic [WWW Document], n.d. URL https://www.mayoclinic.org/diseases-conditions/depression/diagnosis-treatment/drc-20356013 (accessed 7.18.21).

Facebook, Twitter, n.d. The 6 Types of Basic Emotions and Their Effect on Human Behavior [WWW Document]. Verywell Mind. URL https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976 (accessed 6.6.21).

Fujiwara and Kawachi - 2008 - A prospective study of individual-level social cap.pdf, n.d.

He, L., Cao, C., 2018. Automated depression analysis using convolutional neural networks from speech. Journal of Biomedical Informatics 83, 103–111. https://doi.org/10.1016/j.jbi.2018.05.007

Islam, Md.R., Kabir, M.A., Ahmed, A., Kamal, A.R.M., Wang, H., Ulhaq, A., 2018. Depression detection from social network data using machine learning techniques. Health Inf Sci Syst 6, 8. https://doi.org/10.1007/s13755-018-0046-0
Keumhee Kang, Chanhee Yoon, Eun Yi Kim, 2016. Identifying depressive users in Twitter using multimodal analysis, in: 2016 International Conference on Big Data and Smart Computing (BigComp). Presented at the 2016 International Conference on Big Data and Smart Computing (BigComp), IEEE, Hong Kong, China, pp. 231–238. https://doi.org/10.1109/BIGCOMP.2016.7425918
Kumar, A., Sharma, A., Arora, A., 2019. Anxious Depression Prediction in Real-time Social Data. SSRN Journal. https://doi.org/10.2139/ssrn.3383359
Lazy learning, 2020. . Wikipedia.
LIWC [WWW Document], n.d. . LIWC. URL http://liwc.wpengine.com/ (accessed 5.9.21).
Logistic regression, 2021. . Wikipedia.
Naive Bayes Classifiers, 2017. . GeeksforGeeks. URL https://www.geeksforgeeks.org/naive-bayes-classifiers/ (accessed 3.20.21).
Narynov, S., 2020. Dataset of depressive posts in Russian language collected from social media. Data in brief 7.
Rule-Based Classifier - Machine Learning, 2020. . GeeksforGeeks. URL https://www.geeksforgeeks.org/rule-based-classifier-machine-learning/ (accessed 3.20.21).
Sadness, Grief, Depression: What’s the Difference? | Premier Health [WWW Document], n.d. URL https://www.premierhealth.com/your-health/articles/women-wisdom-wellness-/sadness-grief-depression-what-s-the-difference- (accessed 9.1.21).
Schäfer, S.K., Sopp, M.R., Schanz, C.G., Staginnus, M., Göritz, A.S., Michael, T., 2020. Impact of COVID-19 on Public Mental Health and the Buffering Effect of a Sense of Coherence. Psychother Psychosom 1–7. https://doi.org/10.1159/000510752
Sep 03, M.V.B.| N.W.| C.N.|, 2020, n.d. Depression triples in US adults amid COVID-19 stressors [WWW Document]. CIDRAP. URL https://www.cidrap.umn.edu/news-perspective/2020/09/depression-triples-us-adults-amid-covid-19-stressors (accessed 1.23.21).
Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T.-S., Zhu, W., 2017. Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Presented at the Twenty-Sixth International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence Organization, Melbourne, Australia, pp. 3838–3844. https://doi.org/10.24963/ijcai.2017/536
Support-vector machine, 2021. . Wikipedia.
Vodacom Bulk SMS Messaging - SMS Dictionary [WWW Document], n.d. URL https://www.vodacommessaging.co.za/dictionary.asp (accessed 3.19.21).
What Is Depression? [WWW Document], n.d. URL https://www.psychiatry.org/patients-families/depression/what-is-depression (accessed 1.23.21).
Zhang, Y., Lyu, H., Liu, Y., Zhang, X., Wang, Y., Luo, J., 2020. Monitoring Depression Trend on Twitter during the COVID-19 Pandemic. arXiv:2007.00228 [cs].

1. Survey Form - Predicting depression using social media posts - Test Data - Google Forms

Dear Sir/Madam,

I would like to invite you to take part in the survey. This research project aims to develop a model to predict depression using social media posts. As well as aims to find whether there is any impact on the people's mental health by pandemic like COVID-19. Your participation will require about 10 minutes of your valuable time and all the information contributed will be regarded with the highest esteem and confidentially. Your personal details are required only for results validation purposes, those are not published, or anyone wouldn't identifiable you have taken part in this research. Also requesting you the consent to access your Facebook profiles' public posts. Please note that your participation is voluntary, and you have the right to withdraw your consent at any time. Your corporation to participate in this study is very much appreciated.

Thank you very much.

Yours Faithfully,
Thathsarani Samaranayake

* Required

1. Occupation is IT related*

Yes	
No	

Below is a list of some ways you may have felt or behaved. Please indicate how often you have felt this way during the last week.

#	Question	Rarely or none of the time (less than 1 day)	Some or little of the time (1-2 days)	Occasionally or a moderate amount of time (3-4 days)	Most or all of the time (5-7 days)
2	I was bothered by things that usually don't bother me. *				
3	I did not feel like eating; my appetite was poor. *				
4	I felt that I could not shake off the blues even with help from my family or friends. *				
5	I felt I was just as good as other people. *				
6	I had trouble keeping my mind on what I was doing. *				
7	I felt depressed. *				
8	I felt that everything I did was an effort. *				
9	I felt hopeful about the future. *				
10	I thought my life had been a failure. *				
11	I felt fearful. *				
12	My sleep was restless. *				
13	I was happy. *				
14	I talked less than usual. *				
15	I felt lonely. *				
16	People were unfriendly. *				
17	I enjoyed life. *				
18	I had crying spells. *				
19	I felt sad. *				
20	I felt that people disliked me. *				
21	I could not get going. *				

Below is a list of some ways you may have felt or behaved. Please indicate how often you have felt this way during the last two weeks.

#	Question	Not at all	Several Days	More than half the days	Nearly everyday
22	Little interest or pleasure in doing things? *				
23	Feeling down, depressed, or hopeless? *				
24	Trouble falling or staying asleep, or sleeping too much? *				
25	Feeling tired or having little energy? *				
26	Poor appetite or overeating? *				
27	Feeling bad about yourself — or that you are a failure or have let yourself or your family down? *				
28	Trouble concentrating on things, such as reading the newspaper or watching television? *				
29	Moving or speaking so slowly that other people could have noticed? Or so fidgety or restless that you have been moving a lot more than usual? *				
30	Thoughts that you would be better off dead, or thoughts of hurting yourself in some way?				

Clinical depression diagnoses history

31. Have you been diagnosed with clinical depression in the past? *

Yes	
No	

32. If so when do you diagnosed with clinical depression?

Consent for access to Facebook posts

33. Will, you agreed to give the consent to use your Facebook profile public posts in this research. *
(This will only access by myself - Thathsarani Samaranayaka and only view the posts and texts of those posts through my account. Any unauthorized actions were not taking place. Only the text of the post was acquired for the research.)

Yes	
No	

34. Username of the Facebook Profile
(Only if you grant the consent to access Facebook profile public posts)

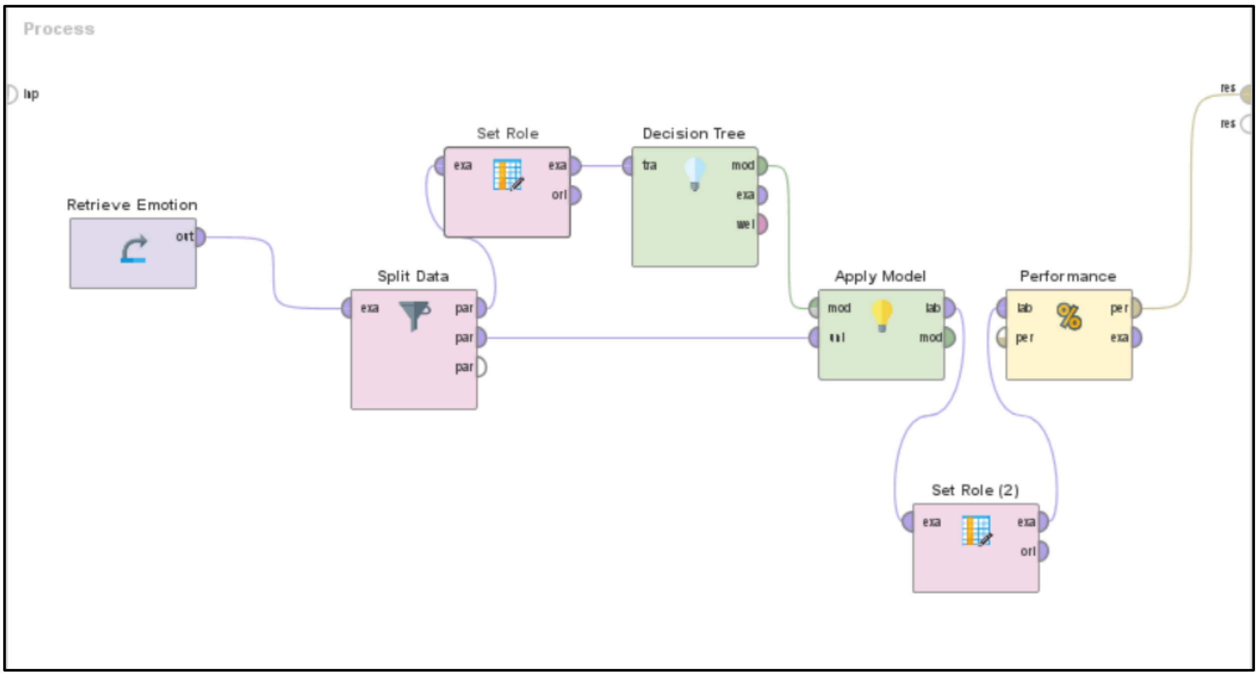
----

2. RapidMiner Process flows

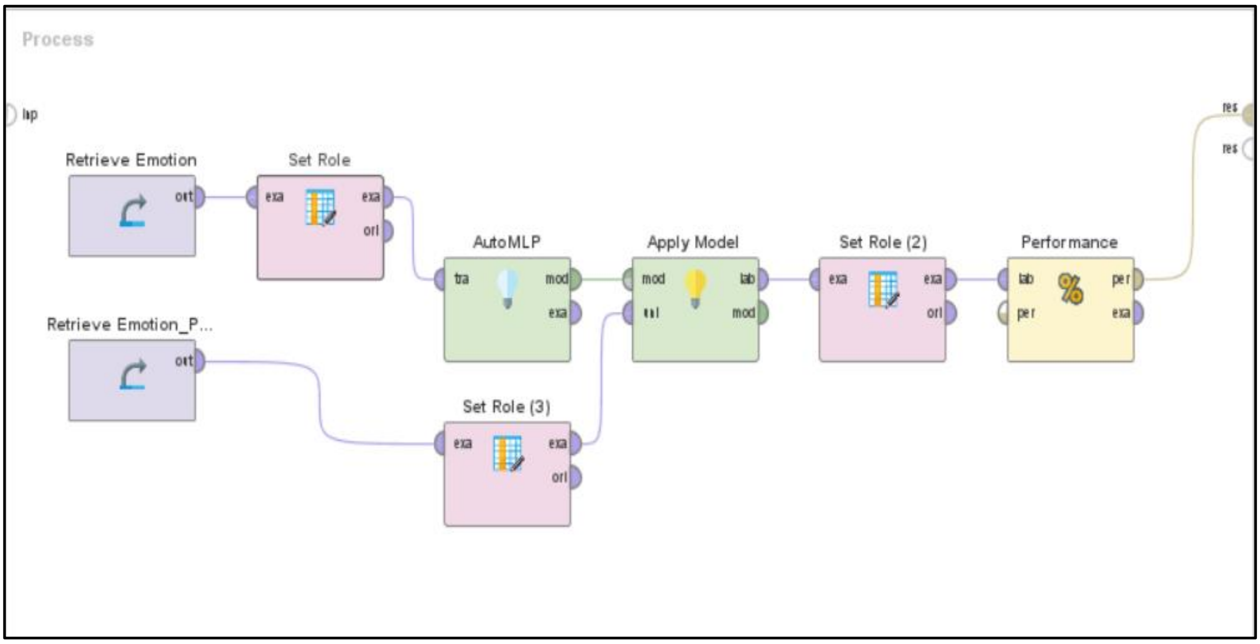
Flows implement by using the UI elements for trained dataset. Pre-covid and In-covid datasets separately. This does not contain all the Figures.

Train Dataset – Decision Tree Sub classifier for Emotion feature

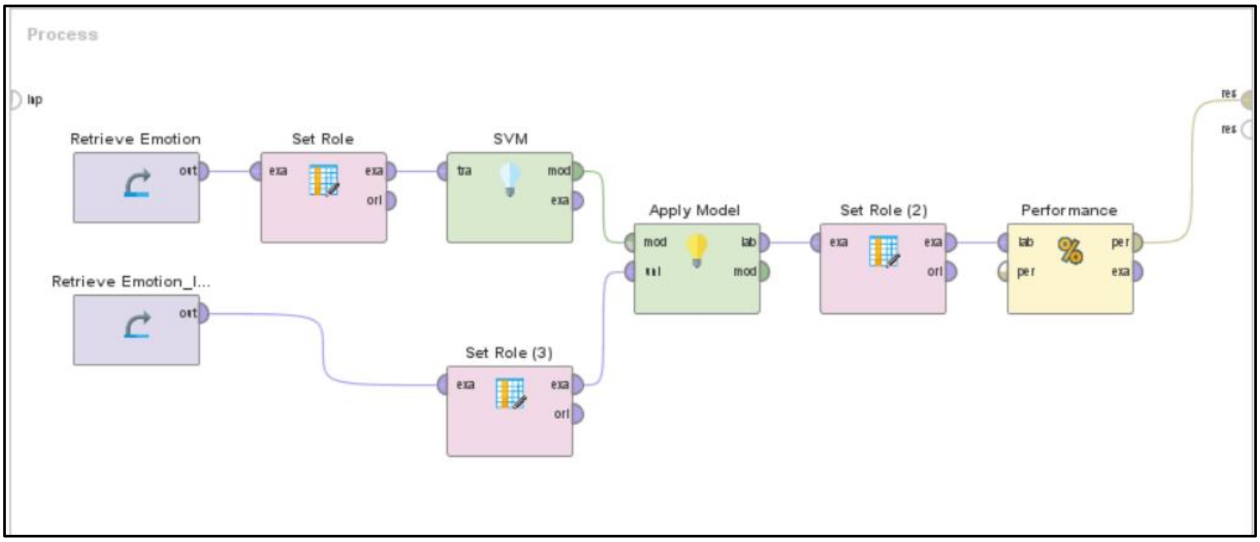
Here the dataset was split using 80:20 proportionate.



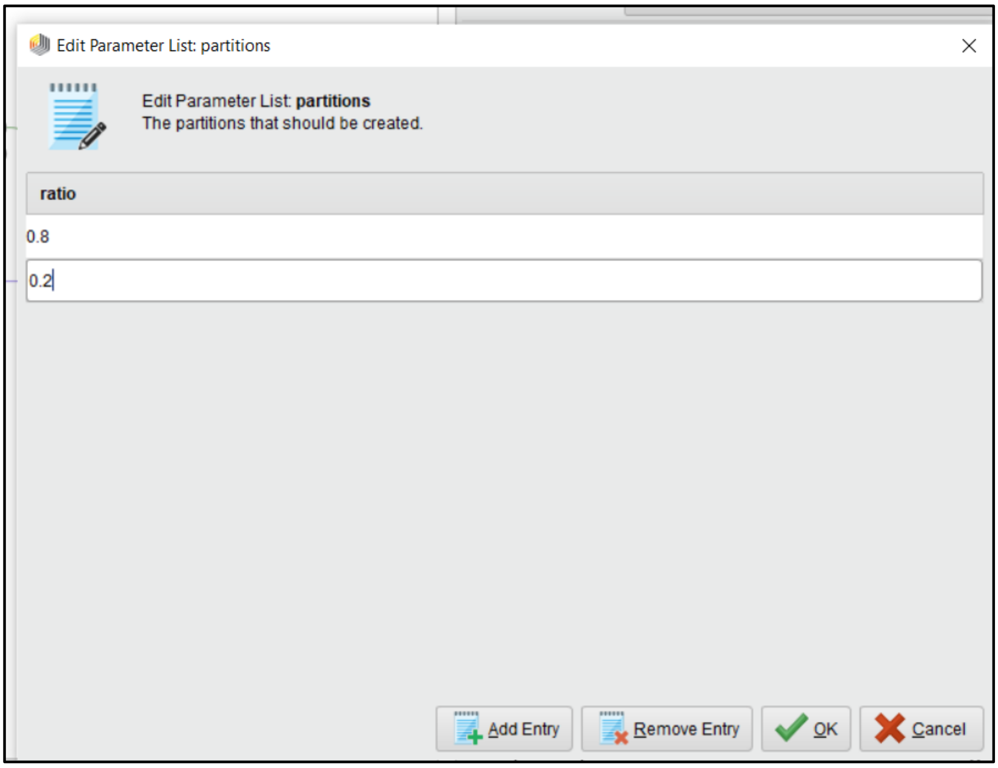
Pre – Covid dataset – AutoMLP sub classifier for Emotion feature



In – Covid dataset – SVM sub classifier for Emotion feature



Proportionate training and test partitions



3. Source Code - preprocessing. depression language and sentimental analysis

Preprocess Data

```
import pandas as pd
import numpy as np
import re
from textblob import TextBlob
from wordcloud import WordCloud
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
import nltk
import string
nltk.download('stopwords')
nltk.download('word_tokenize')
from nltk.corpus import stopwords
nltk.download('punkt')
from nltk.tokenize import word_tokenize

# Install the PyDrive wrapper & import libraries.
# This only needs to be done once per notebook.
!pip install -U -q PyDrive
from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
from google.colab import auth
from oauth2client.client import GoogleCredentials

auth.authenticate_user()
gauth = GoogleAuth()
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

file_id = '1Uj7steu_xdjp8TtAilybPCuvbStaAoh0'
downloaded = drive.CreateFile({'id': file_id})
```

```
downloaded.GetContentFile('Final_Data_Set_ROUND1\Final Dataset.xlsx')

!ls -lha Final_Data_Set_ROUND1\Final Dataset.xlsx

# Now, use pandas read_excel after installing the excel importer.
!pip install -q xlrd

import pandas as pd
df = pd.read_excel('Final_Data_Set_ROUND1\Final Dataset.xlsx')
df

def textLowercase(post):
  return post.lower()

df['Post'] = df['Post'].apply(textLowercase) #convert to lower case

def removeNumbers(post):
    post = re.sub(r'\d+', '', post)
    return post

df['Post'] = df['Post'].apply(removeNumbers) #Remove numbers
def removePunctuation(post):
    translator = str.maketrans('', '', string.punctuation)
    return post.translate(translator)

df['Post'] = df['Post'].apply(removePunctuation) #Remove punctuations

def removeWhitespace(post):
    return  " ".join(post.split())

df['Post'] = df['Post'].apply(removeWhitespace) #Remove white spaces

# remove stopwords function
def removeStopwords(post):
```

```
    stopwords = nltk.corpus.stopwords.words('english')
    word_tokens = nltk.word_tokenize(post)
    filtered_text = [word for word in word_tokens if word not in stopwords]
    return filtered_text

df['Post'] = df['Post'].apply(removeStopwords) #Remove stop words

print(df)
```

Depression Language:

```
import pandas as pd
import numpy as np
import re
from textblob import TextBlob
from wordcloud import WordCloud
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
import nltk
import string
nltk.download('stopwords')
nltk.download('word_tokenize')
from nltk.corpus import stopwords
nltk.download('punkt')
from nltk.tokenize import word_tokenize

# Install the PyDrive wrapper & import libraries.
!pip install -U -q PyDrive
from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
from google.colab import auth
from oauth2client.client import GoogleCredentials

auth.authenticate_user()
gauth = GoogleAuth()
```

```
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

file_id = '1Uj7steu_xdjp8TtAilybPCuvbStaAoh0'
downloaded = drive.CreateFile({'id': file_id})

downloaded.GetContentFile('FINAL_DATA_SET_ROUND1\Final Dataset.xlsx')

!ls -lha FINAL_DATA_SET_ROUND1\Final Dataset.xlsx



# Now, use pandas read_excel after installing the excel importer.
!pip install -q xlrd

import pandas as pd
df = pd.read_excel('FINAL_DATA_SET_ROUND1\Final Dataset.xlsx')
df

def textLowercase(post):
  return post.lower()

df['Post'] = df['Post'].apply(textLowercase) #convert to lower case

# print(df)



def removeNumbers(post):
    post = re.sub(r'\d+', '', post)
    return post

df['Post'] = df['Post'].apply(removeNumbers) #Remove numbers

df.to_excel('Round2-CleanedDataset.xlsx', sheet_name = 'New_sheet')
```

```
#Here I have checked whether the text contains the depression words or not, if it contains words that text will be save d to excel files

df[df['Post'].str.contains(" anxiety")].to_excel('AnxietyDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" dead ")].to_excel('DeadDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("suicide")].to_excel('SuicideDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("fatigue")].to_excel('FatigueDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("hopeless")].to_excel('HopelessDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("depression")].to_excel('DepressionDataSet.xlsx'
, sheet_name = 'New_sheet')

df[df['Post'].str.contains(" hate")].to_excel('HateDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" stress")].to_excel('StressDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("unfair")].to_excel('UnfairDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("anxious")].to_excel('AnxiousDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" sad")].to_excel('SadDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" abuse")].to_excel('AbuseDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" tired")].to_excel('TiredDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("dilemma")].to_excel('DilemmaDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("muscle pain")].to_excel('MusclePainDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" die ")].to_excel('DieDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("feel bad")].to_excel('FeelBadDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("lack of sleep")].to_excel('LackOfSleepDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("bad mood")].to_excel('BadMoodDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("terrible mood")].to_excel('TerribleMoodDataSet.xlsx', sheet_name = 'New_sheet')
```

```
df[df['Post'].str.contains("worst time")].to_excel('WorstTimeDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("insomnia")].to_excel('InsomniaDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("disappoint")].to_excel('DisappointDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("broke")].to_excel('BrokeDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" cry ")].to_excel('CryDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" pain ")].to_excel('PainDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" blues")].to_excel('BluesDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("bipolar disease")].to_excel('ipolarDiseaseDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("helpless")].to_excel('HelplessDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("hopeless")].to_excel('HopeLessDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("appetite")].to_excel('AppetiteDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("mental disorder")].to_excel('MentalDisorderDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains(" anger ")].to_excel('AngerDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("irritate")].to_excel('IrritateDataSet.xlsx', sheet_name = 'New_sheet')

df[df['Post'].str.contains("concentrate")].to_excel('ConcentrateDataSet.xlsx', sheet_name = 'New_sheet')
```

Sentimental Analysis

```
import pandas as pd

import numpy as np

import re

from textblob import TextBlob

from wordcloud import WordCloud

import matplotlib.pyplot as plt

plt.style.use('fivethirtyeight')
```

```
# Install the PyDrive wrapper & import libraries.
# This only needs to be done once per notebook.
!pip install -U -q PyDrive
from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
from google.colab import auth
from oauth2client.client import GoogleCredentials

auth.authenticate_user()
gauth = GoogleAuth()
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

file_id = '1Uj7steu_xdjp8TtAilybPCuvbStaAoh0'
downloaded = drive.CreateFile({'id': file_id})

downloaded.GetContentFile('Final_Data_Set_ROUND1\Final Dataset.xlsx')

!ls -lha Final_Data_Set_ROUND1\Final Dataset.xlsx



# Now, pandas read_excel after installing the excel importer.
!pip install -q xlrd

import pandas as pd
df = pd.read_excel('Final_Data_Set_ROUND1\Final Dataset.xlsx')

#clean the tweets
def cleanText(post):
    post = re.sub(r'@[A-Za-z0-9]+', '', str(post))
    post = re.sub(r'#', '', post)
    post = re.sub(r'https?:\/\/.*[\r\n]*', '', post)
    post = re.sub(r'Https?:\/\/.*[\r\n]*', '', post)
    post = re.sub(r'http?:\/\/.*[\r\n]*', '', post)
```

```
    post = re.sub(r'[\r\n]+https?:\/\/.*[\r\n]*', '', post)
    post = re.sub(r'http?: +\/\/.*[\r\n]*', '', post)
    post = re.sub(r'https?: +\/\/.*[\r\n]*', '', post)
    post = re.sub(r'[\r\n]*twitter.com*[\r\n]*', '', post)
    return post

df['Post'] = df['Post'].apply(cleanText)

#get subjectivity
def getSubjectivity(post):
  return TextBlob(post).sentiment.subjectivity

# get polarity
def getpolarity(post):
  return TextBlob(post).sentiment.polarity

# create 2 new columns
df['subjectivity'] = df['Post'].apply(getSubjectivity)
df['polarity'] = df['Post'].apply(getpolarity)

def getAnalysis(score):
  if score < 0:
    return 'Negative'
  elif score == 0:
    return 'Neutral'
  else:
    return 'Positive'

df['Analysis'] = df['polarity'].apply(getAnalysis)

print(df)

df.to_excel('Round2-Dataset-Subjectivity_Polarity_Sentiment.xlsx',
sheet_name = 'New_sheet')
```

4. Survey results

This section contains the results of the survey. Totally 45 respondents were responded to the survey.

Occupation is IT related	Have you been diagnosed with clinical depression in the past?	If so, when do you diagnosed with clinical depression?	Will, you agreed to give the consent to use your Facebook profile public posts in this research.
Yes	No	N/A	No
Yes	No	N/A	Yes
Yes	No	N/A	No
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	No
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	Yes	N/A	No
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes

Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	No
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	Yes
Yes	No	N/A	No

Centre of Epidemiology screening test related questions – Scores of Survey respondents

This section contains the responses received for the survey questions of 2 – 21. Here the responses have a weightage, and the weightages are as below:

0 – Rarely or none of the time (less than 1 day)
1 – Some or little of the time (1-2 days)
2 – Occasionally or a moderate amount of time (3-4 days)
3 – Most or all of the time (5-7 days)

Below mentioned table displays how each question gets weightages. And to determine whether patient is depressed or not, have to get the summation and check whether the sum is 16 or more. If respondent acquire score as 16 or more that respondent will be consider as the depressed respondent. If score is less than that respondent will be consider as non-depressed.

	0 - Rarely or none of the time (less than 1 day)	1 – Some or little of the time (1-2 days)	2 – Occasionally or a moderate amount of time (3-4 days)	3 – Most or all of the time (5-7 days)
Q5, Q9, Q13, Q17	3	2	1	0
Other questions	0	1	2	3

Below table contains the weightages, according to the respondents' answers.

As example in the first row, Q1 value is 0. This represents answer for the first question is 'Rarely or none of the time'. In the same row, Q2 value is 1; It represents the answer is 'Some or little of the time'.

Questions												
Q2	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**	**Q12**	**Q13**	**Q14**
0	1	0	1	1	1	1	1	0	1	1	1	0
3	1	1	2	3	3	3	3	3	2	1	3	3
1	0	0	0	1	0	1	0	1	0	0	0	1
1	0	1	0	0	2	0	1	2	0	1	1	1
1	1	2	2	1	0	1	0	1	0	1	1	1
1	1	1	1	1	2	1	1	0	1	1	1	0
3	2	0	1	2	1	1	1	0	0	0	0	2
2	1	1	0	0	0	0	0	0	0	2	0	0
0	1	1	3	0	0	1	3	1	0	0	2	3
0	0	0	1	0	1	0	0	0	1	0	0	0
2	0	1	1	1	1	1	0	1	1	0	0	0
1	1	0	2	3	0	1	1	0	0	2	2	0
0	1	1	2	1	0	1	0	0	0	0	2	2
0	0	1	0	1	0	0	0	0	0	1	0	0
1	2	1	1	2	2	0	0	0	1	1	0	1
2	2	1	2	3	3	3	3	2	1	3	2	2
1	1	0	0	1	0	0	0	0	0	0	1	0
1	0	1	0	0	0	1	1	0	0	1	0	2
1	0	0	2	1	0	0	1	1	0	1	1	0
3	2	0	2	3	2	1	2	1	3	1	3	1
0	0	1	2	1	0	1	0	0	0	0	0	0
1	3	1	2	1	0	1	3	1	0	0	2	3
3	1	2	3	1	0	3	2	0	0	2	2	0
1	0	2	0	0	0	1	0	2	0	0	0	0
1	0	0	0	1	1	1	1	0	1	1	1	0
0	0	0	1	1	0	0	1	0	0	0	0	0
0	0	0	2	1	1	1	3	0	1	0	0	1
0	0	1	1	0	0	1	0	0	0	0	0	0
0	1	0	1	1	0	1	1	0	1	1	1	0
1	1	0	1	1	0	0	3	1	1	1	1	0
2	0	0	1	1	0	1	0	0	1	1	0	0
3	1	3	2	3	3	0	1	1	1	1	2	1
2	1	1	3	3	2	3	2	3	2	3	1	2
2	0	0	0	0	1	0	1	0	2	0	2	1
1	0	2	1	1	0	1	0	0	0	0	0	0
1	1	0	1	3	2	1	2	2	3	2	2	0
1	0	0	0	1	0	1	1	0	2	1	2	0
1	1	1	0	1	0	0	1	0	1	0	1	0
1	1	0	1	1	1	1	0	0	1	0	0	0
1	1	2	1	1	0	0	1	1	0	0	0	0
1	1	0	0	1	0	0	1	1	0	0	1	1
0	1	0	1	0	1	1	1	0	0	0	1	0
1	1	0	0	1	0	0	1	1	0	1	0	0
0	2	0	3	3	2	1	3	1	1	1	1	3
1	0	0	1	1	0	1	1	1	0	1	1	0

Questions							Total Score	Depressed/ Not Depressed
Q15	Q16	Q17	Q18	Q19	Q20	Q21		
0	1	1	1	2	0	0	14	Not Depressed
3	2	2	3	2	3	3	49	Depressed
0	1	0	1	1	0	0	8	Not Depressed
1	1	1	1	0	0	0	14	Not Depressed
0	2	1	0	0	0	0	15	Not Depressed
1	0	1	2	2	0	3	21	Depressed
2	1	1	3	1	0	2	23	Depressed
0	1	0	1	1	1	0	10	Not Depressed
3	3	3	1	2	3	2	32	Depressed
0	0	1	0	0	0	0	4	Not Depressed
1	1	1	0	1	1	1	15	Not Depressed
2	0	0	0	1	0	0	16	Depressed
1	0	0	0	1	0	0	12	Not Depressed
0	0	0	0	1	0	0	4	Not Depressed
0	0	0	1	0	1	1	15	Not Depressed
3	3	2	1	2	3	3	46	Depressed
0	1	0	0	1	0	1	7	Not Depressed
0	1	3	1	0	0	1	13	Not Depressed
0	1	2	0	1	0	1	13	Not Depressed
3	0	3	3	0	2	0	35	Depressed
0	1	0	1	1	0	0	8	Not Depressed
1	2	0	1	2	0	3	27	Depressed
0	3	1	1	1	1	1	27	Depressed
1	2	2	0	0	0	1	12	Not Depressed
0	1	1	0	1	0	0	11	Not Depressed
1	2	3	1	1	0	0	11	Not Depressed
0	1	0	0	1	0	1	13	Not Depressed
0	2	0	0	1	1	1	8	Not Depressed
1	1	2	1	1	0	1	15	Not Depressed
0	0	0	0	1	1	1	14	Not Depressed
1	1	2	1	1	0	2	15	Not Depressed
1	0	2	0	3	0	2	30	Depressed
3	3	2	2	3	1	3	45	Depressed
2	0	1	0	1	0	1	14	Not Depressed
1	2	3	1	0	0	0	13	Not Depressed
3	3	0	0	1	0	0	27	Depressed
1	1	1	1	0	0	0	13	Not Depressed
0	1	3	0	1	0	0	12	Not Depressed
1	1	3	0	0	1	0	13	Not Depressed
1	1	0	1	0	0	0	11	Not Depressed
0	1	2	0	1	0	0	11	Not Depressed
1	2	1	1	1	1	1	14	Not Depressed
2	1	3	0	1	1	1	15	Not Depressed
1	1	3	0	0	3	1	30	Depressed
0	0	0	0	1	0	0	9	Not Depressed

Patient Health Questionnaire – 9

This section contains the responses received for the survey questions of 22 – 30. Here the responses have a weightage, and the weightages are as below:

0 – Not at all
1 – Several days
2 – More than half the days
3 – Nearly every day

Below table contains the weightages, according to the respondents' answers.

As example in the first row, Q22 value is 0. This represents answer for the first question is 'Not at all'. In the same row, Q23 value is 1; It represents the answer is 'Several days'.

And the total score is the summation of all the weightages in a particular row. The depression level was categorized according to below manner.

Depression Level	**Marks**
Minimal depression	0 - 4
Mild Depression	5 - 9
Moderate Depression	10 - 14
Moderately Severe Depression	15 - 19
Severe Depression	20 - 27

Questions and weightages of PHQ – 9 screening test related questions

Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Total Score	Depression Level
0	1	0	0	1	1	1	0	0	4	Minimal Depression
2	2	2	2	3	1	3	3	0	18	Moderately Severe Depression
0	0	0	0	1	1	0	0	0	2	Minimal Depression
1	1	0	0	0	0	0	0	0	2	Minimal Depression
0	0	1	1	0	1	1	0	0	4	Minimal Depression
1	1	1	1	2	3	1	3	3	16	Moderately Severe Depression
1	2	3	3	2	1	3	1	3	19	Moderately Severe Depression
2	1	2	1	3	2	2	1	2	16	Moderately Severe Depression
1	0	1	1	2	3	3	2	2	15	Moderately Severe Depression
1	1	1	0	0	1	0	0	0	4	Minimal Depression
1	1	1	0	0	0	0	0	0	3	Minimal Depression
1	1	1	1	1	1	2	2	1	11	Moderate Depression
0	0	1	0	1	0	0	2	0	4	Minimal Depression
0	0	0	0	0	0	1	0	0	1	Minimal Depression
1	1	0	0	0	1	1	0	0	4	Minimal Depression
2	2	3	3	3	3	2	2	2	22	Severe Depression
1	0	1	0	1	0	1	0	0	4	Minimal Depression

0	0	1	1	0	1	0	0	0	3	Minimal Depression
1	0	0	0	0	2	1	0	0	4	Minimal Depression
2	2	0	2	1	2	2	2	3	16	Moderately Severe Depression
1	0	0	0	0	0	0	0	0	1	Minimal Depression
3	3	1	3	2	3	3	3	2	23	Severe Depression
3	2	1	3	1	2	3	2	3	20	Severe Depression
1	0	0	1	0	0	0	0	0	2	Minimal Depression
1	1	1	1	0	0	0	0	0	4	Minimal Depression
1	0	1	0	0	1	1	0	0	4	Minimal Depression
0	1	0	1	0	1	0	0	0	3	Minimal Depression
1	0	0	1	0	1	0	0	0	3	Minimal Depression
1	0	1	0	0	0	0	0	0	2	Minimal Depression
0	0	0	0	1	1	1	0	0	3	Minimal Depression
1	0	0	1	0	1	1	0	0	4	Minimal Depression
2	3	0	2	2	3	2	2	3	19	Moderately Severe Depression
3	2	2	2	2	2	3	2	1	19	Moderately Severe Depression
0	1	2	0	0	0	0	1	0	4	Minimal Depression
1	0	1	1	0	1	0	0	0	4	Minimal Depression
2	3	1	3	2	2	2	2	3	20	Severe Depression
1	0	0	1	0	1	1	0	0	4	Minimal Depression
0	0	0	0	0	0	1	0	0	1	Minimal Depression
0	1	0	1	1	0	1	0	0	4	Minimal Depression
1	0	0	0	0	1	1	0	0	3	Minimal Depression
1	0	0	0	1	0	1	1	0	4	Minimal Depression
0	1	0	1	0	1	0	0	0	3	Minimal Depression
1	0	1	1	0	0	1	0	0	4	Minimal Depression
3	3	0	3	2	3	3	3	3	23	Severe Depression
0	0	1	1	0	0	1	0	0	3	Minimal Depression