

Mental health analysis during COVID

by Btp Group 39

FILE	REPORT.PDF (821.88K)		
TIME SUBMITTED	01-JAN-2021 10:09PM (UTC+0530)	WORD COUNT	11991
SUBMISSION ID	1482381587	CHARACTER COUNT	61199

Mental health analysis during COVID-19: A comparison before and during the pandemic

1

A THESIS REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF ENGINEERING

IN

DIVISION OF COMPUTER ENGINEERING

Submitted by:

Harsh Yadav (2017UCO1597)

Gurshaan Singh Baja (2017UCO1627)

Sudhanshu Sah (2017UCO1637)

Harshdeep Singh Sahdev (2017UCO1640)

Under the provision of :

Ms. Preeti Kaur



DIVISION OF COMPUTER ENGINEERING

1

NETAJI SUBHAS INSTITUTE OF TECHNOLOGY

UNIVERSITY OF DELHI

DECEMBER, 2020



Department of Computer Engineering

University of Delhi

Delhi-110007, India

CERTIFICATE OF ORIGINALITY

We, HARSH YADAV (2017UCO1597), GURSHAAN SINGH BAJAJ (2017UCO1627), SUDHANSU SAH (2017UCO1637), HARSHDEEP SINGH SAHDEV (2017UCO1640) student(s) of B. Tech. Department of Computer Engineering, hereby declare that the Project-Thesis titled “Mental health analysis during COVID-19: A comparison before and during the pandemic” which is submitted by us to the Department of Computer Engineering, Netaji Subhas Institute of Technology, Delhi (University of Delhi) in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology, is original and not copied from source without proper citation. The manuscript has been subjected to plagiarism check by Turnitin software. This work has not previously formed the basis for the award of any Degree.

Place: Delhi

Date: 01/01/2021

HARSH YADAV (2017UCO1597)

GURSHAAN SINGH BAJAJ (2017UCO1627)

SUDHANSU SAH (2017UCO1637)

HARSHDEEP SINGH SAHDEV (2017UCO1640)

Declaration

This is to certify that report entitled “Mental Health Analysis during COVID-19” being submitted by Harsh Yadav (2017UCO1597), Gurshaan Singh Bajaj (2017UCO1627), Sudhanshu Sah (2017UCO1637) and Harshdeep Singh Sahdev (2017UCO1640), to the Division of Computer Engineering, NSIT, (now upgraded to Netaji Subhas University of Technology) for the award of bachelor’s degree in engineering, is the record of the bona fide work carried out by us in partial fulfillment of the requirements for the award of degree of Bachelor of Engineering in Computer Engineering.

¹ The results contained in this report have not been submitted by either in part or in full to any other university or institute for the award of any degree or diploma.

HARSH YADAV (2017UCO1597)

GURSHAAN SINGH BAJAJ (2017UCO1627)

SUDHANSHU SAH (2017UCO1637)

HARSHDEEP SINGH SAHDEV (2017UCO1640)

Acknowledgement

This B.E. project is a part of the curriculum of seventh semester and plays a pivotal role in the overall grooming by providing real time project experience. We have tried to inculcate all that we have learnt from various experienced professors. We wish to express our sincere regards to individuals who have been a great support and motivation throughout the project tenure and helped us perform better.

We would extend our sincere thanks to Dr. Preeti Kaur for giving us an opportunity to work under her guidance and encouraging us at the same time to innovate and come up with fresh ideas. We would also like to thank each other for putting a collective effort to take the project at its final stage.

We wish to place on record our gratitude towards NSIT for providing us this opportunity for research and development. Working on this project has been a great learning experience. This exposure has enriched us with knowledge and has also introduced us to the attributes of a successful computer engineer.

HARSH YADAV (2017UCO1597)

GURSHAAN SINGH BAJAJ (2017UCO1627)

SUDHANSU SAH (2017UCO1637)

HARSHDEEP SINGH SAHDEV (2017UCO1640)

Abstract

COVID-19 has proved to be yet another challenge for humanity. Due to the presence of COVID, there is a massive reduction in the consumption of services provided by different businesses leading to the loss of millions of jobs. The virus has also proved to be life-threatening as well. Consequently, this has led people to suffer from distress, trauma, and other kinds of mental conditions. In these times of social distancing, lockdowns, and home quarantine, people have taken refuge on the internet to carry on with their daily lives in contrast to before, from carrying out their jobs, educating themselves, shopping, etc. And as we know, there is no better platform for sharing than social networking sites, especially during these times.

²⁴

In this paper, we aim to find the effect of COVID-19 on the mental health of the people, its severity, its extent, and try to draw some conclusions based on the trends of people's activities like posting frequency, subscription gain, etc. For this purpose, we have chosen Reddit as our only data source. Reddit provided us with comparatively easier access to the demographic of people we were trying to target, as it contains sub-groups known as subreddits, where like-minded people talk about a particular topic. For our analysis, we divided the timeline of our data into two parts; pre-Covid and post-Covid, i.e., before 15 Feb 2019 and after. Now every classification task requires accurately annotated data so that the results are as correctly classified as possible. From a chosen set of subreddits, we extracted a list of users who were members of these subreddits related to depression, anxiety, or any other mental illness condition and had posted about their conditions because of COVID on these subreddits during the pandemic. We

then extracted their entire activity on Reddit in the pre-covid timeline, i.e., before the pandemic. Using these posts and comments, we drew statistical results to determine the impact of covid on the mental condition of people by comparing their activity in the two timelines.

For correctly classifying their activity in the pre-COVID timeline, we had to use a classifier for classifying the posts into depressed or non-depressed posts. After going through the papers based on the classification in terms of mental health, we concluded that the transformer-based models would give us the best results. For choosing our classifier, we tried five different types of transformer based models vis a vis Roberta, XLnet, BERT, Longformer, DistilBERT. Out of these, Roberta proved to be the best classifier with 97% accuracy in training.

After the classification task, we drew out some statistical measures like a graph of the percentage of their posts being depressed before covid VS number of users depicting the extent of the impact of COVID. Consequently, we found out that almost 6.4% of people were not depressed before COVID. Then we tried to gain some insights from the trends based on subscription gains, and the amount of people's activities, differences between them before and after COVID. We found conclusive evidence that user engagement was much higher before the pandemic in comparison to times during the pandemic. Also, we tried to find out about the reason or justification behind some contradicting trends like the loss in subscription gain in spite of the fact that people adhered to social networking platforms more during the COVID times.

Table of Contents

CERTIFICATE	i
CANDIDATE(S) DECLARATION	ii
ACKNOWLEDGMENTS	iii
ABSTRACT	iv
7 TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	ix
LIST OF SYMBOLS, ABBREVIATIONS & NOMENCLATURE	x
<u>CHAPTER 1</u>	
1.1 Introduction	11
1.1.1 Motivation	11
1.1.2 Reddit	13
1.2 Related Work	15
1.2.1 Shortcomings of the related works and Key Challenges	18
1.3 Problem Addressed	19
<u>CHAPTER 2</u>	
2.1 Approach	21
2.1.1 Methodology	21
2.2 Data Collection	22
2.4 Classification Model	26
2.4.1 Comparison between Transformers and RNN/LSTM	26
2.4.2 Transformer Architecture	28
2.4.3 Transfer Learning using Transformers	30

2.5 Tools and Libraries Used	31
2.5.1 Python	31
2.5.2 Matplotlib	31
2.5.3 Pushshift	32
2.5.4 PRAW	33
2.5.5 Transformers: Hugging-Face	34
2.5.6 Simple Transformers by Thilina Rajapakse	35
2.6 Experiment and Evaluation	35
2.6.1 Training Dataset	35
2.6.2 Experimental Setup	36
2.6.3 Training Results	37
CHAPTER 3	
3.1 Results	38
3.2 Conclusion	43
3.3 Key Challenges Faced	44
3.4 Future Work	46
APPENDIX	
A.1 Transformer based Transfer Learning	50
A.1.1 BERT	50
A.1.2 RoBERT-a	51
A.1.3 XL-Net	51
A.1.4 Longformer	52
A.1.5 Distil-BERT	53

List of Figures

Chapter 1

- Figure 1.1 Social media has become a refuge for depressed people
- Figure 1.2: The Reddit Logo
- Figure 1.3 Depression and anxiety are harmful for mental health

Chapter 2

- Figure 2.1: Praw API
- Figure 2.2: Plot of number of authors with given number of posts
- Figure 2.3: Plot of number of authors with greater than or equal to a given number of posts
- Figure 2.4: LSTM for text classification
- Figure 2.5: Transformer and attention mechanism
- Figure 2.6: Transformer Architecture
- Figure 2.7: The encoder and decoder operation for translation tasks
- Figure 2.8: Python Logo
- Figure 2.9: The Matplotlib library
- Figure 2.10: Praw API
- Figure 2.11: Hugging Face Transformer Library

Chapter 3

- Figure 3.1: Ratio of Reddit Users who were depressed before COVID
- Figure 3.2: Distribution of users having posts which amount to a fixed ratio before COVID-19
- Figure 3.3: Number of Users per month posting in subreddits for the first time
- Figure 3.4: Number of subscribers on r/depression subreddit per month for the year 2019 and 2020
- Figure 3.5: Number of comments in the r/depression subreddit per month for 2019 and 2020

Appendix

- Figure A.1: The BERT architecture
- Figure A.2: XLNet Permutation processing
- Figure A.3: Longformer sliding window

18

List of Tables

Chapter 1

Table 1.1: The selected subreddits and their number of subscribers

Chapter 2

Table 2.1: Pushshift API Parameters

Table 2.2: The Transformer based model and their architecture used

Table 2.3 Transformer based model with their respective accuracies

Appendix

Table A.1: Comparison between various models used

List of Symbols, Abbreviations & Nomenclature

<u>Abbreviation</u>	<u>Definition</u>
COVID	CoronaVirus Disease
NLP	Natural Language Processing
OSN	Online Social Network
SMHD	Self-reported Mental Health Diagnoses
ADHD	Attention deficit hyperactivity disorder
WU3D	Weibo User Depression Detection Dataset
WHO	World Health Organization
HAN	Hierarchical Attention Network
PTSD	Post Traumatic Stress Disorder
API	Application Programming Interface
RNN	Recurrent Neural Network
LSTM	Long Short Term Memory
seq-2-seq	Sequence to Sequence
BERT	Bidirectional Encoder Representations from Transformers
PRAW	Python Reddit API Wrapper

22
CHAPTER 1

1.1 Introduction

1.1.1 Motivation

In December 2020, a cluster of pneumonia cases were found because of the SARS-CoV-2 infection in Wuhan, China and was officially relegated to the name of COVID-19 by the WHO (World Health Organization) has accomplished the severity level of a typical epidemic, it has affected multiple nations across the world and every continent of the planet. As of today (29 December, 2020), in excess of 81 million positive cases and 1.7 million deaths because of this COVID-19 pandemic are confirmed. Because of this worldwide medical emergency, severe public health measures are being taken to reduce the spread of COVID-19.



Figure 1.1 Social media has become a refuge for depressed people

Uncontrolled spreading of highly contiguous infections such as COVID-19, are often associated with mental uneasiness and indications of mental sickness (Bao et al., 2020)^[4]. Mental health specialists across the world ought to keep track of symptoms, their correlation, and techniques to store and organize them, they should also be aware of needs of specific groups of people (Yang et al., 2020)^[6] and, the special precautions required to keep this pandemic in control (Liu et al., 2020a)^[7].

Communication is the basic requirement of modern society, and at present, communication of information through the methods of web-based media like micro blogging websites is one of the most prevalent forms of communication and way to share our thoughts and opinions about things going around us. We utilize web based media to express our thoughts or opinions, to share our mental well being and our aspirations (Calvo et al. 2017)^[9]. People suffering from psychological illness are bound to find themselves sharing their thoughts and opinions on the web, either through microblogging websites like Reddit, Twitter, Facebook, etc, or Social Network Service (SNS), or explicit public discussions (Gowen et al., 2012)^[10]. As we compose carefully to an ever increasing extent and communicate with our knowns on the web, these large volumes of text blobs could be directly used to analyse the mental state of the author that wrote that piece of text.

Early detection and intervention in the lives of people with mental disorders provides the opportunity to incorporate new kinds of treatments like better counselling or connecting people who are not well with peers having similar interest since it could lead to improvement in the quality of treatment for these conditions as there has been an undersupply of traditional treatments (Coppersmith et al., 2015)^[12]. Utilizing online social interaction of these people has yet one more favorable advantage, these people don't need to feel uneasy due to social pressure as they do in case of physical screening for mental conditions. Using web based interaction, we can simply capture their data from the platform we are aiming and then we can screen that content. Doing so opens up new opportunities for early detection and prevention of people suffering from mental conditions like depression, stress or ADHD.

Psychological wellness represents a major test for a person's health. Text investigation of rich assets, similar to web-based media, can add to a more profound comprehension of diseases and give better insights to their early detection. According to the WHO, as much as 300 million individuals complained of health related issues, the Online Social Network (OSN) provides analysts another point of view for recognizing people influenced by the aforementioned mental conditions. With the rapid development of the web social network (OSN) like Twitter, Facebook, and Reddit, people are more frequently using the OSN to precise opinions and emotions.

1.1.2 Reddit

Reddit may be a popular social media platform for discussing a good range of topics, but it's a crucial source of data for data analysis on social media because it provides rich structure, an abundance of knowledge, and covers a broad range of topics. Reddit is employed by approximately 330 million users, with 2.8 million comments written every day. Alexa.com ranks it because of the 21st hottest website worldwide. Despite its popular and rich data, few have considered Reddit as a source of knowledge about users' personal traits like their professions and hobbies. Prior work has focused on Reddit as a source of demographic information, whereas we consider rich attributes like profession and hobbies additionally to demographic ones (age, gender, family status). Such data has many applications, including personalized healthcare (Gyrard et al., 2018)^[14], recommendations, search, and conversational agents.

Reddit is a web network of communities that supports people's interests. The various Reddit communities are mentioned as subreddits. Each subreddit is dedicated to a selected topic. People resort to such communities for several purposes. Some communities promote sharing recovery experiences and emotional support. We have used Reddit data to look at and analyze the increase of a spread of psychological state conditions like depression and other clinical diagnoses during the pandemic of COVID-19. One major reason for using Reddit is that the presence of subreddit rules and moderators on the platform causes the users to stay to the subject or risk banning their posts.



Figure 1.2: The Reddit Logo

Plenty of subreddits were chosen for our analysis. These subreddits were chosen on the basis of their rules, which restricted the users to talk about their mental health struggles on the subreddit. These subreddits were; r/ADHD, r/Anxiety, r/Autism, r/BrosOnToes, r/Neurodiversity, r/PTSD, r/SocialSkills, r/SuicideWatch, r/Depression, r/Stress, r/Offmychest.

Subreddit	Subscriber Count (as of December 25, 2020)
r/ADHD	1,101,745
r/Anxiety	417,585
r/Autism	91,054
r/BrosOnToes	1,334
r/Neurodiversity	14,028
r/PTSD	53,629
r/SocialSkills	2,175,818
r/SuicideWatch	247,995
r/Depression	712,870
r/Stress	9,430
r/Offmychest	2,287,821

Table 1.1: The selected subreddits and their number of subscribers

1.2 Related Work

The paper Ravi et al., (2020)^[11] reviews and summarizes the literature available in relation to mental health and its symptoms shown by people during COVID. There were a total of 26 journals out of which 5 different general topics were categorized, then they were used to structure the study.. The aim of this paper is the main motivation behind this project, as our project aims to find the effects of COVID on mental health of people during this period. Preliminary evidence in the study suggested that general effects of COVID19 pandemic on the mental conditions of the people included hysteria, depression, anxiety and stress.

19

In Ivan et al ., (2020)^[1], they made use of the SMHD dataset (Cohan et al., 2018)^[15], which contains a large number of users on Reddit that have experienced or suffered from more than 1 mental condition. Then they explore the limitations of their model with respect to the data required to attain successful results, particularly, the amount of users and their posting frequency.

For extracting 9 different mental illnesses which were reported by the users themselves, the linguistics patterns were used to obtain the users. For instance , let's say someone writes “I was officially diagnosed with depression last year”, then the person has been diagnosed by depression.

9

They adapt a neural Network (HAN), which is generally applied for classifying documents, and classifying people on sites like reddit, twitter etc. The network has a word sequence encoder, a word-level attention layer, a sentence encoder and a sentence-level attention layer. An embedding of a sentence is generated which is then passed forward for creating representation of a document. Attention mechanism is used in both types of encoders to obtain more correct embeddings.

So they concluded that various mental conditions have some common symptoms, the relation between them can help us in getting better results.

For getting better results, we can employ transformer based models that can be used to encode posts of different users on social media.

In Yiding et al ., (2020)^[2], they manually built a dataset on Sina Weibo called as WeiboUser DepressionDetectionDataset (WU3D), containing 30k users out of which 66.67% users are non-depressed and 33.33% are suffering from depression. The annotation is done and verified by people having proficiency in the domain. A total of ten attributes are obtained by studying the user's behaviour, their online activity, text in the posts, photos uploaded etc.

Many current studies aren't user-oriented modeling. Those works usually aim to research and model the language sort of the user. Through sentiment analysis and engineering of the tweet text, a classification model is developed to detect whether a selected tweet features a depressive tendency. These works analyzed fine-grained features and achieved pretty good results. However, such results can't be directly applied to user-level depression detection, or it's going to cause an incorrect prediction.

So how they collected the user IDs for depressed candidates is, according to their observation, most of the depressed candidates post heavily on depression related sub-groups. So they crawled through every sub-group whose title contained 'depression'. This ensured that the users labeled as depressed are actually depressed.

Also they have maintained the information of how many posts are posted by a user between 6pm to 6am, which is the high time for users in distress to have social media activity. There are also some statistical features which help us get more accurate results

and some more insights about the trend are frequency of depression related posts, as I just said proportion of late night posts, posting frequency per week.

So the reason we chose Reddit was that it also has sub-groups called subreddits where they have pre-defined rules for posting which ensure that we get the right labels for the depressed users. As any classification task it is very important to have correctly labeled data for training to have best results and this paper mainly focused on creating the dataset that is why this paper was a huge inspiration for our data collection methods.

In Razan et al ., (2020)^[8] Detecting topics of interest for people suffering from Anorexia Nervosa who are reddit users and declared themselves as the patient of AN. They provided a dataset of annotated posts labelling the relevant topics for the post, Baselines to predict the different posts categories based on the labeled dataset. In this paper, they are targeting the users suffering from anorexia nervosas and detected topics of interest i.e topics which are most commonly talked about among them. Provided a dataset of annotated posts labelling the relevant topics for the post. This work helped us get our step of data collection right, as instead of just looking for sub-reddits named ‘depression’ or any other disorder, we can also look for subreddits which are based on the topics of interest of people suffering from a disorder.

Prasadith et al ., (2019)^[5] studied the effect of trends and patterns pointed out by professionals working in the field of medicine, to improve the classification results of a model built using ANN. Users suffering from depression tend to put phrases or words containing negative sentiments. An empath is a text classification application which categorizes the text into 200 categories which are very common in a dataset available online .These categories are verified manually. They used this software on a set of posts of a user by combining them into one document. As a result, some interesting attributes are :

users suffering from depression generally use themselves as the subject in their post .

Also they generally post longer posts as compared to normal users.

1.2.1 Shortcomings of the related works and Key Challenges

After having a basic understanding of the related work done above, it had been observed that they mostly just classified the user base into whether or not they are unsound or not. Some people went a step further to seek out what specific problems or conditions they're browsing. But none of them put within the effort to seek out what proportion are people suffering from the explanations they're concluding. what's the trend that's being followed by the number of affected people, or what's the severity of the condition an individual is suffering from?

Also consistent with Ivan et al. (2020)^[1], for detecting depression in the users on social networking websites, all their activity/posts are combined into one. Then a hybrid of convolutional and recurrent neural networks²⁸ are used for classifying this document which do so by partitioning the whole document into segments. This method provides better results in case of classifying scientific papers, but having a straightforward partitioning of a user's activity is unintuitive and does not give optimum results.

In Yiding et al. (2020)^[2], The characteristics and behavior patterns of depressed users are going to be further analyzed. they're going to propose simpler feature solutions for user-level depression detection on the OSN.

Alina et al.^[3], the method of predicting depressed users wasn't automated. Had it been an automated procedure, those users might get additional treatment or a more thorough examination.

Also, some papers relied on the self-declaration of the user to be depressed or not, counting on whether or not they have specifically stated that they're affected by depression or other mental conditions. At the same time, this detection should be

automated and will use user behavior and his semantic and linguistic attributes in his posts to detect his condition.

1.3 Problem Addressed

There has been a surge in stress and anxiety levels among the overall public during the pandemic. There are a number of effects caused by stress during times like covid, like:

- frightened about the health of your close ones.
 - Fear and worry about your financial situation or job.
 - effect on sleep cycle of the people, or daily routines.
 - lack of concentration.
 - consumption of addictive substances like alcohol surges.

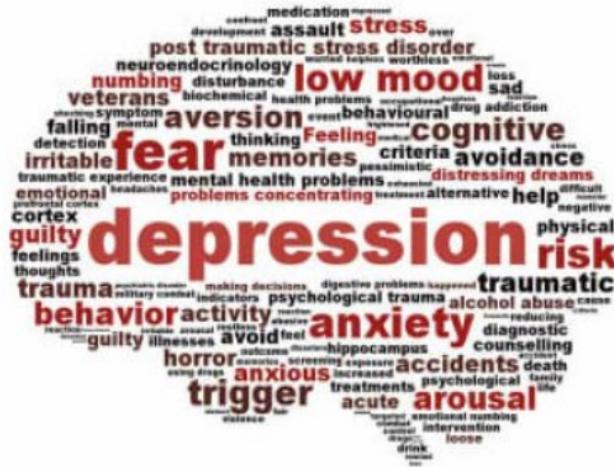


Figure 1.3 Depression and anxiety are harmful for mental health

The problem of psychological state deterioration is the main focus of our project. Through this project, we aim to know how the pandemic has affected the psychological state of the people and the way they're coping up with it.

We have kept the scope of our study pretty long, i.e., nearly 8-7 months because consistent with Ravi et al. (2020) [11], the long-term psychological state effect of

COVID-19 often takes weeks to become fully observable and visible, and tackling or treating these effects requires a collaborative approach from doctors, health workers, therapists and entire healthcare system

This problem may be a lot more serious than it first appears to be because COVID has caused damage to the psychological state services in a major part of the world in spite of increasing demand, according to the WHO. This survey conducted in 130 countries shows the impact of the pandemic on the psychological services and the need for financial aid through statistical data..

And the pandemic is increasing demand for psychological state services. Deprivation, loneliness, unemployment and panic are causing psychological state conditions or worsening the current condition .As described earlier some common consequences of stress during this time like, changes in sleep cycle, increased alcohol consumption, anxiety etc are being experienced by most of us. Some mental conditions like dementia, restlessness, anger, and stroke are direct effects of covid itself. People that are already suffering from these kinds of disorders are more vulnerable to severe effects or death in certain cases.

This project and therefore the analysis that it'll provide will give us very crucial insight in telling us how the pandemic has affected the psychological state of the population and what level of intervention and steps are required by the govt and also on a private level for the well-being of the population.

CHAPTER 2

2.1 Approach

This project approach that we are demonstrating involves various different phases. These phases being data collection and identification of mental health impacted users because of COVID-19 and then collection of their respective posts. The second phase involves creation of an annotated depression classification data-set. The next phase takes us through the journey of Natural Language Processing (NLP) and by the end of that phase we have a depression classification model. The last stage involves classification of the selected user posts that we have collected using the model.

2.1.1 Methodology

To analyse the depressed users on Reddit more effectively, we propose a novel framework. This framework mainly consists of three parts.

1. **Data Collection and labelling:** The first step is to choose a certain number of people on Reddit that fall in the category of people suffering from some other mental health disorders like anxiety, insomnia, depression etc. To get users that are suffering from any mental issues, we reasonably assumed that users posting in mental health subreddits are suffering from the same illness as its subreddit. This premise streamlined the data collection process by eliminating the manual labour needed for labelling users. The subreddits which were used for collections purposes were ADHD, Anxiety, Autism, BrosOnToes, Neurodiversity, PTSD, SocialSkills, SuicideWatch, Depression, Stress, Offmychest. To further make sure that the people we are choosing are seriously affected by COVID-19 one way or the other, we selected posts only from a certain time period (15' Feb 2020 - 30'Sep 2020) and with keywords like 'covid', 'coronavirus', 'pandemic'.
2. **Model Training and predicting:** We created a natural language processing model which can accurately predict whether a particular user is suffering from

mental illness. This is a crucial part of our project as it enables us to analyse the magnitude of impact from the pandemic by labelling users. Our machine learning model utilises the state-of-the-art transformers. The main motive of the transformer models is to provide an accurate dependency between the input and output word embeddings and determining which words to provide attention to. The self-attention or the attention mechanism used in the transformers are the main highlight. They help us to determine how the words relate to each other and which words to focus on.

3. **Visual Depiction of Results:** On getting results from our machine learning model, we plotted our findings using matplotlib library in a Jupyter notebook. Matplotlib is a graph plotting library for Python language which along with another library NumPy opens a lot of gateways for visually representing any set of data. We depicted the ratio of reddit users who were depressed before COVID-19 as compared to after it. Apart from our results we also plotted some general data which can further support our thesis like a bar graph which shows the distribution of users having posts which amounts to a fixed ratio before COVID-19, user engagement on particular subreddits distributed along different months , number of subscriptions of a subreddit for each month between March and September for both 2019 and 2020.

2.2 Data Collection

As described in the introduction of the report, the aim of this project is to study the effect of COVID-19 on the mental health of the general public. Therefore it was crucial to gather an inclusive dataset for our study. That's why the first step is to choose a certain number of people on Reddit that fall in the category of people suffering from some other mental health disorders like anxiety, insomnia, depression, etc. To get users that are suffering from any mental issues, we reasonably assumed that users posting in mental health subreddits are suffering from the same illness as its subreddit. This premise streamlined the data collection process by eliminating the manual labor needed for labeling users. The subreddits which were used for collection purposes were ADHD,

Anxiety, Autism, BrosOnToes, Neurodiversity, PTSD, SocialSkills, SuicideWatch, Depression, Stress, Offmychest.



Figure 2.1 PRAW API

Hence, if a user is found to post in the subreddit anxiety, then the user is labeled to be having anxiety problems. As our focus is aimed towards analyzing changes specifically due to the COVID-19 pandemic, we looked for the keywords ‘covid’, ‘coronavirus,’ ‘covid19’, ‘pandemic’ in user-submitted posts from 15’ Feb 2020 to 30’Sep 2020 and considered only such users and their posts. This time period (15’Feb 2020 to 30’Sep 2020) is hereafter referred to as the pandemic period, and the period before 15’Feb 2020 is referred to as the pre-COVID period. The actual collection of posts was done by a python script using a Reddit API push shift, and afterward, we extracted the users who posted those posts and created a set of those unique users.

Pushshift is an openly available RESTful API that provides abundant functionality for searching Reddit data and incorporates making sturdy data aggregations. Using this API, we could promptly find the info we were interested in along with finding fascinating correlations. There are two main ways of obtaining and handling the Reddit comment and submission database. Pushshift API is packed with useful features like handling paging with sophisticated sort options. It also holds rate-limiting and exponential backoff subject to maximum retries and maximum backoff limits. The default rate provided by the API is one request per second.

We obtained 99334 unique authors and 147034 posts mentioning the given keywords after filtering the removed posts and deleted users. Now, after this step, we have all the users who are suffering from some kind of mental health problems but are also impacted in some way by the pandemic. The collected posts and users were preprocessed to remove any URLs or usernames that could potentially contain sensitive information.

Later, using a python script and praw API for Reddit, we fetched the previous posts before our prescribed pandemic timeline, i.e., before the pandemic(before 15'Feb 2020) for every user in the list. PRAW, which is ³ an acronym for “Python Reddit API Wrapper,” is a Python package that facilitates easy access to Reddit’s API. PRAW aims to be easy to use and internally follows all of Reddit’s API rules. Various different parameters of a post were recorded, such as id, URL, author, created_timestamp, upvotes, title, text, subreddit.

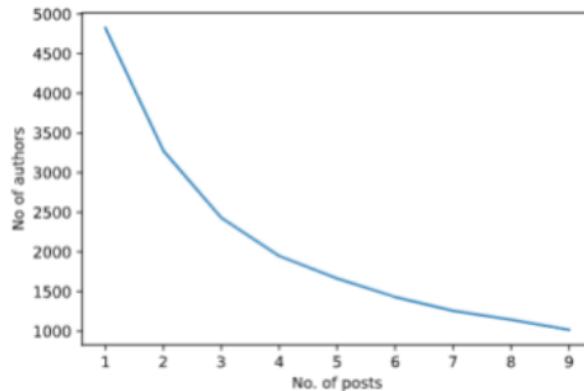


Figure 2.2: Plot of number of authors with given number of posts

It proved to be a tedious task and was parallelized by running simultaneously on four computers, with each processing approximately 25000 users and taking a total time of 12 hours. We selected only those users who have sufficient Reddit activity to get an insight into their mental health before the pandemic. Since there can be a lot of new Reddit users among our authors who did not have any past posts for our evaluation, we visualized the number of posts from our set of 99334 users. It was found that most of

the users only had one post, and a threshold was decided, which denoted the minimum number of posts a user must have before our timeline to be considered for evaluation. As there were nearly 32000 users having four or more posts before the pandemic, we decided to put the threshold at 4.

Now that we have our users on whom we are going to base our study. The next step was to detect their condition before COVID-19. For nearly 32000 users, we extracted close to 2.9 million posts and 2.73 GB of data, and these posts were later passed through a classifier to get some insight into the mental conditions of their authors before the pandemic.

To be able to get a robust machine learning model, our collected data should be diverse with depressive (posts from people with mental health issues and which highlight these problems) as well normal posts (posts from casual or happy users). Therefore, we also collected posts from some of the subreddits that are known to have positive content. These subreddits include wholesome, positivity, positivethoughts, humansbeingbros, etc. We didn't maintain a record for authors of these posts as these posts were only beneficial for model training and to label whether a post is depressing or not.

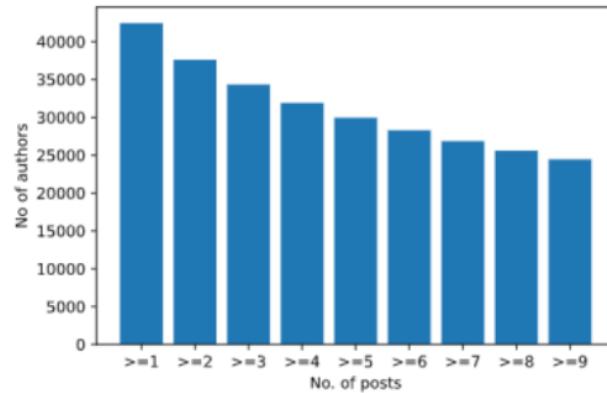


Figure 2.3: Plot of number of authors with greater than or equal to a given number of posts

Even though we rely on publicly available Reddit posts in our work, mental health is a delicate matter. Data collection should always consider measures to prevent individuals' risk in social media research (Hovy and Spruit, 2016; Suster et al., 2017). The perils

associated with our data collection methods are minimum. We abstained from openly posting any extracts of the data. We also made no such attempt to contact users and did not try to identify or associate users to other social media accounts. We replaced usernames with arbitrary identifiers to prevent users' identifications from being identified.

2.4 Classification Model

2.4.1 Comparison between Transformers and RNN/LSTM

Five years ago, RNN-LSTM based models were the go-to choice for NLP classification tasks. RNN-LSTM, alongside semantic and lexical features and various word embeddings present, was quite successful in various NLP domains.

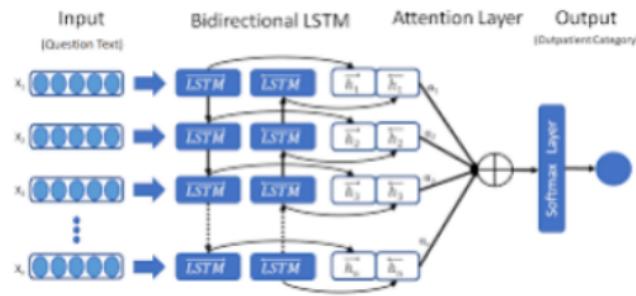


Figure 2.4: LSTM for text classification

15 Recurrent neural networks (RNN) and long-short term memory (LSTM) models are 6 traditional Natural language processing (NLP) techniques and they have very similar characteristics. Few of these characteristics are:

- Sequential Processing: Sentences or data provided to these models are processed sequentially i.e. either in original or reverse order. These sequences are mainly processed in a word by word manner.
- Past information retained through past hidden states (Vanishing or Exploding gradients): Sequence to Sequence (seq2seq) models follow the property where

each state or event is expected to be dependent on any previously observed state or event.

The sequential processing property is the reason why RNN and LSTM can only be processed in a linear fashion and not in parallel. So, when we compute the embeddings for the third word, we require the word embeddings for the first and second word and also of the previous line. The property of Information from the hidden states can be very easily understood conceptually. The information that is coming from the hidden states in RNN and LSTM can be arriving from the states too way back such that their effect on the new state will either be too profound or it will be almost negligible. When the effect is too profound, we say that there are exploding gradients in the model and when these effects are negligible, there is presence of vanishing gradients. LSTM (and also GruRNN) can solve the problem of vanishing and exploding gradients to some extent by simply allowing the previous gradients to pass through without any processing but these steps make the models more slow without providing any out of the box improvement to the results. To solve some of these problems various Bi-directional models have been introduced like the bi-directional RNN and the bi-directional LSTM but these models do not have a true bi-directional nature. The final result is simply the aggregation of the result from the two directional processings and does not provide any extra contextual information.

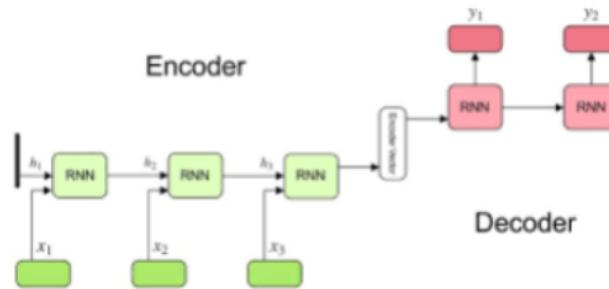


Figure 2.5: Transformer and attention mechanism

The Transformer model and its application of self attention was proposed within the paper [Attention Is All you need](#) ([Ashish Vaswani et al., 2017](#))^[14]. Quoting from the paper:

"The Transformer is that the first transduction model is relying entirely on self-attention to compute representations of its input and output without using sequence-aligned RNNs or convolution."

The word "transduction" in ² the above quotation means the conversion of input sequences into output sequences. The main motive of the transformer models is to provide an accurate dependency between the input and out word embeddings and determining which words to provide attention to.

Some of the important characteristics of transformers are:

- Non-sequential Nature: Sentences or the data provided are processed by the transformer as single entity and every word of that sentence is processed in a parallel manner at the same time
- Self Attention: This is the most fascinating feature and novel technique introduced by transformers. The self attention feature processes all the words of the sentence and then decides which words require the most attention in that sentence to understand its context.
- Positional embeddings: Positional embeddings like any other embedding provide extra insight to the true context of the sentence by processing the position in which the words occur in the sentence.

The property of non-sequential nature helps the transformers to counter the problem of vanishing and exploding gradients as seen in RNN and LSTM and also allow for parallel processing. The self attention information along with the positional embeddings, both provide information about the context and meaning of the sentence and how various words interact with each other.

2.4.2 Transformer Architecture

Focussing on the transformer architecture through the image below, the transformer consists of an encoder block which consists of one layer of Multi-Head Attention. This

2

encoder block then transfers the embeddings to the layer of Feed Forward Neural Network. The decoder has a similar architecture to that of the encoder but has an additional Masked Multi-Head Attention. The masked here means that while calculating attention in decoder, it only takes into account the words that it has already decoded.

In reality, in a transformer there are multiple units of encoder and decoder that are stacked on top of each other. The amount of encoder and decoder units in a model can often be decided and used as a hyperparameter. Within the paper, six encoders and decoders are used.

The encoder and decoder in the transformer works in the following way:

- The word embeddings of the sentence after passing through initial embeddings and positional encoder are passed to the primary encoder.
- The processed embeddings are then forwarded to the next encoder and so on.
- The output from the last encoder from the encoder-stack is passed to either all or some of the decoders within the decoder-stack.

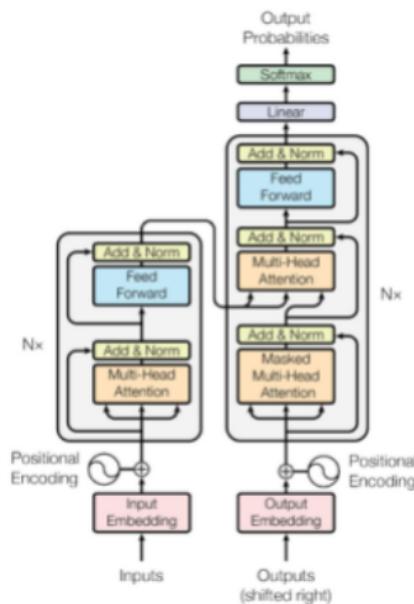


Figure 2.6: Transformer Architecture

The self-attention or the attention mechanism used in the transformers are the main highlight. They help us to determine how the words relate to each other and which words to focus on. According to the paper:

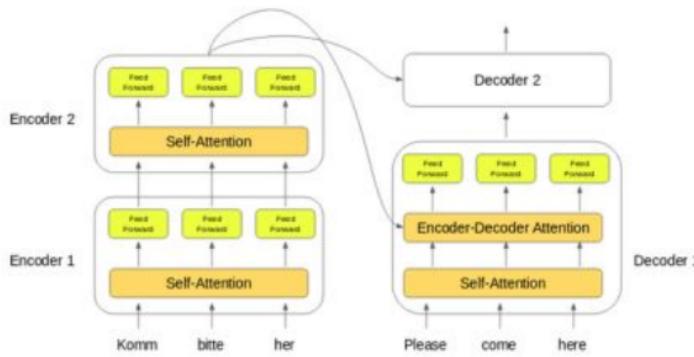


Figure 2.7: The encoder and decoder operation for translation tasks

“Self-attention, sometimes called intra-attention, is an attention mechanism relating different positions of one sequence to compute a representation of the sequence.”

²⁷ Self-attention allows the transformer models to determine how the words relate to each other and which words to focus on.

2.4.3 Transfer Learning using Transformers

Earlier transfer learning was a technique to design very powerful and effective classifiers with minimal effort in the field of computer vision only. But because of ²⁵ transformers, transfer learning is now possible in Natural Language processing (NLP) tasks as well. Various versions of these transformers can be trained on massive text corpus by companies with resources and resulting word embeddings can be used by end users to train classifiers at almost no cost.

The BERT framework is one such example, which is a transfer leaning transformer based model from Google AI. BERT, which has 12 layers and 768 hidden units and 109

million parameters has been pre-trained by Google on massive Wikipedia English text corpus.

Other transfer learnings and transformer-based frameworks that are available are XLnet, Longformer, DistilBERT, and Roberta.

2.5 Tools and Libraries Used

2.5.1 Python

Python is a scripting, high-level, object oriented programming language and can also be used for general purposes. It uses an interpreter for compiling the code. whose design philosophy emphasizes code readability with its notable use of whitespace. Its ability to write code fast is beneficial for developers to write prototype code as soon as possible. In our project, we primarily used Python for its ease of use as well as writing down the code swiftly. Moreover, the plethora of libraries which are available for various use cases, especially machine learning like TensorFlow, scikit-learn, matplotlib, nptk, etc.



Figure 2.8: Python Logo

2.5.2 Matplotlib

Matplotlib is a utilitarian library used by machine learning practitioners and data analysts for generating static, animated, or interactive visualizations in Python language. Matplotlib takes advantage of another famous library NumPy to open a lot of gateways for visually representing any set of data. Matplotlib was originally developed by John

Hunter, but now it is mainly supported by many contributors, who have put an uncountable number of hours and energy into producing a software that is utilized by thousands of scientists worldwide.



Figure 2.9: The Matplotlib library

2.5.3 Pushshift

Pushshift is an openly available RESTful API that provides abundant functionality for searching Reddit data and incorporates making sturdy data aggregations. Using pushshift, we could promptly find the information we were interested in along with finding interesting correlations. There are two main ways of obtaining and handling the Reddit comment and submission database. Pushshift API is packed with useful features like handling paging with sophisticated sort options. The default rate provided by the API is one request per second.

One crucial capability which this API provided was to specify a time frame within which we want to collect our data as well as the ability to focus our query on a particular subreddit.

The API call used for our use case was:

```
'https://api.pushshift.io/reddit/search/submission/?query='+str\(query\)+'&size=500&after='+str\(after\)+'&before='+str\(before\)+'&subreddit='+str\(sub\)'"
```

Where, query is the keyword to filter the results, after is the time in Unix timestamp after which we want the result, before is the time in Unix timestamp before which we want the result and subreddit is the subreddit we want to query.

The API by default sorts by the most recently submitted posts. After running the above mentioned call, 500 results were returned. Data was returned in a JSON (JavaScript Object Notation) format and results were present in the "data" key.

2.5.4 PRAW

PRAW stands for “Python Reddit API Wrapper,” ³ is a Python package that facilitates easy access to Reddit’s API. It intends to be easy to handle and follows all of Reddit’s API rules. It is basically a wrapper that hides actual calls to the official Reddit API and instead exposes methods and classes from which a developer can query Reddit.

PRAW requires a Reddit instance, which is made by first enabling developer tools for a Reddit account, then creating an app using old Reddit. On successfully creating an app, Reddit grants us with a Client ID and Client Secret. These two values are required to access Reddit’s API as a script application. After creating a Reddit instance, PRAW provides different classes and methods to query subreddits and users effectively. The python code used for gathering past submissions from a specific Redditor is shown as follows:

```
reddit.redditor(author).submissions.new(limit=1000)
```



Figure 2.10: Praw API

2.5.5 Transformers: Hugging-Face

Transformers is a library by the hugging Face team that gives incredible results in NLP based tasks. It is a library that is built on top of Pytorch by Facebook and TensorFlow by Google. Hugging Face ¹⁶ Transformers library provides transfer learning transformer based architectures for Natural Language Generation (NLG) and Natural Language Understanding (NLU) and with more than 32 models that are pre-trained in around 100

spoken and written languages and provides deep ability to exchange and use information between TensorFlow and PyTorch.

- The computation overhead is less, and hence less carbon footprint:
- Users can use a common model instead of taking turns to train increasing redundancy of the task.



Figure 2.11: Hugging Face Transformer Library

- Users can further improve the computation time and thus create optimal models

2.5.6 Simple Transformers by Thilina Rajapakse

Simple Transformers by Thilina Rajapakse is a natural language Processing (NLP) library whose main feature is its 'just works' capability. This library is a very simple way to apply transfer learning and work with transformers. This library has been built on top of the transformer library by hugging face. This library makes use of apex library by NVIDIA for very fast training procedures. This library has been designed to simplify the application of Transformer models without losing out on accuracy and performance of the models. The most prominent reason to use this library for this project to coach our models is that the library's just works nature.

Simple transformers are designed to interpret the application of Transformer models without compromising performance and efficiency. They pertain to certain benefits like optimizing input data formats for specific tasks, generating a clean and ready-to-use output.

Simple transformers make it very easy to train models and then later save those models and also helps in transporting them across python files with minimal configuration required.

2.6 Experiment and Evaluation

2.6.1 Training Dataset

To train our classifier we needed an annotated data-set containing two fields; one, the text or some Reddit post text that we need to classify in the end, two, a label saying whether that text sounds depressed or not. To create the dataset, we extracted depression related posts from the subreddits listed in earlier slides. All these posts were given the label of “Depressed”.

Later we extracted posts from many different subreddits like r/Thoughts, r/ReviewThis, r/culture, r/positivepsychology, r/love, r/travel and r/MovieSuggestion to get a wide variety of non-depressed posts.

The final data-set had 50,242 samples containing an equal number of posts from the two labels.

2.6.2 Experimental Setup

Due to lack of a dedicated graphics card for model training, we had to resort to Kaggle platform for training purposes. Kaggle is an online community aimed towards data scientists and machine learning enthusiasts. It provides its users with free GPU sessions to train their models. We had access to 12GB of RAM and 40 hours of GPU with 15GB of graphic memory. Uploadation of the training and test data sets were needed prior to moving to the training phase. The training dataset amounting to 2GB was uploaded along with a testing dataset of 500 MB. To accelerate the training process, the training set was divided into four parts and were trained individually by each one of us.

The model architecture used for training was simple. The first layer being the transfer learning transformer based layer which was followed by a dense layer.

Following were the different specifications that were used for different frameworks

	Variation	Architecture
BERT	<i>bert-base-cased</i>	12-layer, 768-hidden, 12-heads, 109M parameters.
RoBERTa	<i>roberta-base</i>	12-layer, 768-hidden, 12-heads, 125M parameters
XLnet	<i>xlnet-base-cased</i>	12-layer, 768-hidden, 12-heads, 110M parameters.
Longformer	<i>allenai/longformer-base-4096</i>	12-layer, 768-hidden, 12-heads, ~149M parameters
DistilBERT	<i>distilbert-base-cased</i>	6-layer, 768-hidden, 12-heads, 65M parameters

Table 2.2: The Transformer based model and their architecture used

Hyper-parameters

1. 'Learning_rate':1e-5
2. 'num_train_epochs': 5
3
3. 'reprocess_input_data': True
4. 'process_count': 10
3
5. 'train_batch_size': 10
6. 'eval_batch_size': 10,
7. 'max_seq_length': 512
8. 'fp16': True

2.6.3 Training Results

The above stated frameworks were used to create respective models and were trained on the above created dataset. The results of training are as follows for various models.

The ROBERTa transformer based transfer learning framework gave the best results upon training on our dataset and therefore will be used for evaluation purposes.

Model Name	Accuracy	Macro-AVG	Weighted-AVG	Label "Depressed"			Label "Non-depressed"		
				Precision	Recall	F1-score	Precision	Recall	F1-score
BERT	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.95	0.95
RoBERTa	0.97	0.97	0.97	0.96	0.97	0.97	0.97	0.96	0.96
XLnet	0.94	0.94	0.94	0.95	0.94	0.94	0.94	0.95	0.94
Longformer	0.96	0.96	0.96	0.95	0.96	0.96	0.96	0.95	0.96
DistilBERT	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95

Table 2.3 Transformer based model with their respective accuracies

CHAPTER 3

3.1 Results

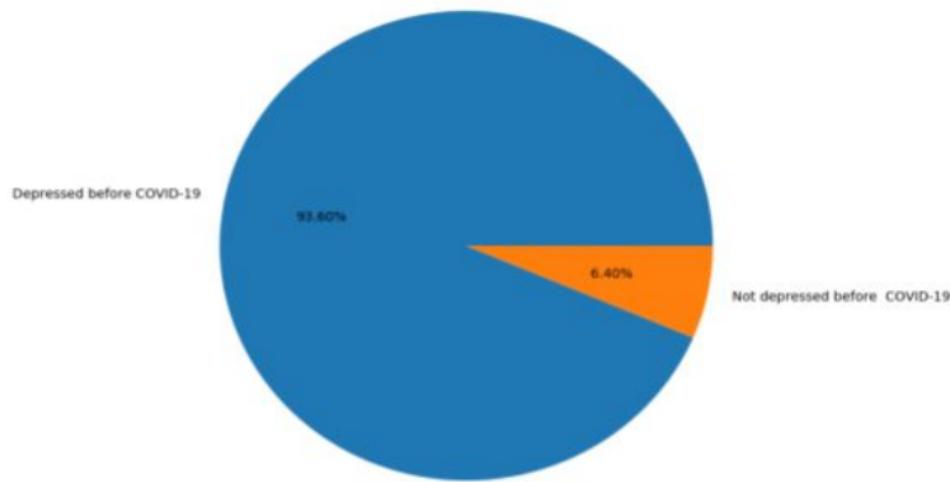


Figure 3.1: Ratio of Reddit Users who were depressed before COVID

This pie chart shows the ratio of Reddit users who were depressed before COVID-19 as compared to after it. We have divided the timeline as pre and post COVID-19 by choosing 15 Feb 2020 as the dividing date. This date was chosen after referring to other research papers and analyzing the trend of new COVID-19 cases. We found that after 15 Feb 2020, the new COVID-19 cases started growing exponentially throughout the world.

As shown in the pie chart, 93.60% (28915) of the total users (30892) we collected were already depressed before COVID-19, whereas 6.40% (1978) users got depressed after the pandemic struck. This shows that there could be a direct correlation between new depression cases and COVID-19. We have collected the entire history of this 6.40% of the users, and we can analyze and prove the effect of COVID-19 on their mental state.

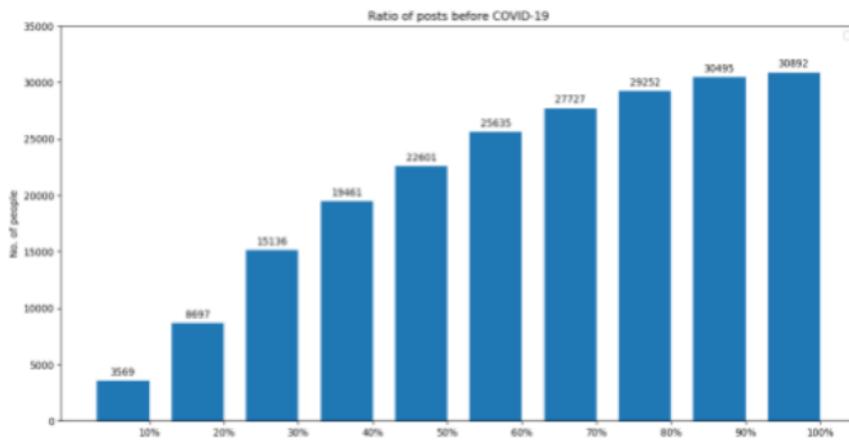


Figure 3.2: Distribution of users having posts which amount to a fixed ratio before COVID-19

This bar graph shows the distribution of users having posts, which amounts to a fixed ratio before COVID-19. This shows that 3569 users have 10% of their posts before COVID-19 and the rest 90% after COVID-19; similarly, 8697 users have 20% of their posts before COVID-19 and the rest 80% after COVID-19, and so on. We used this data to determine whether we are getting enough data to run our analysis model on it or not.

This data can be used to infer how new users are behaving on the platform and how COVID-19 has affected users' behavior on Reddit. The 3569 people that we found having only 10% of their posts before COVID-19 proves that we got active after the COVID-19 pandemic struck the world, and hence it can be used to check the correlation between COVID-19 and the mental state of people throughout the world.

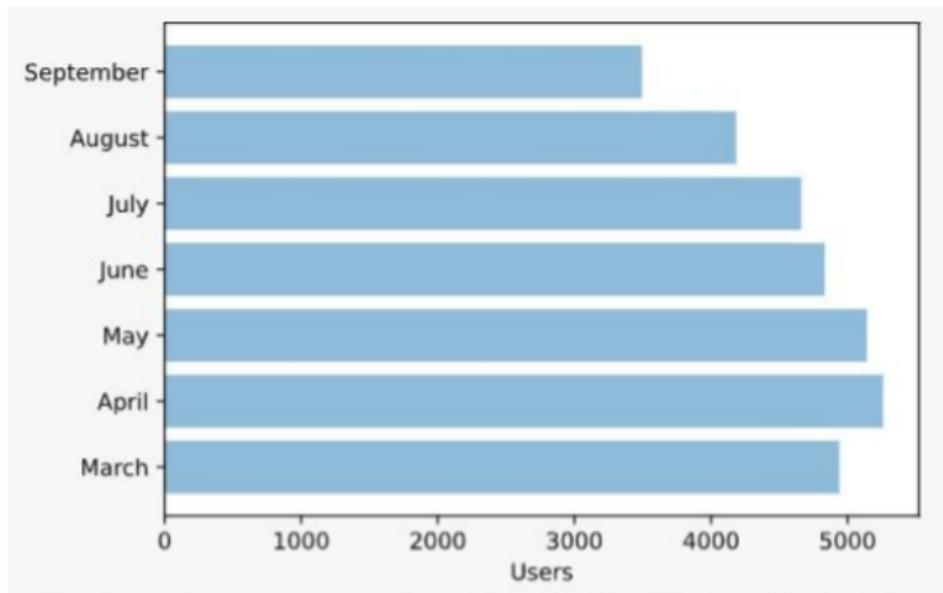


Figure 3.3: Number of Users per month posting in subreddits for the first time

This horizontal bar graph shows new user activity after COVID-19 per month.

It could be seen that in March, around 5000 new users logged on to depression-related subreddits for the first time; similarly, in April, more than 5000 users posted for the first time. This graph also shows that as the COVID-19 pandemic progressed, we continued to see high activity by new users, which shows that COVID-19 affected the mental state of users. Also, as the pandemic progressed, this effect slightly diminished, and we might continue to see a downward trend in the new users' activity.

Even in September, we saw around 3500 new users posting on r/depression for the first time. This means that they were not depressed before the pandemic. A vast majority of the new posts that were coming on r/depression were by new users only, old users were also active on this subreddit, but they were most active in the comments section of these new posts.

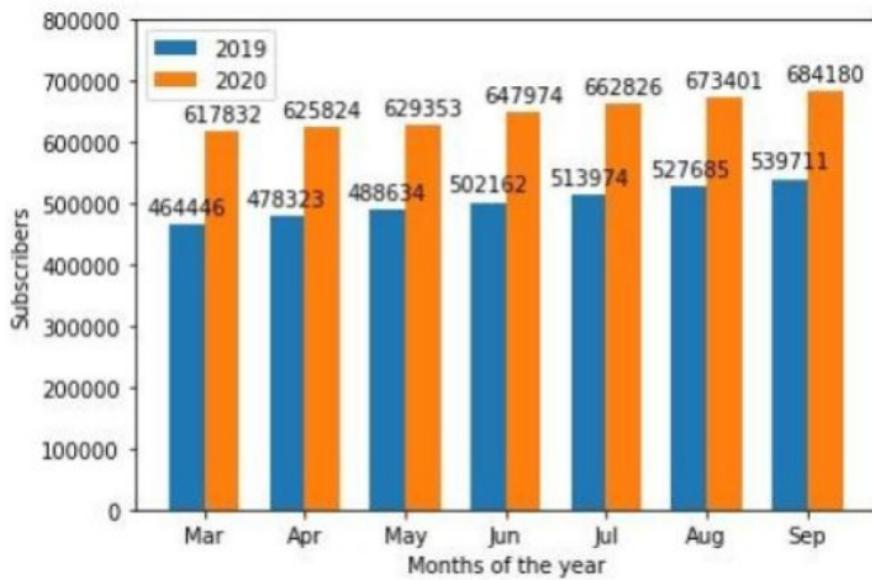


Figure 3.4: Number of subscribers on r/depression subreddit per month for the year 2019 and 2020

This bar graph shows the number of subscribers on the r/depression subreddit per month for the years 2019 and 2020. It could be seen that the total number of subscribers of r/depression was lesser in 2019 as compared to 2020. In March 2019, there were 464446 subscribers, and in March 2020, there were 617832 subscribers. We saw an increase of 153386 subscribers in just one year. Similarly, from April 2019 to April 2020, there was an increase of 147501. Also, from May 2019 to May 2020, there was an increase of 140719. In Jun 2019, an increase of 145812. This shows that in March, the growth was higher as compared to other months. This month correlates with the starting of the pandemic, so a large number of new users started posting for the first time when the pandemic started.

In both years, we can see growth in the number of subscribers. But, the growth was much larger in 2019 at 75265, whereas we see a growth of 63648 in 2020. We will try to figure out why this is in our future work.

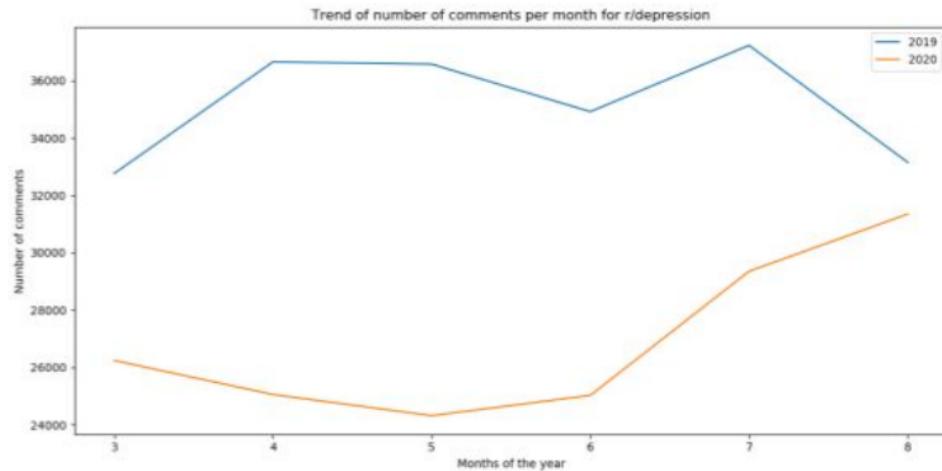


Figure 3.5: Number of comments in the r/depression subreddit per month for 2019 and 2020

This line graph shows the trend of the number of comments in the r/depression subreddit per month for 2019 and 2020.

We can see that in all months of 2019, the new comments coming on r/depression posts were higher than any of the months in 2020, therefore, in general, the engagement was higher in 2019 as compared to 2020, but in 2020, we can see a larger growth towards the second half of the year. It shows that users got more active on r/depression as the COVID-19 pandemic progressed.

This graph shows a similar trend to the previous graph as it shows that engagement was higher in 2019 as compared to 2020. We will try to figure out why this is in our future work.

3.2 Conclusion

As the results have shown, COVID 19 has been one of the major reasons of depression among people which is evident by the results that 6.4% of our user base showed normal behaviour before the pandemic. Though this figure seems small but generalizing this result on a larger scale reveals that this is a significant amount. There are many factors which could be included for better classification results which may improve the legitimacy of the result.

The increasing number of new users posting on the reddit can also be supported by the fact that in the period from february to april or may, COVID was on its peak with still no significant developments finding the cure. This led to a state of fear among people, and getting accustomed with the shift to online methods of business and communication. Then the number started to fall which seems true, that people started to adjust with the situation and had started to work around the problems which were initially a huge change for them.

Now naturally as we have shown, there is a significant effect on the mental health of the people, we can see that there is an increase in the subscribers on the depression subreddits as compared to each month of 2019, with a difference of at least 1.5 lakh new subscribers.

But one thing which is not evidently clear is that although there are significant new subscribers , the user activity shows the opposite trend. The number of posts and comments are less than what they were in 2019. Hence as our future work, finding the cause for this trend is one of the goals.

For improving the accuracy of classification of depressed and non-depressed users, one of our steps was to use a transformer based model which outperforms other RNN-LSTM based models. Out of 5 Different models RobertA proved to be the best classifier. For further improving the results we will take into account other statistical features and linguistics trends in the posts i.e the pattern in which they use phrases in the text to determine the severity of different conditions to get a detailed insight in the effects of COVID.

3.3 Key Challenges Faced

This line graph shows the number of comments in the r/depression subreddit per month for 2019 and 2020.

We can see that in all months of 2019, the new comments coming on r/depression posts were higher than any of the months in 2020, therefore, in general, the engagement was higher in 2019 as compared to 2020, but in 2020, we can see a more extensive growth towards the second half of the year. It shows that users got more active on r/depression as the COVID-19 pandemic progressed.

This graph shows a similar trend to the previous chart. Reddit doesn't provide API to get posts/comments data for a time frame directly. We had to use third-party APIs to make multiple API calls for fragments of time. This caused difficulty during data collection from Reddit. We tried two different APIs; the first was PRAW, and the second was pushshift. None of the APIs were "full" in the sense that we required them to work. We made our code to utilize both APIs' best features to fulfill our use case.

Some subreddits are restricted, so getting data from them is not possible even with third-party APIs. We initially had a long list of subreddits that we were going to analyze and download data from. Due to these difficulties mentioned above, we had to drop some subreddits from our list to ensure our data collection script works fine. Earlier, it was giving errors if we were trying to collect data from these restricted subreddits.

Reddit doesn't store demographic data about users, so we can't get users' age, sex, nationality, income bracket, etc. We tried using Twitter to collect data by streaming new tweets and storing them into CSV files, but it was challenging to collect demographic data from Twitter. Also, it is challenging to collect historical tweets from Twitter. There is another issue with Twitter, i.e., Twitter tweets are not categorized; therefore, we had to go through an extra process of classifying the data before we can use it.

Since Reddit is an old platform, there are many dormant users. When we tried to collect all data from specific subreddits, we found many users who stopped using Reddit. Hence they don't have any new posts coming on the platform. We can't use these users' data to get any insights regarding pandemic on the users. Therefore, we had to filter out the unique authors that we found. We first collected the entire Reddit history of every user and then tried to filter them out based on their activity before and after the pandemic. We only chose those authors with a total number of posts above a specific threshold determined after examining the dataset we create.

Reddit doesn't restrict users from creating multiple accounts, so people sometimes have more than five stores. It isn't easy to track all their performances. This poses a problem during data collection and data cleaning, as well.

Reddit doesn't allow us to get a particular user's activity in a specific subreddit; we must collect the entire history and filter it accordingly. We wanted to manage every author's complete record, but we only needed their posts related to depression. We can't put any search query while collecting the entire history of Reddit users. Therefore we had to manage their full account of posts and comments and then filter them accordingly. This leads to a loss of wasted effort in collecting data that we will not use in our analysis.

Sometimes, useful Reddit posts are hidden due to their non-compliance with subreddit rules. There are specific rules that must be followed by Reddit users while posting any content on any subreddit. If these rules aren't followed, then the posts are liable to be deleted by moderators.

Due to the lack of a powerful dedicated graphics card for model training, we had to resort to the Kaggle platform for training purposes. ² Kaggle is an online community of

data scientists and machine learning enthusiasts and practitioners. Kaggle allows users to find, upload and publish data sets, explore and build models in a web-based data-science environment, collaborate with other data scientists and machine learning engineers, and enter competitions to solve data science challenges. We had access to 12GB of RAM and 40 hours of GPU with 15GB of graphic memory. Uploadation of the training and test data sets were needed before moving to the training phase. The training dataset amounting to 2GB was uploaded along with a testing dataset of 500 MB. To accelerate the training process, the training set was divided into four parts and trained individually by each of us.

As it shows, that engagement was higher in 2019 as compared to 2020. We will try to figure out why this is in our future work.

3.4 Future Work

As shown in the graphs in “Results” section, there have been more daily comments in 2019 as compared to 2020, also the growth of number of subscribers for r/depression has been more in 2019 as compared to 2020. We need to figure out what the reason behind this is. We can start reading research papers written in 2019 to try to figure it out. We can also read various news articles related to mental health written in 2019 and compare them with articles written in 2020. Another approach we can take is, talking with experts in this field, we can try to figure out the reason by asking people who have been working in this field for a long time. Figuring out this anomaly is important for the sake of completeness of our study.

In future, we can try to gather more data by focussing not only on Reddit for opinion and thought of people suffering from mental illness but also from other micro blogging sites like Twitter, Facebook, Quora, etc. Each of the above mentioned sites have their own sets of pros and cons. For example, in Twitter we can stream the tweets by users in real time but that also poses the issue of categorizing these tweets. Only after categorizing, we can really use the data collected from Twitter. Similarly, from Facebook, we can go to specific groups to collect data and there are lots of suitable

groups available on Facebook. But, these groups are often unmoderated hence the data collected must be filtered before using. Quora is another good site in terms of data generated by the users.

Another possible use case in future could be to generalize our approach to measure the effect of any event on the mental state of people. For example, we can try to judge the effect of specific social or political movements on the mental state of people.

If we can collect the demographic data of the people who are writing the posts then we can more precisely measure the effect of events on the mental state of people. Right now, we are using Reddit as a platform to collect data but we can't get data like nationality, age, sex, etc of the authors who are writing the posts, therefore our dataset is not complete. If we can use this data then we can precisely tell the effect of pandemic on every segment of population.

We can work on our code to create a fully functional API for Reddit which can collect data from Reddit for a specific segment of time and also automatically filter and categorize the data on the basis of search queries provided.

During the data collection stage, we had to give you two separate APIs to collect and categorize data. Using the aforementioned API, we can release our own API which combines the best features from both APIs and act as a standalone and fully functional API for Reddit.

Another possible use case of our research work could be to analyse any post from any source to look for signs of depression or other depression related traits.

We have built our analytical model training it on numerous posts collected from reddit by 30892 unique authors. In future, we will try to feed more data into this model to make it more accurate in detecting various linguistic patterns found in posts by people suffering from some mental condition like acute depression, stress, ADHD, etc. We'll also try to make our model classify the posts into various kinds of mental conditions.

We want our model to be able to parse any post/transcript/message and be able to accurately predict and classify the mental state of the author of that specific piece of text.

We can also deploy our machine learning model to various online microblogging platforms like Facebook, Twitter, Reddit, etc to monitor the posts in real-time in order to provide early detection of aforementioned mental conditions. Early detection will lead to timely intervention and may save some lives.

Our model would also be able to analyse these posts and dynamically generate the sets of linguistic patterns found in the post/transcript/message that point to a depressed state of mind. It would also be able to classify various cases, where depression was found, into severity level and hence we would be able to tell which particular linguistic patterns are found in mild depression cases and similarly, which ones are found in severe cases of depression.

We will compare these linguistic patterns with a control group, certified as “non depressed” and try to analyse the difference to be able to find the major depression causing factors in today's time. We can also extend this functionality to be able to parse any set of older text and tell what factors were major depression causing factors in older time. We can then try to figure out the trend in these depression causing factors.

Since we'll generate a set of symptoms that the depressed people are facing using our machine learning model. We can also cross verify these symptoms with the posts found on platforms like Reddit, because, in many cases, the posts on various platforms also list the symptoms the author is facing therefore, after doing the aforementioned analysis, we'll be able to cross-verify whether the post is genuine or is it simply a post to gain virtual internet points, as offered by many online platforms like Reddit.

Therefore, in future, we'll be able to analyse posts on online platforms in real-time and also generate the set of linguistic patterns found in the posts by the author having depression or depression related mental condition. Similarly, we'll also be able to generate a set of symptoms these people are facing.

These linguistic patterns can then be directly used to classify a post as depressed or not. Similarly, the set of symptoms could be used to diagnose a person for depression or depression related mental condition.

APPENDIX A

A.1 Transformer based Transfer Learning

¹² A.1.1 BERT

BERT (Bidirectional Encoder Representations from Transformers) is a breakthrough paper by Google AI Language. ⁶ BERT has achieved state-of-the-art results in almost all the language modelling benchmarks. BERT is based on the transformer architecture that has been pre-trained on massive text corpus by companies with resources and resulting word embeddings can be used by end users to train classifiers at almost no cost. BERT has allowed us to use transformers using transfer learning for Natural language processing tasks. BERT, which has 12 layers and 768 hidden units and 109 million parameters has been pre-trained by Google on massive Wikipedia English text corpus.

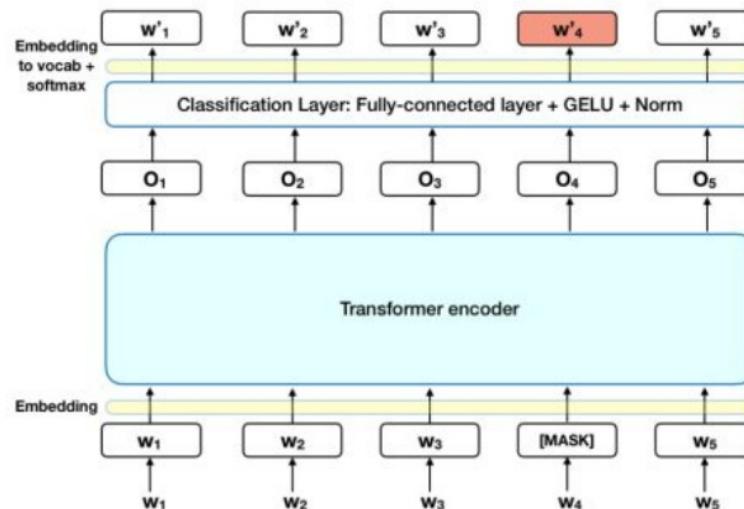


Figure A.1: The BERT architecture

A.1.2 RoBERT-a

³ RoBERT-a is a robustly optimized method for pretraining natural language processing (NLP) systems that provide improvements on Bidirectional Encoder Representations from Transformers, or BERT, which was released by Google in 2018. ⁸ RoBERTa, which was implemented in PyTorch by Facebook, ⁸ modifies important hyperparameters in BERT, and trains with much larger mini-batches and learning rates.

RoBERTa is an extension of BERT with changes to the pretraining procedure.

RoBERT-a's results show that tuning the BERT training procedure can significantly improve its performance on a spread of NLP tasks while also indicating that this overall approach remains competitive with alternative approaches.

A.1.3 XL-Net

XLNet is another pre-trainer transformer based transfer learning model just like BERT and RoBERT-a. XLNet is an auto-regressive (AR) model. An autoregressive model is a model that uses the context word to identify or predict the next word in both the direction i.e., both forward and backward direction. XLNet is another model that has

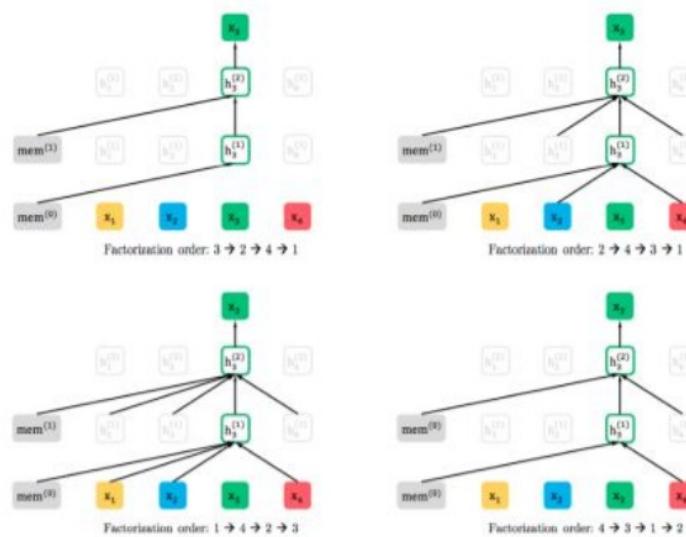


Figure A.2: XLNet Permutation processing

26

state-of-the-art results in almost all the language modelling benchmarks and has also out-ranked BERT in around 20 tasks. Auto-regressive models like XLNet are very good generative NLP tasks because it only involves processing of text in one direction only. The only disadvantage of auto-regressive models like XLNet is that they can only determine the context in one direction only at a time i.e. either forward or backward direction.

A.1.4 Longformer

One major drawback of BERT based models is that they can't "attend" to longer sequences. For instance, BERT is restricted to a max of 512 tokens at a time. To beat these long sequence issues, many novel methods and approaches have been used. Models like Transformer-XL and Reformer propose novel ways to scale back the model parameters, and hence, the complexity. The Longformer essentially combines several attention and self-attention patterns to beat the above drawbacks:

1. Sliding Window
2. Dilated window
3. Global Attention (full self-attention)

In the sliding window approach, we take an arbitrary window size x , and every token within the sequence will only provide attention to some w tokens (mostly $x/2$ to the left and $x/2$ to the right). To know the working of this attention pattern, let's consider the instance of convolutions. Say we take a kernel of size x and slide it through all the tokens within the sequence. After this process, we'll have the hidden state embeddings of all the tokens within the sequence when attended with w adjacent tokens.

The main idea of the Transformer was the parallelization of eye computation over the sequential nature of the RNNs. The window attention takes away the entire point of using transformer-based architectures. To feature parallelization, one can implement an enormous matrix with all zeros apart from the diagonals (as shown within the attention figures), indicating windowed Attention.

To answer this, the authors have implemented a custom CUDA kernel for the proposed attention pattern. The authors have written a sort of banded matrix operation in python that the TVM generates the required CUDA code and compiles it for the precise GPU.
The parallelization ⁶ is achieved by making use of GPU threads, and only the non-zero values of the large matrix (i.e., just the window values) are stored.

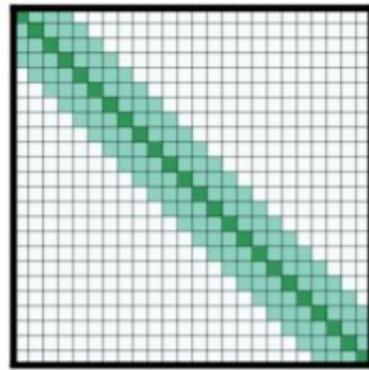


Figure A.3: Longformer sliding window

A.1.5 Distil-BERT

DistilBERT is a distilled or an approximate version of BERT, is another pre-trainer transformer based transfer learning model just like BERT and RoBERT-a. DistilBERT is a very optimal and a light-weight transformer transfer learning model which achieves 95% performance of BERT but makes use of half the parameters as BERT i.e., close to 60 million parameters. DistilBERT also keeps only half of layers in its architecture as that of Google's BERT. DistilBERT uses a way called distillation, which approximates Google's BERT, i.e., the massive neural network by a smaller one. DistilBert achieves the process of distillation or approximation through the process of knowledge distillation which is a form of compression techniques to convert a big model into a small one.

Mental health analysis during COVID

ORIGINALITY REPORT

% 6	% 5	% 2	% 4
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

- | | | |
|---|-------------------------------------------------------------------------------------------------------------------|------|
| 1 | Submitted to Netaji Subhas Institute of Technology | % 1 |
| 2 | www.analyticsvidhya.com
Internet Source | % 1 |
| 3 | github.com
Internet Source | % 1 |
| 4 | fr.scribd.com
Internet Source | % 1 |
| 5 | www.kdnuggets.com
Internet Source | <% 1 |
| 6 | "Natural Language Processing and Chinese Computing", Springer Science and Business Media LLC, 2019
Publication | <% 1 |
| 7 | studentsrepo.um.edu.my
Internet Source | <% 1 |
| 8 | Submitted to East Carolina University
Student Paper | <% 1 |

9	docplayer.net Internet Source	<% 1
10	divostar.com Internet Source	<% 1
11	Arianna Maever L. Amit, Veincent Christian F. Pepito, Bernardo Gutierrez, Thomas Rawson. "Tracking changes in reporting of epidemiological data during the COVID-19 pandemic in Southeast Asia: an observational study during the first wave", Cold Spring Harbor Laboratory, 2020 Publication	<% 1
12	Submitted to University of Sussex Student Paper	<% 1
13	www.reddit.com Internet Source	<% 1
14	Submitted to Macao Polytechnic Institute Student Paper	<% 1
15	Submitted to University of Hull Student Paper	<% 1
16	algorithmxlab.com Internet Source	<% 1
17	www.kdd.org Internet Source	<% 1

18	scholarspace.manoa.hawaii.edu	<% 1
Internet Source		
19	www.aclweb.org	<% 1
Internet Source		
20	repository.upi.edu	<% 1
Internet Source		
21	Submitted to Liverpool John Moores University	<% 1
Student Paper		
22	ir.lib.uwo.ca	<% 1
Internet Source		
23	scholarworks.iupui.edu	<% 1
Internet Source		
24	"Advanced Data Mining and Applications", Springer Science and Business Media LLC, 2017	<% 1
Publication		
25	"Neural Information Processing", Springer Science and Business Media LLC, 2017	<% 1
Publication		
26	"Neural Information Processing", Springer Science and Business Media LLC, 2019	<% 1
Publication		
27	medium.com	<% 1
Internet Source		

28

Sepp Hochreiter, Jürgen Schmidhuber. "Long
Short-Term Memory", Neural Computation,
1997

<% 1

Publication

EXCLUDE QUOTES ON
EXCLUDE ON
BIBLIOGRAPHY

EXCLUDE MATCHES OFF