

Text Analytics Of Peer Group - WhatsApp Communications For Mental Health Analysis Of Peers

Anubhav Behera^{*}, Dibakar Upadhyaya^{**}, Manjusha Pandey^{***}, Siddharth Swarup Rautaray^{****}

^{*} School Of Computer Engineering, KIIT University, 2005011@kiit.ac.in

^{**} School Of Computer Engineering, KIIT University, 2005798@kiit.ac.in

^{***} School Of Computer Engineering, KIIT University, manjushafcs@kiit.ac.in

^{****} School Of Computer Engineering, KIIT University, siddharthfcs@kiit.ac.in

Abstract- Mental health encompasses emotional, psychological, and social well-being. It influences cognition, perception, and behaviour. It also determines how a person handles stress, interpersonal relationships, and decision-making. When a number of people are communicating with each other through different forms of channels it is known as group communication. The members of this group are likely to influence the individual's beliefs and behaviour. WhatsApp Messenger is a cross-platform instant messaging application that allows smartphone users to exchange text, image, video and audio messages for free. It also allows users to communicate in peer-to-peer or form groups where information can be exchanged via media and text. As WhatsApp groups may contain individuals from different backgrounds, fields of interest and age groups as well there is an exchange of a variety of opinions. Humans being social animals do get influenced by this kind of communication to fit in or agree to the flow of discussion. Hence it does create an impact on peers' mental health. Text analytics is the process of translating large volumes of text into quantitative data to uncover insights, trends, and patterns. This project would be utilising text analytics techniques to analyze the WhatsApp group communications to peer mental health.

Index Terms- COVID-19, Mental Health, Text Analytics, WhatsApp Communications

I. INTRODUCTION

The COVID-19 pandemic brought us close yet miles apart. Everyone's life was drastically affected during this monstrosity with a rapid phase change to the lockdown period where people were not only affected by the virus but also by the aftermath of it [3]. This included the whole world experiencing numerous emotions such as fear of the virus, sudden cut-off from the physical world, loneliness from rare physical human interaction etc [4]. As we know, Mental Health determines one's psychological, emotional and social well-being. During these times people's mental health was challenged which caused them to shift to online means of relationships and exchange of information, particularly text-based chatting platforms like WhatsApp [6].

WhatsApp Messenger is an instant messaging and multimedia-sharing application with around 2 billion users worldwide that allows smartphones and personal computer users to exchange text, image, video and audio messages for free with new additions to its application, as we read. It saw an immense increase in usage during the recent pandemic. Text-based group conversations became the common ground for discussion and sharing information with the introduction of this application, creating a virtual interactive environment.

These text-based groups consist of people from various backgrounds with unique opinions discussing a plethora of topics and contributing their perspectives since these conversations became a part of daily life and even more so during these difficult times it confirmed the possibility of individuals getting mentally and physically affected by the same [11].

WhatsApp generates a humongous amount of text data every single minute and we are focusing on making use of the same. This paper will demonstrate how WhatsApp group chat data can be analyzed by utilizing text-analytics techniques to determine the effect of these group communications on peer mental health. The proposed model would be based on the 'Bag of Words' model.

II. STATEMENT OF THE PROBLEM AND OBJECTIVES

The sudden cut-off from the whole world due to the COVID-19 pandemic led individuals to start to resort to digital platforms to support the lack of physical communication. This encapsulates various people from different racial and financial backgrounds, ethnicity, age, languages etc. Consequently leading to a humongous amount of information being exchanged, and perspectives being put forward which in turn affects the mental health of the people participating in the same. In this paper, we are analysing the text data taken from a peer-group chat, in a chat platform (WhatsApp) and flagging the peer group on the basis of its positive, neutral or negative effects on any individual's mental health who is part of these groups.

The dataset that we have used contains a Formula One Racing (a.k.a. F1 Racing) based peer-group chat which has 254 participants from different backgrounds, ages ranging from 17 - 43 years old belonging to the aforementioned countries of England, India, USA, Morocco, Mauritius, Kenya, Zimbabwe, Netherlands, Belgium, France, Spain, Ireland, Italy, Switzerland, Germany, Mexico, Brazil, Australia and Portugal. The group was created on 28th August 2021 for which we have taken the data from the months of August, September and October of the year 2022 while implementing our model. The participants of this group are from different cultural backgrounds, and hence they give an insight into the different opinions of various people about the given topic. They also have different ages and gender which can affect their attitude towards a particular situation and/or have a unique opinion about the same giving us a wide variety of conversations to analyze and conclude our results. The chat is derived from WhatsApp, as it is the most used chat application in the world currently.

We start with converting the text chat into a CSV file. The words from the texts are then pre-processed i.e. they are tokenized as individual words, furthermore, they are converted into lower-case letters with the removal of punctuation and stop words along with performing word stemming on the various strings generated. The frequency of these occurring words is then found out from which a document-term sparse matrix is made. A bar graph is then plotted showcasing the frequency of the top most used words with the previously attained matrix.

We disambiguate the words based on the probability that a word occurs with a particular tag, i.e., we need to mark up the words in text format for particular parts of the message based on their definition and context. We then utilize the process of chunking in making groups of noun phrases. The resulting group of words are called chunks.

The possible tags are -

FW - (foreign word), JJR - (adjective, comparative), JJS - (adjective, superlative), LS - (list marker), MD - (modal), NN - (noun, singular), NNS - (noun, plural), NNP - (proper noun, singular), NNPS - (proper noun, plural), PDT - (predeterminer), POS - (possessive ending), PRP - (personal pronoun), PRP\$ - (possessive pronoun), RB - (adverb), RBS - (adverb, superlative), RP - (particle), TO - (infinite marker), UH - (interjection), VB - (verb).

We will consider positive words to be the ones which are encouraging or uplifting in nature. Neutral words are ones which come under the tag such as NN, NNS, NNP or NNPS, meaning they exist to give meaning to the sentence and are not targeted to portray a positive or negative image. Negative words are the ones which are abusive in nature such as cuss words.

From the chunks formed we determine the nature of the words which have occurred most frequently in the chat. We see the nature of these words and based on this observation we flag the group into one of the three possible categories - positive, neutral or negative.

III. BACKGROUND STUDY AND TECHNOLOGY GAPS IDENTIFIED

Year/Author	Title	Proposed	Advantage	Disadvantage
[1] [12] [13] 2022/Taquet, Maxime et al.	Neurological and psychiatric risk trajectories after SARS-CoV-2 infection: an analysis of 2-year retrospective cohort studies including 1 284 437 patients	To find out how COVID-19 is related to the increased risks of neurological and psychiatric sequela.	Discovered that there is an increased risk of a range of neurological and psychiatric diagnoses in the first 6 months after COVID-19 diagnosis. They found drastic differences in the results of these risks within the first 2 years after diagnosis. The results were much different in children compared with adults and older ages differing between variants of SARS-CoV-2.	COVID-19 appeared to be more severe than other respiratory infections, and the severity of the illness was not taken into consideration during the analysis. Only individuals who were diagnosed early in the pandemic contributed to this dataset for the whole 2-year follow-up.

[2] 2022/Sophie Lythreatis et al.	The digital divide: A review and future research agenda	A systematic review of the prevailing digital divide i.e, unequal access to digital technology including two new potential forms apart from 'type-of-internet access' are 'algorithmic awareness' and 'data inequalities'.	Structurally analyzed and critically reviewed the previous studies regarding the digital divide reviewing what determines while discovering new levels and forms of the existing concepts, which is going to help organizations, governments and policymakers to assess, adjust and develop ways how technology creates an inclusive prosperous and sustainable society.	<p>The coding for the results was done by hand which could include subjectivity.</p> <p>The results of 2021 are not complete as all the articles for the same were not completed at the time this paper was written.</p> <p>Their methodological approach was eliminating certain articles; also only the articles in the English language were included in the same.</p>
[3][14][15][16] 2022/Ali Cheshmehzangi et al.	The digital divide impacts on mental health during the COVID-19 pandemic	A study on how the COVID-19 pandemic brought an intensifying scale change in the digital divide (DD) whilst affecting people's mental health unnaturally.	Revealed that the impact of the COVID-19 pandemic, digital exclusion and social exclusion driven by DD had worsened people's mental health and people are prone to depression and loneliness as a consequence.	<p>They did not shed light on the study that should overlap the factors between DD-prone people and/or social exclusion.</p> <p>The impact of DD on mental health was not backed up as expected.</p>
[4][17][18][19] 2020/Nina Vindegaard, Michael Eriksen Benros	COVID-19 pandemic and mental health consequences: Systematic review of the current evidence	A systematic review of the literature to provide an overview of the psychiatric complications of COVID-19 infection and how	<p>Revealed that there are indications of PTSD and depression following the COVID-19 pandemic.</p> <p>Evidence backs up an increase in</p>	<p>Restrictions were recognized in the dataset as the studies were conducted mostly in Asia.</p> <p>The studies were not case-control</p>

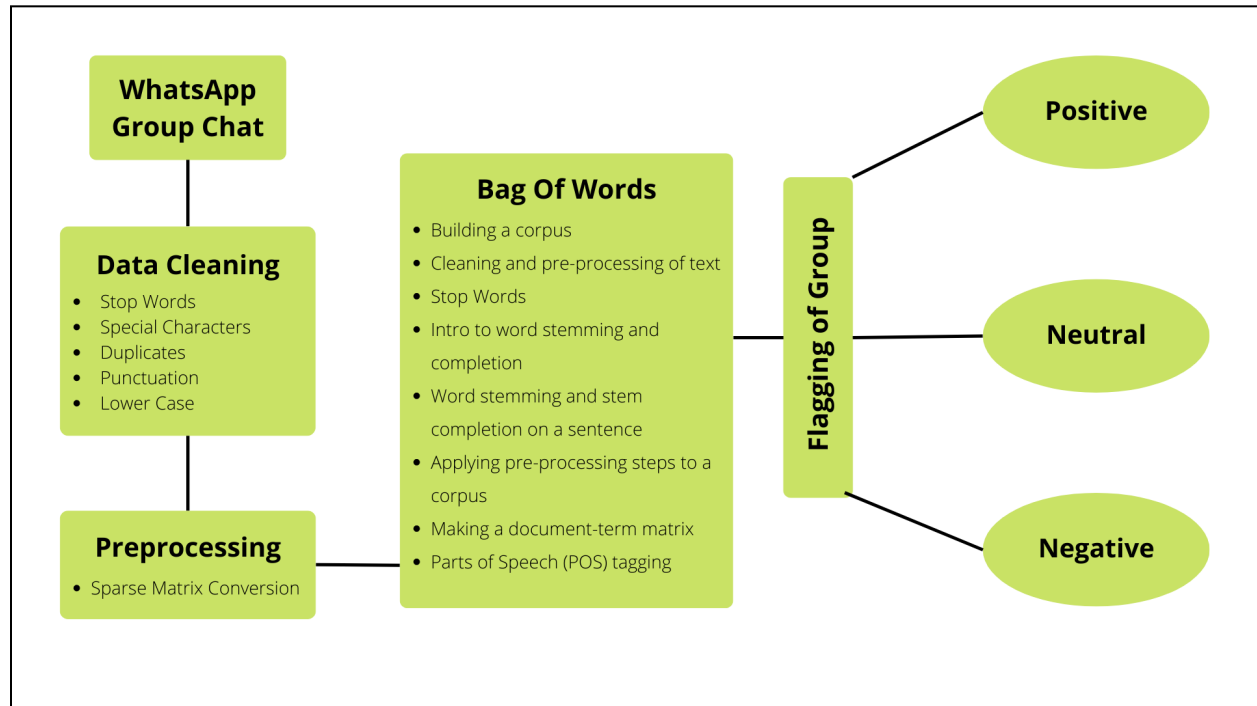
		it is, directly or indirectly, affecting the mental health of psychiatric patients and the general public.	depressive and anxiety symptoms and adverse effects on general mental health.	<p>studies and with a variety of reported outcomes, measuring of outcomes and statistical analysis were revealed.</p> <p>The literature on psychiatric symptoms among patients with SARS-CoV-2 was limited.</p>
[5][20][21] 2020/Shen Zhu et al.	The immediate mental health impacts of the COVID-19 pandemic among people with or without quarantine managements	To test mental health problems using early survey data with a mobile app “Sojump” to determine how it was related to COVID-19 quarantine.	Results showed a relatively high prevalence of mental health problems in the sample dataset, which were closely related to the impacts on daily life that were created and not the control measure of quarantine itself.	The study couldn’t focus on the potential psychological risk factors associated with mental health problems under quarantine limited by their early data.
[6] 2016/J A Naslund et al.	The future of mental health care: peer-to-peer support and social media.	Offering a perspective on how online peer-to-peer connections among people with serious mental illness could advance efforts to promote mental and physical well-being in this group.	<p>The results revealed that people with serious mental illness did benefit from interacting with peers online from greater social connectedness, and feelings of group belonging and sharing their stories helped them cope with the daily challenges of living with their condition.</p> <p>Individuals also gained insights on important healthcare decisions which promoted healthcare-seeking behaviours for their mental illness</p>	<p>With the benefit comes the risk of getting exposed to misleading information, facing hostile or derogatory comments, or feeling uncertain about one’s health condition. These are the potential risks that tag along.</p> <p>They also need to determine that the skills learned through this method of peer-to-peer support bring an impact on their</p>

			overall.	offline lives and overall physical and mental well-being.
[7][22][23] 2020/Anna Lavis, Rachel Winter	#Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media.	To know the potential detrimental impact of social media on young people's mental health, with a focus on online harm and getting to know the reason behind the same.	The results of the data shed light on peer support being the central component of online interactions around self-harm.	The authors have not taken physical communication as a factor affecting mental health and were only limited to digital platforms.
[8][24][25] 2022/Conor J. Wild et al.	Disentangling the cognitive, physical, and mental health sequelae of COVID-19	Study on the suggestion of the fact that the subjective experience of "long COVID" related to physical symptoms and cognitive deficits, especially executive dysfunction.	<p>The results supported the existence of cognitive impairments in the aftermath of COVID-19 disease and include other important findings.</p> <p>Emphasized the fact that the physical, emotional, mental and cognitive sequelae of COVID-19 were not bound together as a single neurocognitive syndrome and that executive function and verbal abilities were the key domains that were affected in COVID-19 survivors.</p>	<p>There was selection bias in the sample dataset is drawn by them.</p> <p>Several other outcomes of COVID-19 were not measured such as PTSD which might have been interrelated with the sets of symptoms discussed in their study.</p>
[9][26][27] 2022/Stephanie Scott et al.	Australian Psychologists Experiences with Digital Mental Health: a Qualitative Investigation	This study explores psychologists' attitudes and experience with digital mental health intervention by taking a	They discovered that their study was fruitful for the future design of digital mental health resources and service delivery	They focused mainly on the larger cities for the samples of the psychologists whereas the experiences of rural healthcare professionals

		qualitative exploratory approach with the help of thematic analysis.		<p>should have also been included.</p> <p>The sample of 10 psychologists acquired by them did not represent the broader population.</p>
[10][28][29] 2022/Jonathan Campion, FRCPsych et al.	Public mental health: required actions to address implementation failure in the context of COVID-19	It outlines the specific actions to improve the coverage of Public Mental Health (PMH) which had an increase in interest due to the onset of the COVID-19 pandemic.	<p>They confirmed that evidence-based, cost-effective PMH interventions exist to prevent mental disorders from arising, which have broad health, social, and economic impacts that also support the achievement of a range of policy objectives, failing to achieve the same breaches results in the breaching of the right to health leading to large-scale associated impacts and economic costs.</p> <p>These actions are also being found to be more important due to the COVID-19 pandemic consequently piquing the interest of the government and people in the mental health of the population.</p>	They broadly focused on the high-income countries for this study and the low-income and medium-income countries were not taken into consideration, which breaches the idea of a general opinion of public mental health showcasing the prevalence of selection bias.
[11][30][31] 2020/Shweta Singh, Deblina Roy et al.	Impact of COVID-19 and lockdown on mental health of children and adolescents: A	It aims at reviewing multiple articles related to mental-health aspects of	The infection of COVID-19 among young children and adolescents is low however the stress confronted by the	The review articles were taken during the time of global lockdown when the issues and

	<p>narrative review with recommendations</p>	<p>children and adolescents impacted by the COVID-19 pandemic and enforcement of nationwide or regional lockdowns to prevent further spread of infection.</p>	<p>same is considered to be highly vulnerable.</p> <p>The study revealed that the pandemic and its preventions which led to lockdown and quarantine affected children by them having disturbed sleep, clinginess, poor appetite, inattentiveness and significant separation problems.</p> <p>The containment measures not only showed a rise in loneliness, anxiety and stress among the children and youth but also the economically underprivileged children were prone to exploitation and abuse.</p>	<p>challenges were new with the global crisis at its peak.</p> <p>They were unable to track the measures of the management targeted towards the children.</p> <p>The strategies reported were isolated to geopolitical conditions.</p> <p>There was also selection bias which persisted along with a short period of data collection restricted by only the use of electronic databases for their research.</p>
--	--	---	--	---

IV. PROPOSED MODEL / TOOL



The above flowchart represents the proposed model for the process of implementation of all our objectives which we have defined in section II of this paper.

Python - Python is a high-level programming language which is widely used for its easy-to-read syntax. It also has multiple libraries which cater to different fields hence is a general-purpose language with its aforementioned statistical and data analysis libraries and modules it is the perfect choice for our paper.

Collections - Collections in Python are containers used for storing data and are commonly known as data structures which include tuples, lists, dictionaries, etc.

Pandas - Pandas is a software library written for the Python programming language for data manipulation and analysis.

Matplotlib - It is a data visualization library in Python and its numerical extension, NumPy. It helps in plotting graphs and charts, across various platforms.

Seaborn - The Seaborn package in Python allows the creation of statistical visuals. It consolidates closely with Pandas data structures and is built upon the Matplotlib library.

Natural Language Toolkit (NLTK) - It is a platform to develop Python programs that incorporate data collected from human language for statistical natural language processing (NLP). It contains text-processing libraries for parsing, tokenization, stemming, classification, tagging, and semantic reasoning.

Scikit-learn (Sklern) - It provides a diverse range of effective methods in Python for machine learning and statistical modelling, including classification, clustering, regression, and dimension reduction.

V. IMPLEMENTATION AND RESULTS

PYTHON

We have implemented the entire code in Python with the help of Jupyter Notebook. Various libraries which have been mentioned above have been used to implement the Bag of Words model on our dataset to acquire the results.

PANDAS

We utilize Pandas to display the data in a tabular format.

	Date	timestamp	Person	title	day	hour	dayofweek
0	8/6/22	2022-11-08 18:59:00	A	Piastri has dominated everything he has done e...	2022-11-08	18	Tuesday
1	8/6/22	2022-11-08 18:59:00	A	Wc in all categories	2022-11-08	18	Tuesday
2	8/6/22	2022-11-08 18:59:00	A	Would definitely learn this new car faster	2022-11-08	18	Tuesday
3	8/6/22	2022-11-08 21:03:00	A	W11 was a beast	2022-11-08	21	Tuesday
4	8/6/22	2022-11-08 21:03:00	A	Rocket on rails	2022-11-08	21	Tuesday
5	8/7/22	2022-11-08 19:40:00	A	Will never happen sadly	2022-11-08	19	Tuesday
6	8/8/22	2022-11-08 15:32:00	A	In all this chaos u forget u are still aiming ...	2022-11-08	15	Tuesday
7	8/8/22	2022-11-08 15:33:00	A	They should announce it already.	2022-11-08	15	Tuesday
8	8/9/22	2022-11-08 05:08:00	A	What I meant is. He's already competing at the...	2022-11-08	5	Tuesday
9	8/11/22	2022-11-08 17:39:00	A	How long is this summer break	2022-11-08	17	Tuesday
10	8/11/22	2022-11-08 23:12:00	A	Kimi	2022-11-08	23	Tuesday
11	8/11/22	2022-11-08 23:26:00	A	https://nypost.com/2022/08/09/ferrari-recalls-...	2022-11-08	23	Tuesday
12	8/12/22	2022-11-08 21:13:00	A	Why r fans even using Twitter.	2022-11-08	21	Tuesday
13	8/12/22	2022-11-08 21:13:00	A	Nothing good ever comes from that platform	2022-11-08	21	Tuesday
14	8/12/22	2022-11-08 21:14:00	A	Even Chris medland posts are put here by prana...	2022-11-08	21	Tuesday
15	8/12/22	2022-11-08 21:14:00	A	Id prefer such groups and Instagram anyday.	2022-11-08	21	Tuesday
16	8/12/22	2022-11-08 21:15:00	A	Had to be either a mercedes or a rb fan	2022-11-08	21	Tuesday
17	8/13/22	2022-11-08 15:10:00	A	no wonder elon is hesitating to buy	2022-11-08	15	Tuesday
18	8/13/22	2022-11-08 17:17:00	A	Hope no rain this year.. should be a good race	2022-11-08	17	Tuesday
19	8/16/22	2022-11-08 18:18:00	A	Is he talking about himself	2022-11-08	18	Tuesday

NLTK

We utilize this library to convert all strings into lower-case, remove all punctuations, tokenize strings into individual words, remove stop words, and perform word stemming on the dataset.

COLLECTIONS

This Python module was used to calculate the frequency of individual words appearing in the corpus.

```
Counter({'piastri': 5,
        'dominated': 2,
        'everything': 11,
        'done': 3,
        'earlier': 1,
        'wc': 1,
        'category': 1,
        'would': 31,
        'definitely': 2,
        'learn': 1,
        'new': 18,
        'car': 70,
        'faster': 6,
        'beast': 1,
        'rocket': 2,
        'rail': 1,
        'never': 18,
        'happen': 5,
        'sadly': 2,
        'chaos': 2,
        'u': 33,
        'forget': 7,
        'still': 26,
        'aiming': 1,
        'drive': 5,
        'pinnacle': 1,
        'sport': 3,
        'maybe': 12,
        'guy': 13,
        'know': 25,
```

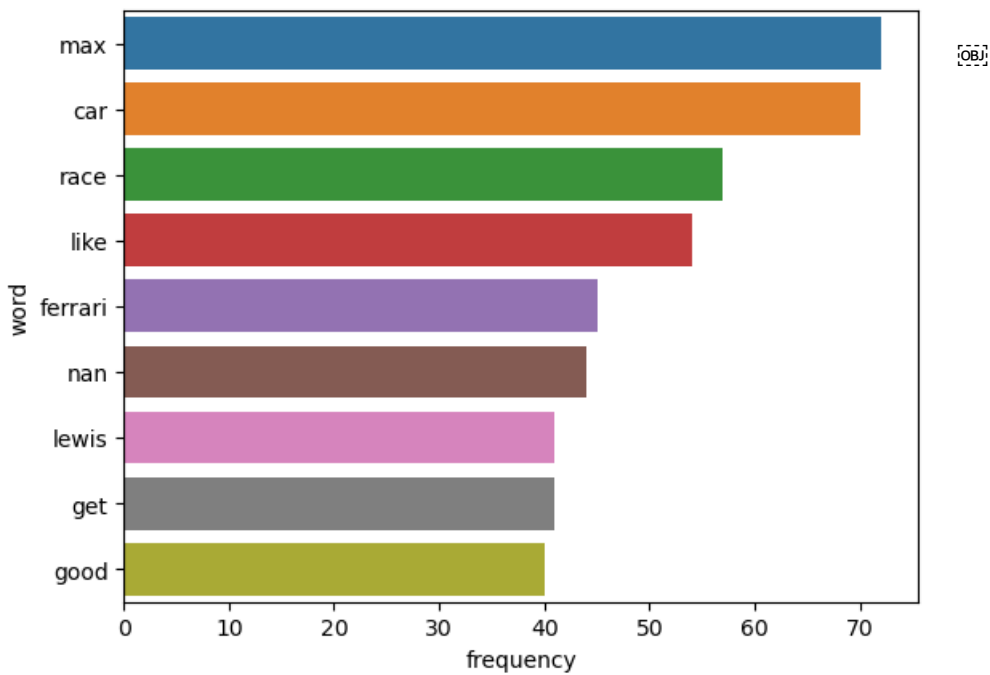
SCIKIT-LEARN

We used this tool to convert individual strings into vectors along with Pandas library we utilize them together to obtain the document-term sparse matrix.

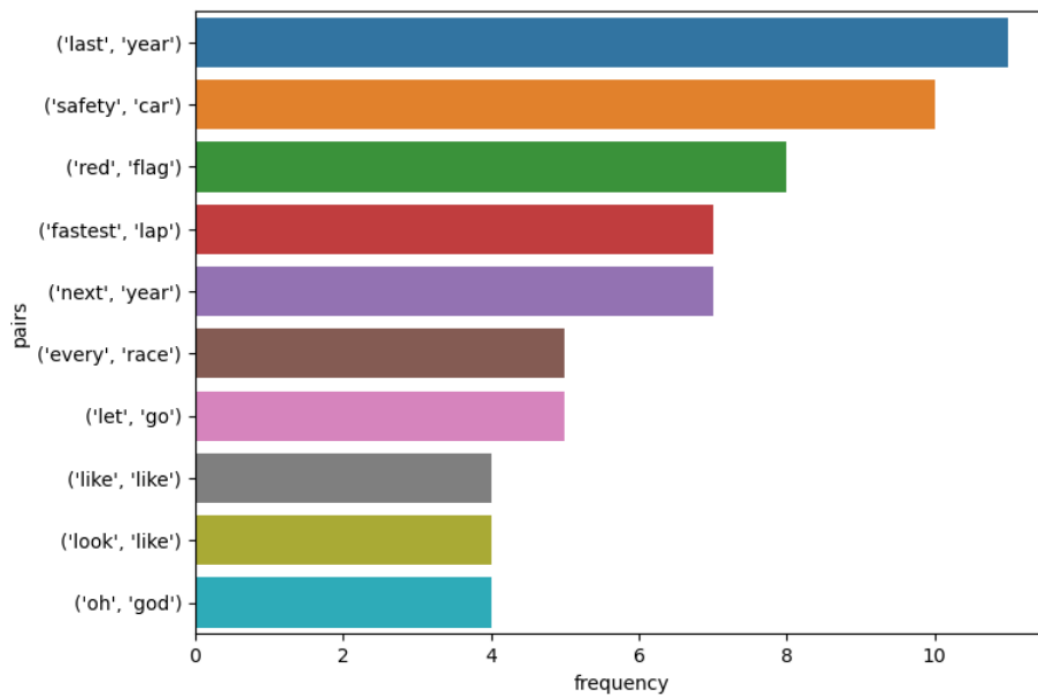
	08	09	17	2022	ago	aiming	amp	announce	anyday	anymore	...	wc	weekend	winning	wins	won	wonder	worth
Piastrì has dominated everything he has done earlier. From F4 to f2	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
Wc in all categories	0	0	0	0	0	0	0	0	0	0	0 ...	1	0	0	0	0	0	0
Would definitely learn this new car faster	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
W11 was a beast	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
Rocket on rails	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
Will never happen sadly	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
In all this chaos u forget u are still aiming to drive at the pinnacle of the sport. Maybe these guys know that they could get a seat in other formats of racing. Also he's loaded already. It's not an L , just disappointment ig	0	0	0	0	0	1	0	0	0	0	0 ...	0	0	0	0	0	0	0
They should announce it already.	0	0	0	0	0	0	0	1	0	0	0 ...	0	0	0	0	0	0	0
What I meant is. He's already competing at the highest level in racing. But fighting for more/wins. Didn't say it's disappointing. He can always do something else if there is situation where no one signs him. For people like us , feeling bad for daniel not being in formula 1 is stupid.. he is in this position due to his own choices. It's disappointing he's not at the top anymore winning like before.	0	0	0	0	0	0	0	0	0	1	0 ...	0	0	1	1	0	0	0
How long is this summer break	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
Kimi	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0
https://nypost.com/2022/08/09/ferrari-recalls-17-years-worth-of-cars-over-possible-brake-failure/ amp/	1	1	1	1	0	0	1	0	0	0	0 ...	0	0	0	0	0	0	1
Why r fans even using Twitter.	0	0	0	0	0	0	0	0	0	0	0 ...	0	0	0	0	0	0	0

SEABORN AND MATPLOTLIB

We utilize these libraries to plot a word-frequency bar graph.



Below is the bar graph which was made using ngrams for pairs of words which occur the most frequently.



We utilize the Parts of Speech (POS) tagging technique and tag the various chunks which are formed into their individual categories.

After Split: ['Piastrì has dominated everything he has done earlier. From F4 to f2', 'Wc in all categories', 'Would definitely learn this new car faster', 'Will was a beast', 'Rocket on rails', 'Will never happen sadly', 'In all this chaos u forget u are still aiming to drive at the pinnacle of the sport. Maybe these guys know that they could get a seat in other formats of racing. Also he's loaded already. It's not an L , just disappointment ig', 'They should announce it already. ', 'What I meant is. He's already competing at the highest level in racing. But fighting for more/wins. Didn't say it's disappointing. He can always do something else if there is situation where no one signs him. For people like us , feeling bad for daniel not being in formula 1 is stupid.. he is in this position due to his own choices. It's disappointing he's not at the top anymore winning like before.', 'How long is this summer break', 'Kimi', 'https://nypost.com/2022/08/09/ferrari-recalls-17-years-worth-of-cars-over-possible-brake-failure/amp/', 'Why r fans even using Twitter.', 'Nothing good ever comes from that platform ', 'Even Chris medland posts are put here by pranav and others.', 'Id prefer such groups and Instagram anyway.', 'Had to be either a mercedes or a rb fan', 'no wonder elon is hesitating to buy ', 'Hope no rain this year.. should be a good race', 'Is he talking about himself ', 'Maybe one day when lewis retires they could be better friends ', 'Mercedes have won for so long they got very cocky very soon.', 'It's time Ferrari and rb come back up again.', 'Does max prefer oversteer ? Heard this somewhere long ago', 'I mean a little bit..', 'Since last season his cars rears tend to slide around , visually atleast..', 'Interesting.', 'At first I thought he doesn't have the right balance but he keeps winning with this setup', 'Maybe even gets more temp on tyres that way considering the sliding around', 'Let's gooo', 'Atleast now they must announce about piastrì', 'Raining this weekend at spa wtf ?', 'Both days ', '']

After Token: [('Piastrì has dominated everything he has done earlier. From F4 to f2', 'NNP'), ('Wc in all categories', 'NNP'), ('Would definitely learn this new car faster', 'NNP'), ('Will was a beast', 'NNP'), ('Rocket on rails', 'NNP'), ('Will never happen sadly', 'NNP'), ('In all this chaos u forget u are still aiming to drive at the pinnacle of the sport. Maybe these guys know that they could get a seat in other formats of racing. Also he's loaded already. It's not an L , just disappointment ig', 'NNP'), ('They should announce it already. ', 'NNP'), ('What I meant is. He's already competing at the highest level in racing. But fighting for more/wins. Didn't say it's disappointing. He can always do something else if there is situation where no one signs him. For people like us , feeling bad for daniel not being in formula 1 is stupid.. he is in this position due to his own choices. It's disappointing he's not at the top anymore winning like before.', 'NNP'), ('How long is this summer break', 'NNP'), ('Kimi', 'NNP'), ('https://nypost.com/2022/08/09/ferrari-recalls-17-years-worth-of-cars-over-possible-brake-failure/amp/', 'JJ'), ('Why r fans even using Twitter.', 'NNP'), ('Nothing good ever comes from that platform ', 'NNP'), ('Even Chris medland posts are put here by pranav and others.', 'NNP'), ('Id prefer such groups and Instagram anyway.', 'NNP'), ('Had to be either a mercedes or a rb fan', 'NNP'), ('no wonder elon is hesitating to buy ', 'JJ'), ('Hope no rain this year.. should be a good race', 'NNP'), ('Is he talking about himself ', 'NNP'), ('Maybe one day when lewis retires they could be better friends ', 'NNP'), ('Mercedes have won for so long they got very cocky very soon.', 'NNP'), ('It's time Ferrari and rb come back up again.', 'NNP'), ('Does max prefer oversteer ? Heard this somewhere long ago', 'NNP'), ('I mean a little bit..', 'NNP'), ('Since last season his cars rears tend to slide around , visually atleast..', 'NNP'), ('Interesting.', 'NNP'), ('At first I thought he doesn't have the right balance but he keeps winning with this setup', 'NNP'), ('Maybe even gets more temp on tyres that way considering the sliding around', 'NNP'), ('Let's gooo', 'NNP'), ('Atleast now they must announce about piastrì', 'NNP'), ('Raining this weekend at spa wtf ?', 'NNP'), ('Both days ', 'NNP'), ('', 'NN')]

After Chunking (S
(mychunk
Piastrì has dominated everything he has done earlier. From F4 to f2/NNP
Wc in all categories/NNP
Would definitely learn this new car faster/NNP
Will was a beast/NNP
Rocket on rails/NNP
Will never happen sadly/NNP
In all this chaos u forget u are still aiming to drive at the pinnacle of the sport. Maybe these guys know that they could get a seat in other formats of racing. Also he's loaded already. It's not an L , just disappointment ig/NNP
They should announce it already. /NNP
What I meant is. He's already competing at the highest level in racing. But fighting for more/wins. Didn't say it's disappointing. He can always do something else if there is situation where no one signs him. For people like us , feeling bad for daniel not being in formula 1 is stupid.. he is in this position due to his own choices. It's disappointing he's not at the top anymore winning like before./NNP
How long is this summer break/NNP
Kimi/NNP
https://nypost.com/2022/08/09/ferrari-recalls-17-years-worth-of-cars-over-possible-brake-failure/amp//JJ)
(mychunk
Why r fans even using Twitter./NNP
Nothing good ever comes from that platform /NNP
Even Chris medland posts are put here by pranav and others./NNP
Id prefer such groups and Instagram anyway./NNP
Had to be either a mercedes or a rb fan/NNP
no wonder elon is hesitating to buy /JJ)
(mychunk
Hope no rain this year.. should be a good race/NNP
Is he talking about himself /NNP
Maybe one day when lewis retires they could be better friends /NNP
Mercedes have won for so long they got very cocky very soon./NNP
It's time Ferrari and rb come back up again./NNP
Does max prefer oversteer ? Heard this somewhere long ago/NNP
I mean a little bit../NNP
Since last season his cars rears tend to slide around , visually atleast../NNP
Interesting./NNP
At first I thought he doesn't have the right balance but he keeps winning with this setup/NNP
Maybe even gets more temp on tyres that way considering the sliding around/NNP
Let's gooo/NNP
Atleast now they must announce about piastrì/NNP
Raining this weekend at spa wtf ?/NNP
Both days /NNP
(NN))

As we can see from the results most of the chunks are categorized in either the NNP category or the NN category in the dataset meaning the words with the highest frequency are either Nouns or Pronouns pointing to an object or a person. Since the words with the maximum frequencies are neither positive nor negative, hence we flag the group to be neutral.

VI. CONCLUSION

A lot of research has been done on text analysis for the identification of emotions and predicting the choice of people based on the text generated by them. Mostly this emotional analysis has been done for products, services and political implications which is thoroughly profit-oriented and has been capitalized by commercial giants for their benefits. None of the studies to date has considered the impact of this emotional analysis on human beings and specifically students so that the mental health and physical well-being of the current and next generation could be enhanced. Hence, this project of ours concentrates on this loop of the virtual world and its impact on students, and their physical and mental health specifically during the times of pandemic when they were compelled to be part of this virtual world without any choice.

The presented study could be enhanced and could be utilized further for making the virtual interactions flagged with their mental health implications as to be positive implications, negative implications or neutral implications. The same could be utilized for preparedness for any other natural calamity like the COVID-19 pandemic to be managed in a way which will have little or no impact on the mental and physical well-being of the current and next generations.

This paper also comes with its limitations. Firstly, the dataset taken was a Formula One-related peer group which might arise a selection bias. Secondly, emojis, images and videos were not taken into consideration hence we won't get the full picture of what was occurring in the group. Thirdly, abbreviations were kept intact which might have affected the maximum frequency of certain words. These are a few limitations which we hope to be worked on as future studies and improvements.

REFERENCES

- [1] Taquet, M., Sillett, R., Zhu, L., Mendel, J., Camplisson, I., Dercon, Q., & Harrison, P. J. (2022). Neurological and psychiatric risk trajectories after SARS-CoV-2 infection: an analysis of 2-year retrospective cohort studies including 1 284 437 patients. *The Lancet Psychiatry*, 0(0). [https://doi.org/10.1016/S2215-0366\(22\)00260-7](https://doi.org/10.1016/S2215-0366(22)00260-7)
- [2] Lythreathis, S., El-Kassar, A.-N., & Singh, S. K. (2021). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, 175, 121359. <https://doi.org/10.1016/j.techfore.2021.121359>
- [3] Cheshmehzangi, A., Zou, T., & Su, Z. (2022). The digital divide impacts on mental health during the COVID-19 pandemic. *Brain, Behavior, and Immunity*, 101, 211–213. <https://doi.org/10.1016/j.bbi.2022.01.009>
- [4] Vindegaard, N., & Eriksen Benros, M. (2020). COVID-19 pandemic and mental health consequences:

- Systematic review of the current evidence. *Brain, Behavior, and Immunity*, 89(4).
<https://doi.org/10.1016/j.bbi.2020.05.048>
- [5] Zhu, S., Wu, Y., Zhu, C., Hong, W., Yu, Z., Chen, Z., Chen, Z., Jiang, D., & Wang, Y. (2020). The immediate mental health impacts of the COVID-19 pandemic among people with or without quarantine managements. *Brain, Behavior, and Immunity*, 87, 56–58.
<https://doi.org/10.1016/j.bbi.2020.04.045>
- [6] Naslund, J. A., Aschbrenner, K. A., Marsch, L. A., & Bartels, S. J. (2016). The Future of Mental Health care: peer-to-peer Support and Social Media. *Epidemiology and Psychiatric Sciences*, 25(02), 113–122. <https://doi.org/10.1017/s2045796015001067>
- [7] Lavis, A., & Winter, R. (2020). #Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 61(8). <https://doi.org/10.1111/jcpp.13245>
- [8] Wild, C. J., Norton, L., Menon, D. K., Ripsman, D. A., Swartz, R. H., & Owen, A. M. (2022). Disentangling the cognitive, physical, and mental health sequelae of COVID-19. *Cell Reports Medicine*, 3(10). <https://doi.org/10.1016/j.xcrm.2022.100750>
- [9] Scott, S., Knott, V., Finlay-Jones, A. L., & Mancini, V. O. (2022). Australian Psychologists Experiences with Digital Mental Health: a Qualitative Investigation. *Journal of Technology in Behavioral Science*. <https://doi.org/10.1007/s41347-022-00271-5>
- [10] Champion, J., Javed, A., Lund, C., Sartorius, N., Saxena, S., Marmot, M., Allan, J., & Udomratn, P. (2022). Public mental health: required actions to address implementation failure in the context of COVID-19. *The Lancet Psychiatry*, 9(2), 169–182.
[https://doi.org/10.1016/s2215-0366\(21\)00199-1](https://doi.org/10.1016/s2215-0366(21)00199-1)
- [11] Singh, S., Roy, D., Sinha, K., Parveen, S., Sharma, G., & Joshi, G. (2020). Impact of COVID-19 and Lockdown on Mental Health of Children and Adolescents: A Narrative Review with Recommendations. *Psychiatry Research*, 293(113429).
<https://doi.org/10.1016/j.psychres.2020.113429>
- [12] Al-Aly, Z., Xie, Y., & Bowe, B. (2021). High-dimensional characterization of post-acute sequelae of COVID-19. *Nature*, 1–8. <https://doi.org/10.1038/s41586-021-03553-9>
- [13] Rogers, J. P., Chesney, E., Oliver, D., Pollak, T. A., McGuire, P., Fusar-Poli, P., Zandi, M. S., Lewis, G., & David, A. S. (2020). Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: a systematic review and meta-analysis with comparison to the COVID-19 pandemic. *The Lancet Psychiatry*, 7(7). [https://doi.org/10.1016/s2215-0366\(20\)30203-0](https://doi.org/10.1016/s2215-0366(20)30203-0)
- [14] Mathrani, A., Sarvesh, T., & Umer, R. (2021). Digital divide framework: online learning in developing countries during the COVID-19 lockdown. *Globalisation, Societies and Education*, 1–16. <https://doi.org/10.1080/14767724.2021.1981253>
- [15] Spanakis, P., Peckham, E., Mathers, A., Shiers, D., & Gilbody, S. (2021). The digital divide:

- amplifying health inequalities for people with severe mental illness in the time of COVID-19. *The British Journal of Psychiatry*, 219(4), 1–3. <https://doi.org/10.1192/bjp.2021.56>
- [16] Zheng, Y., & Walsham, G. (2021). Inequality of what? An intersectional approach to digital inequality under Covid-19. *Information and Organization*, 31(1), 100341. <https://doi.org/10.1016/j.infoandorg.2021.100341>
- [17] Asadi-Pooya, A. A., & Simani, L. (2020). Central nervous system manifestations of COVID-19: A systematic review. *Journal of the Neurological Sciences*, 413, 116832. <https://doi.org/10.1016/j.jns.2020.116832>
- [18] Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The Psychological Impact of Quarantine and How to Reduce it: Rapid Review of the Evidence. *The Lancet*, 395(10227), 912–920. [https://doi.org/10.1016/S0140-6736\(20\)30460-8](https://doi.org/10.1016/S0140-6736(20)30460-8)
- [19] Gao, J., Zheng, P., Jia, Y., Chen, H., Mao, Y., Chen, S., Wang, Y., Fu, H., & Dai, J. (2020). Mental health problems and social media exposure during COVID-19 outbreak. *PLOS ONE*, 15(4). <https://doi.org/10.1371/journal.pone.0231924>
- [20] Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., McIntyre, R. S., Choo, F. N., Tran, B., Ho, R., Sharma, V. K., & Ho, C. (2020). A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. *Brain, Behavior, and Immunity*, 87. <https://doi.org/10.1016/j.bbi.2020.04.028>
- [21] Zhang, J., Lu, H., Zeng, H., Zhang, S., Du, Q., Jiang, T., & Du, B. (2020). The differential psychological distress of populations affected by the COVID-19 pandemic. *Brain, Behavior, and Immunity*, 87, 49–50. <https://doi.org/10.1016/j.bbi.2020.04.031>
- [22] Arendt, F., Scherr, S., & Romer, D. (2019). Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media & Society*, 21(11-12), 146144481985010. <https://doi.org/10.1177/1461444819850106>
- [23] Baker, D., & Fortune, S. (2008). Understanding Self-Harm and Suicide Websites. *Crisis*, 29(3), 118–122. <https://doi.org/10.1027/0227-5910.29.3.118>
- [24] Hampshire, A., Trender, W., Chamberlain, S. R., Jolly, A. E., Grant, J. E., Patrick, F., Mazibuko, N., Williams, S. C., Barnby, J. M., Hellyer, P., & Mehta, M. A. (2021). Cognitive deficits in people who have recovered from COVID-19. *EClinicalMedicine*, 0(0). <https://doi.org/10.1016/j.eclinm.2021.101044>
- [25] Zanin, L., Saraceno, G., Panciani, P. P., Renisi, G., Signorini, L., Migliorati, K., & Fontanella, M. M. (2020). SARS-CoV-2 can induce brain and spine demyelinating lesions. *Acta Neurochirurgica*, 162(7), 1491–1494. <https://doi.org/10.1007/s00701-020-04374-x>
- [26] Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for

- common mental disorders. *World Psychiatry*, 13(1), 4–11. <https://doi.org/10.1002/wps.20083>
- [27] Balcombe, L., & De Leo, D. (2021). Digital Mental Health Challenges and the Horizon Ahead for Solutions. *JMIR Mental Health*, 8(3), e26811. <https://doi.org/10.2196/26811>
- [28] Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., Üstün, T. B., & Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *Epidemiologia E Psichiatria Sociale*, 18(1), 23–33. <https://doi.org/10.1017/s1121189x00001421>
- [29] Wang, Q., Xu, R., & Volkow, N. D. (2020). Increased risk of COVID -19 infection and mortality in people with mental disorders: analysis from electronic health records in the United States. *World Psychiatry*, 20(1), 124–130. <https://doi.org/10.1002/wps.20806>
- [30] Jiao, W. Y., Wang, L. N., Liu, J., Fang, S. F., Jiao, F. Y., Pettoello-Mantovani, M., & Somekh, E. (2020). Behavioral and Emotional Disorders in Children during the COVID-19 Epidemic. *The Journal of Pediatrics*, 221(1). <https://doi.org/10.1016/j.jpeds.2020.03.013>
- [31] Lee, J. (2020). Mental health effects of school closures during COVID-19. *The Lancet Child & Adolescent Health*, 4(6), 421. [https://doi.org/10.1016/s2352-4642\(20\)30109-7](https://doi.org/10.1016/s2352-4642(20)30109-7)