NLP approach for Mental Health Problems Associated with Social Media Activities

Daniyal Hotiana Brianni Duncan Deepthi Tamma Will Courtney

***Abstract*—While social media has brought countless social benefits, the increased inter connectivity has introduced unforeseen social issues. Among younger people in particular, where social media use is ubiquitous, there is concern that social media produces an adverse effect on mental health. According to an article regarding Social Media published by the National Center for Health Research, "With 13% of 12–17-year-olds reporting depression and 32% reporting anxiety, mental illness is a concern for adolescent health. It is a concern for young 2adults as well since 25% of 18–25-year-olds report having some form of mental illness".We use machine learning models to evaluate these effects.**

***Index Terms*—Social Media, Mental Health, Depression, Social Benefit**

I. INTRODUCTION

In today's world, a majority of individuals depend on social media platforms such as Facebook, Twitter, Snapchat, YouTube, and Instagram to associate with one another. Social media is a tool that offers users entertainment, creative expression, networking opportunities, more accessible access to an abundance of resources. Ironically, for a technology intended to unite individuals, engaging a tremendous amount of time can potentially leave users feeling lonely and insecure, worsening mental health issues like Anxiety and Depression. Given that mental health is an imperative aspect of an overall healthy lifestyle, it is vital to consider how millions of people utilize and rely on social media applications daily, making them vulnerable to mental illnesses associated with social media use. In addition to this, a particular demographic may be at high risk: a target age group, people of color, underprivileged communities, and a particular gender. Overall, Big Data and Social Networking do not seem to be going anywhere. If companies desire to achieve longevity in customer/user satisfaction, they must accept responsibility for this phenomenon and seek improvisation.

Coexisting in a world containing approximately 7.8 billion humans , it is common to suggest that humanity thrives on social interaction. Human interaction is inevitable as it is an element that is practiced daily: through work, homelife, academia and even occurs in random encounters with strangers. Social interaction is the driving force that encourages humans to self-expression, collaborates with others, and develops cognitive solid skills. Socialization is a vital human need theorized by humanist Abraham Maslow []. It has continuously been proven over many years by researchers that the deficiency of social needs can place heinous effects on an individual's wellbeing. On the contrary, due to social progression, humanity sought out countlessdiscoveries and constructed ingenious inventions. social Bots can interacts with human users [1]–[6] and cause mental health problems.

With the rapid advancement of technology, one of the most arguably impactful innovations developed during the present Information Age is described as the World Wide Web; the platform grants millions of people access to copious amounts of information at one's fingertips. [7] This platform was established by Tim-Berners Lee in 1989 [8]. Social media, one of the WWW's most significant byproducts, has created a communication pipeline between users and technology. We decided to examine the relationship between humans and their social media activity, mainly focusing on adolescent age users. Many online users, especially teenagers, rely on social media platforms to connect; however, recent studies find that increased use can deter mental health. While there are many positive aspects of social networking, the number of studies suggesting that excessive use is causing a decline in mental health is increasing exponentially. literature research is essential to our study. Mental Health like depression and anxiety can be discovered based on online posts [9]–[12] [13].

Social media engagement can make people feel closer to society and their near and dear, but this also can lead to FOMO (Fear of Missing Out) when left alone, potentially causing disruptions in mental health. It can cause financial problems as well [14] Ironically, social media activates the brain's reward center by releasing dopamine, which is recognized as a happy chemical. Dopamine releases during pleasurable experiences, such as eating a favorite food, interacting with loved ones, and exercising; when it is lacking, this leads to Anxiety and Depression.

Similarly, the social media activities are designed in such a way that it is addictive and brings Anxiety and Depression when it is unavailable. However, where there can be positive effects of social media, there are even more negative impacts on children and adults. [15]

One of the most prevalent age groups that utilize social media is teenagers. With social media existing as a platform for competition for popularity, teenagers respond to societal beauty norms, popular, attractive, and appealing. A study conducted by Maria Clark in 40+ Frightening Social Media and Mental Health Statistics shows a linkage where the suicide rates of teenagers have risen from 2011 by 150 percent, and the depression rate has grown by 112 percent. It is pretty noticeable that the rate is increasing per the introduction of social media networking platforms such as Facebook, WhatsApp, Instagram. In an effort to respond, Instagram has suppressed likes to curb the comparisons and hurt feelings associated with attaching popularity to sharing content. Nevertheless, this made the most negligible impact on teenagers' mental health changes, and rates have not improved effectively. [13]

II. Literature review

Numerous articles exist suggesting a correspondence between social media use and mental health risks. This literature review will evaluate two articles: Social Media Use and Its Connection to Mental Health: A Systematic Review and Social Media and Adolescents' and Young Adults' Mental Health. We assess these particular articles for several reasons. The first being to adopt data collection and processing methods from knowledgeable researchers. Transfer Learning approaches in Social media helps to detect different trends [16] [17]–[32]. The second reason is to assist in defining a research problem and proposing a solution. Lastly, our group intends to seek a clear understanding of how best technology utilizes in research strategies.

In Social Media Use and Its Connection to Mental Health: A Systematic Review, authors Fazida Karim, Azeezat Oyewande, Lamis Abdalla, Reem Ehsanullah, and Safeera Khan make a strong claim by suggesting that social media is a great contributor to mental health issues. Our proposition is also similar to Karim's, suggesting that increased social media usage propels mental health decline. As for their research strategy, the authors supported their claim by first collecting fifty papers from Google Scholar databases. The authors targeted articles containing specific keywords, such as "social media," "mental health", "social media AND mental health," "social networking," and "social networking," OR "social media" AND "mental health." The keywords searched generated hundreds of thousands of results in the Google Scholar database.

Of the hundreds of thousands of articles retrieved, fifty were selected and put through an inclusion-exclusion criteria process, resulting in 28 articles narrowed down for final selection. The selection process's requirements included chronological relevance, relation to thesis statements, English language, and discarding of duplicates. Correspondingly, 16 articles were selected where the focus prevailed on adults and their gender and preadolescents. Both quantitative and qualitative studies were conclusively selected for a full investigation. The researchers found that anxiety and depression were the most frequent mental disorders discussed in the pool of 16 articles examined.

Our research group adopted a similar method in our systematic study by extracting one hundred scholarly articles highlighting the relationship between social media usage and mental health. The two central databases accessed were Google Scholar and George Mason University's online library. Collectively, we gathered twenty-five articles per student, focusing on content containing similar keywords of "social media," "social networking," "mental health," "mental disorder," and "technology and mental health." The second article we are reviewing, Social Media and Adolescents' and Young Adults' Mental Health, studies the relationship between social media use among adolescents and how their mental health is in decline. This article supports the first composition conclusions claiming that 25% of adolescent users believe their social media habits negatively impact their mental health. A strong link also exists between sleep patterns and high social media usage. Poor sleep patterns commonly lead to depression and anxiety, two mental illnesses prominently cited in the first article. [16] [16]

Our group intends to produce results that are consistent with the two articles. Similar to the two studies discussed, we evaluate the correlation between social media use and mental health and their associated risks with high usage. Our goal in this research is to identify the diagnoses and symptoms of mental illness prominent in high social media use. Although anxiety and depression were of frequent discussion, we will not limit our findings to those two and plan to explore additional potential diagnoses. Ultimately, our theoretical research problem entails that there is not enough evidence to prove that social media usage causes a decline in mental health. Applying articles published within the past five years, we hope to research this topic and find improved studies and research methods to support our thesis.

III. Problem Description

A. Problem Statement

Thousands of published articles suggest a linkage between social media usage and mental illness. However, the articles contain information gaps. Many of these publications do not explicitly address adequate elements of the argument. For example, publications will not address all of the Five W's (who, what, where, when, and why). To better understand the mental illness caused by social media usage, and whom they target, we will locate the demographics at high risk through social media platforms such as Twitter, Reddit, Facebook, and Instagram. We also determine an association between social networking and depressive symptoms, fluctuations in self-esteem, and other potential psychiatric problems and issues. Lastly, we intend on identifying shortcomings, gaps, and weaknesses in the research of our topic and propose a solution.

B. Problem Investigation

[33]

Mental health crises can lead to tragic results like suicide, and social media must therefore be studied to determine how significant a contributor to the problem is. Our group will investigate factors and activity based on social media, which has affected this age group's mental health because we believe Social Media is a critical player in affecting mental health. To investigate our problem statement, we have collected various articles with mental health and social media in their title. Our research will explore how social media has affected the mental health of people living in the United States, focusing on teenagers and young adults. By extracting the articles' qualitative data, we prepared our dataset containing text providing insights into our problem statement. [34]

C. Software Required:

1) R-Studio
2) Python
3) MS Excel
4) Tableau
5) MonkeyLearn

1) R-Studio: R-Studio is an open-source tool used for data modeling. The software provides a broad range of statistical computing and graphical techniques and is profoundly extensible. Here, we have used R-Studio to extract the data from the collected articles and web pages to the Excel File. Here, we have used various libraries such as rvest, pdftools, stringr, xlsx, and openxlsx. [35]

2) Python: Python Programming prepares text information for the requirements in varying data analyses. An important area of application of Python's text processing ability is NLP (Natural Language Processing). Here, we have used Python to clean the text, i.e., remove the special characters, stop words, and punctuate. In our project, we have used libraries like nltk, sklearn, and BeautifulSoup. [36]

3) Tableau: Tableau is one of the best visualization tools used for Text Analytics. The software provides vivid visualizations and pleasing results of the text once the data is well organized, i.e., after data pre-processing. Here, we have used Tableau to create visualizations such as Word Cloud, Tree Diagram, and Word Drill. [37]

4) MonkeyLearn: MonkeyLearn is a machine learning tool designed for text analysis. It allows for the training of a custom machine learning tool that we require. It includes tools for text extraction, text classification, and integration. [38]

5) MS Excel: Excel is one of the powerful tools which we used for text extraction. Excel is mainly used to convert our huge amount of unstructured data that is extracted from our gathered articles into structured data. In addition to that we have split the body of text into single words using Excel. Excel provides various features such as Sentence Counts, Sentiment Analysis, Word Counts and word cloud. [39]

IV. Approach

Analytics begins with raw data, so our main priority was finding reliable data sources to support our thesis. We have gathered articles from various resources and used R Studio for web scraping and converting into CSV files. Once all our data converts into a single structured format, we have performed data pre-processing on the extracted data. Data is pre-processed and cleaned using Python, eliminating any anomalies, including removing the stop words, connective words, and other special characters or missing data. Since the data gathered comes from various resources, we will be creating a standard format to analyze data further and gain valuable insight.

Our research explores how social media has affected the mental health of people living in the United States, focusing on teenagers and young adults. Our group investigated the factors and activity based on social media, which has affected this age group's mental health because we believe Social Media is a critical player in affecting mental health. To address questions guiding us in understanding why and how social media affects mental health, various approaches will be applied to narrow down our answers.

We have applied a combination of qualitative and quantitative analysis approach to understand our problem at hand. To identify people at risk, we first took a qualitative approach and organized people into age groups, and their corresponding genders who have identified their mental health have been affected by social media use. This process is achieved through text mining. We have selected articles where social media and mental health are both in the title. We analyzed data to determine if social media usage over the last decade is correlated to the number of mental issues after the year 2010 because that is when the social media platform became mainstream. We have created visualizations representing the relationship between a user's mental health and social media activity. [40]

Our preliminary analysis shows that factors such as cyberbullying, lack of sleep, disconnect from real-world relationships, low self-esteem, and self-comparison to influencers are all potentially damaging effects of Social Media on mental health. We took a quantitative approach to perform descriptive analysis using R. Once we have generated our results, we will be moving our data to Tableau for visualizations to communicate our findings.

V. Data Pre-Processing

We have gathered the articles that incorporate the terms "Social Media" and "Mental Health" in the title for pre-processing the data. These articles were gathered from various sources such as GMU Resources (Online Database and libraries). Once we gathered all the related articles we used R Studio for web scraping using the HTML Method. We used multiple libraries such as Rvest and PDFtools in R studio inorder to extract data from websites as well as pdf. Codes for 3extraction shown below for websites and Pdf:

Codes were repeated for all the websites and PDFs that we chose based on the category that we selected. The above code stores the text from the website and pdf in the list format in R environment, so for our analysis purposes we had to convert the list into a dataframe. Once all the lists were converted into dataframes, data frames were exported to our local hard drive for the purpose of combining all the data frames. Using MS Excel CSV format for the exported files to combine data. Once all the data frames were combined, cleaning processes started where any empty row was taken out, any row which had irrelevant data was deleted such as references, author names, title etc. That took most of our time as we went in manually to verify that once data is loaded it does not have any unnecessary text that can affect the end result which is referred to as noise.

Even after converting our dataframes into csv after cleaning we realised that the code that we used stored each line from the text was separated by comma not each word so we used MS excel feature data tab, text to column feature is used. through which each word was separated into its own column. Since there were still a lot of unnecessary words in our data such as stop words we used MS excel feature find and replace on all the stop words that were present through which a lot of unnecessary data was deleted we finally had our clean data ready for Text Analysis.

VI. Text Analysis

Once we had created a clean dataset, we were ready to begin analysis of the text. Importing the dataset into Monkeylearn, a text analysis tool that uses machine learning, we were able to generate meta-data about the combined dataset. First was a simple word count to show which areas were focused by the researchers (Figure 1). From the word count, it is simple to see which areas were commonly addressed in the papers: social media, mental health, trangender people, and more. Many of the studies were focused on adolescents and other young people, also demonstrated in the word count. A word count such as this one allowed us to quickly see which areas of classification to focus on moving forward.

After a count of single words, the next step would be a count of particular phrases that showed up multiple times across the dataset. (Figure II) This dataset assists with stringing concepts together and giving more context in certain areas. For example, ”face to face communication,” ”minority stress,” ”parental monitoring,” ”self esteem,” and other contextual clues emerge that are not obvious in the single-word count, particularly in terms of sentiment analysis. For example, ”minority stress” is evocative and meaningful, giving better understanding of a negative sentiment than the words ”minority” and ”stress” convey individually. This expansion of the metadata allows the reader to better understand the areas of research and concern without having to directly read the dozens of journal articles. These areas will help in accurate classification for machine learning.

For more advanced analysis, we turned to the machine-learning capabilities of Monkeylearn. While the tool could easily produce basic metadata about the text, it had not done any learning and could not understand the context of the text, which we needed to provide. To do so, we manually tagged a portion of the text (about 10-20%) in 13 different categories, ranging from anxiety and depression to transgenderism and bullying. Eventually, the machine had learned enough to classify different phrases with reasonable accuracy, as displayed in Figures 3 and 4. After training, the system was able to determine the instances of certain concepts more accurately than a simple word count was ever able to do. For example, the trained machine was able to find 20 instances of cyberbullying being mentioned, while the word count only found six. Social media was found in over 100 instances, instead of the 21 counted in the simple word count. Machine learning essentially

Fig. 1. Figure III

allowed for a more accurate and better understanding of the metadata contained within the paper.

VII. Challenges and Future Goals

Our team faced challenges throughout the project. Occasionally we had to change our methodology to accommodate those challenges. However, we anticipated challenges working with a large unstructured dataset and were not surprised when those concerns were realized. The first challenge that we came across was identifying the relevant data to be processed, as mentioned above only the articles with certain criteria were selected. After the data was identified we had to identify how best to extract the data using R Studio. We decided to extract from PDF files rather than HTML as we found the text to be cleaner. Once the data was extracted and converted into .CSV format, the biggest challenge was in cleaning the data by removing noise as different formats, empty rows, stopwords etc. Although cleaning was partially successful, we learned that additional pre-processing will be necessary to meet the ideal conditions for machine learning.

We also had challenges in the text analysis. Despite our efforts to clean the data, there were still instances of commas and missing spaces which confused the computer trying to read the data. This is reflected in word counts being generally

Fig. 2. Figure IV

less than the "true" word counts would be if they were counted by a human. While training Monkeylearn's machine learning algorithm helped alleviate these issues, the software still incorrectly labelled certain areas or didn't label them at all. Particularly troublesome were the false negatives, i.e the system not detecting a key phrase when it exists. The false negative count was very high before training, and was still the biggest weakness after training. False positives also exist, as demonstrated in Figure III where "suicidal ideation" was classified in the 'anxiety' tag instead of the more appropriate 'suicide' tag. As more time was spent training the algorithm, cases of false negative and false positives both decreased. With a larger dataset and more time spent training, the machine learning algorithm would have no doubt been more accurate.

For our future deliverables, we seek to improve on our methodology and provide even more accurate and robust findings. As previously mentioned, we will be doing further cleaning of the data to help the machine learning algorithms in classifying their categories correctly. We plan on selecting a smaller subset of the most valuable articles we found, performing a thorough cleaning, and then seeing what more advanced analysis we can run on them, e.g a true semantic analysis and multiple regression. Our discussions and lessons learned to data now give us a clear path ahead to maximize the amount of metadata we can produce for the data we have.

VIII. Analytical Conclusion

Our project seeks to determine the factors that researchers find to be most associated with social media and mental health. Although we included a large number of articles within general parameters, and were able to find meaningful results with both basic tests and machine learning algorithms. A simple word count like Figure I or word cloud like Figure IV can give clues as to which topics are most commonly seen in these papers. We found a wider variety of symptoms and issues, such as minority or transgender vulnerabilities, and concerns about children and young teens than we originally anticipated. We also found cyberbullying and addiction to social media to have a wide footprint across the scholarly literature on the subject. These findings are key to our classification and further work we are doing through machine learning.

We found machine learning to be a powerful tool to assist with categorization and classification of the myriad subjects within our dataset. Once it was developed, our model was able to rapidly categorize the sentiments of unstructured text with reasonable accuracy. The results were clear, that the warnings and concerns of social media and mental health by academic and medical professionals far overshadowed the benefits, particularly among young people. The social negatives of media addiction, cyberbullying, loneliness, and other damaging social phenomena are found consistently throughout the academic literature. While there are certainly benefits to social media, as in the data, there is certainly enough content from just our analysis thus far to show that social media is a double-edged sword at best. As we conduct further data processing, analysis, and machine learning, we aim to have even more quantifiable evidence to demonstrate our conclusion in the final product.

References

[1] M. Heidari, J. H. J. Jones, and O. Uzuner, "Deep contextualized word embedding for text-based online user profiling to detect social bots on twitter," in *IEEE 2020 International Conference on Data Mining Workshops (ICDMW), ICDMW 2020*, 2020.

[2] U.S.Gov, "U.S. and World Population Clock." https://www.census.gov/popclock/, 21 April 2021.

[3] M. Heidari and J. H. Jones, "Using bert to extract topic-independent sentiment features for social media bot detection," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pp. 0542–0547, 2020.

[4] M. Conway and D. O'Connor, "Social media, big data, and mental health: current advances and ethical implications," *Current Opinion in Psychology*, vol. 9, pp. 77–82, June 2016.

[5] Dr. C. George Boeree, "Abraham maslow." http://webspace.ship.edu/cgboer/maslow.html, 2006.

[6] M. Heidari, J. H. J. Jones, and O. Uzuner, "An empirical study of machine learning algorithms for social media bot detection," in *IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021*, 2021.

[7] World Wide Web Foundation, "History of the web." https://webfoundation.org/about/vision/history-of-the-web/, 11 April 2021.

[8] Harvard Medical School Affiliate, "The social dilemma: Social media and your mental health." https://www.mcleanhospital.org/essential/it-or-not-social-medias-affecting-your-mental-health, February 9, 2021.

[9] S. Zad, M. Heidari, J. H. J. Jones, and O. Uzuner, "Emotion detection of textual data: An interdisciplinary survey," in *IEEE 2021 World AI IoT Congress,AIIoT2021*, 2021.

[10] B. Dao, T. Nguyen, S. Venkatesh, and D. Phung, "Nonparametric discovery of online mental health-related communities," in *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1–10, 2015.

[11] B. Dao, T. Nguyen, S. Venkatesh, and D. Phung, "Effect of social capital on emotion, language style and latent topics in online depression community," in *2016 IEEE RIVF International Conference on Computing Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*, pp. 61–66, 2016.

[12] S. Zad, M. Heidari, J. H. J. Jones, and O. Uzuner, "A survey on concept-level sentiment analysis techniques of textual data," in *IEEE 2021 World AI IoT Congress,AIIoT2021*, 2021.

[13] E. Mir, C. Novas, and M. Seymour, "Social media and adolescents' and young adults' mental health." https://www.center4research.org/social-media-affects-mental-health/, 2021.

[14] M. Heidari, S. Zad, and S. Rafatirad, "Ensemble of supervised and unsupervised learning models to predict a profitable business decision," in *IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021*, 2021.

[15] M. Clark, "40+ frightening social media and mental health statistics." https://etactics.com/blog/social-media-and-mental-health-statistics, 12 November 2020.

[16] C. Berryman, C. J. Ferguson, and C. Negy, "Social media use and mental health among young adults," *Psychiatric Quarterly*, vol. 89, pp. 307–314, Nov. 2017.

[17] M. M. Aldarwish and H. F. Ahmad, "Predicting depression levels using social media posts," in *2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*, pp. 277–280, 2017.

[18] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot, and Y. Kaewpitakkun, "Facebook social media for depression detection in the thai community," in *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 1–6, 2018.

[19] M. Heidari and S. Rafatirad, "Using transfer learning approach to implement convolutional neural network model to recommend airline tickets by using online reviews," in *2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA*, pp. 1–6, 2020.

[20] H. Yang, C. He, H. Zhu, and W. Song, "Prediction of slant path rain attenuation based on artificial neural network," in *2000 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 1, pp. 152–155 vol.1, 2000.

[21] A. Zahura and K. A. Mamun, "Intelligent system for predicting suicidal behaviour from social media and health data," in *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*, pp. 319–324, 2020.

[22] L. Liu, B. Li, I.-M. Chen, T. J. Goh, and M. Sung, "Interactive robots as social partner for communication care," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2231–2236, 2014.

[23] M. Heidari and S. Rafatirad, "Bidirectional transformer based on online text-based information to implement convolutional neural network model for secure business investment," in *IEEE 2020 International Symposium on Technology and Society (ISTAS20), ISTAS20 2020*, 2020.

[24] L. Dobrescu, S. Obreja, M.-C. Vochin, D. Dobrescu, and S. Halichidis, "New approaches for quantifying internet activity," in *2019 E-Health and Bioengineering Conference (EHB)*, pp. 1–4, 2019.

[25] A. J. Majumder, J. W. Dedmondt, S. Jones, and A. A. Asif, "A smart cyber-human system to support mental well-being through social engagement," in *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 1050–1058, 2020.

[26] J. Parraga-Alava, R. A. Caicedo, J. M. Gómez, and M. Inostroza-Ponta, "An unsupervised learning approach for automatically to categorize potential suicide messages in social media," in *2019 38th International Conference of the Chilean Computer Science Society (SCCC)*, pp. 1–8, 2019.

[27] M. Heidari, J. H. J. Jones, and O. Uzuner, "Misinformation detection model to prevent spread of the covid-19 virus during the pandemic," 2021.

[28] F. Murtagh, "Analysing activities, contextualized for general health, depression and demographics," in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 547–554, 2018.

[29] C.-H. Tai, Z.-H. Tan, and Y.-S. Chang, "Systematical approach for detecting the intention and intensity of feelings on social network," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 4, pp. 987–995, 2016.

[30] M. Heidari, S. Zad, B. Berlin, and S. Rafatirad, "Ontology creation model based on attention mechanism for a specific business domain," in *IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021*, 2021.

[31] J. Du, Y. Zhang, C. Tao, and H. Xu, "A pilot study of mining association between psychiatric stressors and symptoms in tweets," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1254–1257, 2017.

[32] M. Heidari and S. Rafatirad, "Semantic convolutional neural network model for safe business investment by using bert," in *IEEE 2020 Seventh International Conference on Social Networks Analysis, Management and Security, SNAMS 2020*, 2020.

[33] L. Robinson and M. Smith, "Social media and mental health." https://www.helpguide.org/articles/mental-health/social-media-and-mental-health.htm, September 2020.

[34] H. Patel, "An introduction to web scraping using R." https://www.freecodecamp.org/news/an-introduction-to-web-scraping-using-r-40284110c848/, 24 October 2018.

[35] tutorialspoint, "Python - text processing." https://www.tutorialspoint.com/python_text_processing/index.htm, 2021.

[36] K. Flerlage, "A starter kit for text analysis in tableau." https://www.flerlagetwins.com/2019/09/text-analysis.html, 28 September 2019.

[37] R. Thorstad and P. Wolff, "Predicting future mental illness from social media: A big-data approach," *Behavior Research Methods*, vol. 51, pp. 1586–1600, Apr. 2019.

[38] A. MacCaw, G. Cabane, R. Fishkin, and S. Blum, "[14] text analysis with monkeylearn." https://monkeylearn.com/.

[39] K. Yap, "Text analysis using excel." https://www.keithyap.com.au/text-analysis-using-excel/#topic-modelling, 25 November 2016.

[40] S. Jain, "Ultimate guide to deal with text data (using python) – for data scientists and engineers." https://www.analyticsvidhya.com/blog/2018/02/the-different-methods-deal-text-data-predictive-python/, 27 February 2018.