# Social Media's Influence on Mental Health

Will Courtney
wcourtn3@gmu.edu

Research Contribution: Will Courtney has contribution in both research and coding part of this work.

***Abstract*—The technological advancement that is continually occurring has brought forth many unique innovations. Social networking is arguably one of the most impactful mediated technologies produced by the Information Age. Social media facilitates a profound approach to communication between humans across the globe. While social media platforms' usage grants countless benefits, researchers hold social media platforms responsible for the decline in mental health, especially adolescent users. Hundreds of thousands of articles published by medical researchers suggest psychological aberration; there has not been enough evidence supporting a direct cause. Qualitative and quantitative methods will be applied to clarify our inquiry to address the questions that prompt us to learn social media's mental health effects. The influence of social media on a user's health will be further analyzed.**

***Index Terms*—Social Media, Social Networking, Engagement, Mental Health**

## I. Introduction

Coexisting in a world containing approximately 7.8 billion humans [1]–[3], 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 [4]. 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 [5] [6].

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. 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, or shows fake trends [**?**], [7]–[13]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. [14], [15]

One of the most prevalent age groups that utilize social media is teenagers [16]–[19]. 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. [20]

## II. Literature review

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. [21], Transfer Learning in Social media helps to detect these trends [21] [22]–[36].

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. Background

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. [37]

## IV. 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

While social media has brought countless social benefits, the increased interconnectivity 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". [38]

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. [39]

### 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. [40]

*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. [41]

*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. [42]

*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. [43]

*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. [44]

## V. 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. [45]

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.

## VI. 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.

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 data frame. 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.

## VII. 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.

| Some top phrases containing 4 words (without punctuation marks) | Occurrences |
|---|---|
| face to face communication | 2 |

| Some top phrases containing 3 words (without punctuation marks) | Occurrences |
|---|---|
| parental monitoring and | 2 |
| and social media | 2 |
| in order to | 2 |
| read and accept | 2 |
| for instance it | 2 |
| of the data | 2 |
| of social media | 2 |
| based on the | 2 |
| face to face | 2 |
| to face communication | 2 |
| minority stress and | 2 |
| with social media | 2 |
| of the models | 2 |
| the present study | 2 |
| model of behavior | 2 |
| self esteem and | 2 |
| blogs have been | 2 |
| it is clear | 2 |
| in the early | 2 |
| parent reported number | 2 |
| tumblr transition blogs | 2 |
| well being and | 2 |
| and so on | 2 |

| Some top phrases containing 2 words (without punctuation marks) | Occurrences |
|---|---|
| social media | 21 |
| of the | 18 |
| mental health | 11 |

Fig. 1. Figure II

## VIII. 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.

## References


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

[2] A. H. Yazdavar, M. S. Mahdavinejad, G. Bajaj, K. Thirunarayan, J. Pathak, and A. Sheth, "Mental health analysis via social media data," in *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 459–460, 2018.

[3] Y. Zhao, Y. Guo, X. He, J. Huo, Y. Wu, X. Yang, and J. Bian, "Assessing mental health signals among sexual and gender minorities using twitter data," in *2018 IEEE International Conference on Healthcare Informatics Workshop (ICHI-W)*, pp. 51–52, 2018.

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

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

[6] 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.

[7] 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.

[8] R. Ramadan, S. Alqatawneh, F. Ahalaiqa, I. Abdel-Qader, A. Aldahoud, and S. AlZoubi, "The utilization of whatsapp to determine the obsessive-compulsive disorder (ocd): A preliminary study," in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 561–564, 2019.

[9] S. T. Sadasivuni and Y. Zhang, "Finding a depressive twitter user by analyzing depress and antidepressant tweets," in *2020 IEEE India Council International Subsections Conference (INDISCON)*, pp. 142–145, 2020.

[10] S. R. Kamite and V. B. Kamble, "Detection of depression in social media via twitter using machine learning approach," in *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)*, pp. 122–125, 2020.

[11] 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.

[12] G. Liu, C. Wang, K. Peng, H. Huang, Y. Li, and W. Cheng, "Socinf: Membership inference attacks on social media health data with machine learning," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 5, pp. 907–921, 2019.

[13] 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.

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

[15] M. Mahat, "Detecting cyberbullying across multiple social media platforms using deep learning," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 299–301, 2021.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

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

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

[42] 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.
[43] A. MacCaw, G. Cabane, R. Fishkin, and S. Blum, “[14] text analysis with monkeylearn.” https://monkeylearn.com/.
[44] K. Yap, “Text analysis using excel.” https://www.keithyap.com.au/text-analysis-using-excel/#topic-modelling, 25 November 2016.
[45] 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.