



SOCIAL MEDIA-BASED DEPRESSION PREDICTION USING DEEP LEARNING

Pratibha

Research Scholar, Dr. K. N. Modi University Newai, Rajasthan

Dr. Neha Gupta

Associate Professor, Dr. K. N. Modi University Newai, Rajasthan

Abstract

Social media sites like Facebook, Instagram, and Twitter have grown in popularity to the point that users often share personal thoughts and feelings with the world. Because of this, social media may be a great tool for finding the first symptoms of mental health issues like depression. A deep learning-based model for depression prediction using user-generated content analysis (UGC)—pictures, captions, hashtags, and text—is the primary focus of this study. In order to identify patterns in behaviour and emotions, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are used, while Natural Language Processing (NLP) methods are used to extract relevant characteristics from textual data. The study's overarching goal is to bolster early diagnosis and intervention by giving medical providers more resources to help them spot those who may be at danger. Improving mental health monitoring systems and enabling prompt psychological help are also possible outcomes of this method.

Keywords: Depression Detection, Social Media Analysis, Natural Language Processing (NLP), Mental Health Forecasting, Predictive Modeling.

INTRODUCTION

With the advent of social media platforms, individuals are able to express themselves in a manner that has never been seen before. These platforms provide a glimpse into society as a whole. As a consequence of the widespread presence of user-generated content on social media platforms, these platforms have become a veritable treasure mine of textual material that represents the most personal ideas, feelings, and experiences of individuals. Depression is a complicated mental condition that has a significant influence on people's health and happiness. When it comes to these feelings, millions of people struggle with clinical depression. A diagnosis of depression at an earlier stage raises the possibility of prompt intervention, support, and therapy being provided. The traditional methods of identifying depressed symptoms, on the other hand, often fail due to the limitations that they include, such as their reliance on subjective self-reporting or their failure to scale. The combination of the capabilities of Natural Language Processing (NLP) and Machine Learning (ML) is a technique that is both unique and optimistic for solving this essential problem.

The purpose of this research is to interpret the complex linguistic patterns and emotional signals that are included within the textual data that was acquired from social media sites. This will be accomplished via the use of natural language processing and machine learning methods. We want to construct a depression detection system that is reliable and accurate, and in order to do so, we seek for useful data among these large streams of unstructured data. This study takes into consideration a wider range of attitudes than merely positive and negative ones. Rather, it goes deeper into the usage of language in order to identify signals of emotional pain that people may otherwise miss. This is done from the perspective of the language itself. We have high hopes that this approach will assist us in gaining a better understanding of mental health in general and will also assist us in overcoming the digital fog that surrounds depression.

Definition

Depression identification from social media textual data is a technique that makes use of Natural Language Processing (NLP) and Machine Learning (ML) to automatically identify instances of sorrow or emotional discomfort from the vast volumes of text-based information that are published by users of social media platforms. In order to interpret the feelings, emotional signals, and linguistic patterns that are included within the unstructured textual material that is provided by users of social media platforms, this cutting-edge technology makes use of the most advanced natural language processing algorithms now available. We are able to discover linguistic patterns that are associated with depression by using machine learning algorithms that have been trained on labelled data. Additionally, we are able to construct a prediction model that is capable of distinguishing between those who exhibit depressive symptoms and those who show emotional well-being.

The process typically involves several key stages:

1. **Data Collection:** When collecting textual data from various social media sites, it is important to keep privacy and ethics in mind. This may be accomplished by using web scraping tools or application programming interfaces (APIs).
2. **Data Preprocessing:** Processing and cleaning the raw text data in order to extract significant linguistic features while simultaneously eliminating extraneous information, personal identities, and background noise.
3. **Feature Extraction:** Performing processing on the text data that has been preprocessed in order to extract valuable characteristics, such as word frequencies, sentiment scores, emotional expressions, linguistic style, and other contextual aspects.
4. **Model Training:** For the purpose of training a machine learning classifier to identify patterns in the textual data that are related with emotional distress and depression, a tagged dataset is used.
5. **Model Evaluation:** determining the accuracy, precision, recall, and F1 score of the trained model in order to evaluate its performance; ensuring that the model is reliable when detecting depression.
6. **Deployment:** Putting the trained model to use by allowing it to automatically evaluate posts posted on social media in real time and classify them as perhaps depressed or not depressing. The application of natural language processing and machine learning for the purpose of identifying

depression in textual data from social media platforms might be of tremendous help to research and aid for mental health. When dealing with sensitive mental health information, researchers and developers have a need to prioritise ethical considerations such as data protection, authorisation, and responsible technology usage.

Background

Depressive disorders have become a significant problem in the field of public health due to the fact that they have a catastrophic impact on millions of individuals all over the world. Identification and treatment of this mental disease at an early stage is the only method to decrease the effect that it has on a person's life. On social media platforms, users are able to freely communicate their thoughts, emotions, and experiences with one another. These platforms have evolved into massive archives of human expression. Scholars and experts in the medical field are in agreement that social media may be able to assist in filling in the gaps in our knowledge about mental health. With all of the textual data that is accessible on these platforms, we are able to have a better understanding of the mental health of individuals. However, manually analysing such enormous data sets is not only impossible but also demands a significant amount of time. This has led to the development of novel methods for detecting depression, which make use of machine learning and natural language processing techniques. In the field of natural language processing (NLP), computers are able to automatically grasp and interpret human speech. On the other hand, machine learning (ML) algorithms may be trained to recognise patterns and classify text input according to certain features. Following the application of natural language processing and machine learning techniques to textual data obtained from social media platforms, the researchers want to identify linguistic markers and emotional signals that are associated with depression. Depression detection models that are accurate and scalable are produced by using features derived from a variety of sources, including sentiment analysis, language style, word usage, and context, amongst others. It is possible that the successful use of such models might have far-reaching repercussions.

Motivation

Depressive disorders have become a significant problem in the field of public health due to the fact that they have a catastrophic impact on millions of individuals all over the world. Identification and treatment of this mental disease at an early stage is the only method to decrease the effect that it has on a person's life. On social media platforms, users are able to freely communicate their thoughts, emotions, and experiences with one another. These platforms have evolved into massive archives of human expression. A consensus has been reached among academics and medical practitioners that social media platforms have the potential to shed light on mental health on a societal level. With all of the textual data that is accessible on these platforms, we are able to have a better understanding of the mental health of individuals. However, manually analysing such enormous data sets is not only impossible but also demands a significant amount of time.

Machine Learning and Natural Language Processing to Assess the Emotional Impact of Influencers' Mental Health Content on Instagram

There are two elements that are contributing to the growth in anxiety: the meteoric development of social media and the pervasiveness of mental health concerns in today's society. Both of these variables are

contributing to the rise in anxiety. To begin, there is no Internet tool that is more symbolic than social networks. These networks provide a continuous flow of information that reflects the current state of society and styles as they occur in real time. People spend over two and a half hours per day on social media, and as of October 2023, there were 4.95 billion users throughout the globe, which is equivalent to almost 61% of the total population of the planet. Despite WhatsApp's predominant position among all users (16.6%), Instagram comes in at a close second with 17.7% of the market share. In point of fact, Instagram has a significant amount of users among the younger generation, commanding the market share of both males and females between the ages of 16 and 24 (24.4%) and between the ages of 25 and 34 (20%). Over the course of the last several years, there has been a steady increase in the proportion of Instagram users who are male and between the ages of 16 and 24. Instagram is the second most popular social media network among Spanish Millennials, who are defined as those who are between the ages of 25 and 34. Seventy-four percent of users are on Instagram. A large section of the population, on the other hand, is now dealing with challenges related to their mental health.

OBJECTIVES

- 1. To examine social media data in order to find behavioural and linguistic indicators of depression.
- 2. To create a deep learning model that can accurately predict the likelihood of depression.

RESEARCH METHODOLOGY

A person will often experience depression more than once in their lifetime, and it has the potential to develop into a permanent recurring condition. The prevention of relapses and recurrences of depression is something that should be recommended strongly. more and more researchers are beginning to acknowledge the factors that increase the likelihood of recurrent depression. Real-time clinical information diagnostic has the potential to bring solutions to a number of challenges that remain unresolved in the field of medical research. Using 109 samples of client data from the National Institutes of Health, this article explores several categories of intelligent data analysis methodologies. The purpose of this investigation is to evaluate the need of intelligent data analysis in medical diagnosis (NIH). A discussion is held at the end of the article on the accuracy of the intelligent data analysis methods that were used and their usefulness in the field of medical research.

RESULT

The term "mental procedure" refers to a complex concept that involves consciousness, bodily sensitivity, and attitude. It is used to bring back the individual significance of a problem, an event, or a thing. It is possible to attach the survey of the mental assertions of social network data in order to develop trustworthy conceptions in a variety of various settings. In this study, we make use of psycholinguistic extents to consider five characteristics of the mental states that are presented in the comments.

Table 1: Disintegration of focus expressions into different segments

LIWC derived cues	Sample expression
-------------------	-------------------

Emotional procedure	
Positive emotion words	Pleased, Elated, Thrilled,
Negative emotion words	Anger, sadness, Loneliness, corona, dead
Sadness words	Unemployment, GDP, grief, unhappy, lockdown, jobless
Anger words	End, objection, trouble, hit, mental health
Anxiety words	Confusion, low mode, suicide, trigger
Lexical style	
Articles	A, an, the
Prepositions	For, in, of, to, with, above
Auxiliary verbs	Do, have, am, will
Adverbs	Quickly, slowly, very, really
Conjunctions	And, but, whereas
Total pronouns	I, them, itself
Personal pronoun	I, them, her
1st person singular pronoun	I, me, mine
1st person plural pronoun	We, us, our
2nd person	You, your
3rd person singular pronoun	He, she, her, him
3rd person plural pronoun	They, their, they'd
Impersonal pronouns	It, it's, those
Verbs	Go, good
Negation	Deny, disho

Table 2: Grouping of depression dataset

Data set details	Quantity
Total number of Twitter user's data	1709
Depression symbolic data	991
Non-depression symbolic data	715
Neutral symbolic data	3

Table 3: Examples of depression typical comments

Examples	Feedback
People with mental illness are not lazy. They're just exhausted from fighting their way through every single day.	YES
I am go in through alot which what ive bern goin through its different from wha everyone goes through cause my situations it never goes away	YES
Facts that's why I don't wanna wake up from a sleep anymore but insomnia is a bitch.	YES
Same here, someone that won't leave me for no reason	YES
The overwhelm tends to make me hyper fix on what's left and I "have" to clean more. Clutter, dirtiness makes my anxiety worse	YES
So Tired to live in this world can I rest?	YES
Why I want to end my life. Is this a joke??	YES
Love some ink that tells a story...	NO
I feel so hopeless. I just want to quit life. I hate it all so much and I hate my mind. Being okay for awhile and then ready to give up. I'm so sick of the pain and bullshit that I keep enduring. I just want to be happy, what's so wrong with that...?	NO
True Motivation for Success in Life	NO

I still love you... I still need you... I miss you... I still want you...	NO
---	----

Temporal process:

Typically, the phrase "temporal process" is used to communicate information about the past, present, and future target groups of the method by which individuals are discussing one another and the frequency with which they are sensitive to one another. Procedures in linguistics: The LIWC psycholinguistic lexicon collection includes a significant amount of sections devoted to linguistic process. It is recommended that LIWC be used for the calculation of the application of words in the classification of things that are internally significant. In the same vein, this has been effectively employed for the purpose of recognising connections between individuals in the context of community-based contributions, in combination with relative condition, craftiness, and the fundamental nature of personal relationships. This is the reason why we make use of nine explicit language qualities in this study. These attributes are articles, prepositions, auxiliary verbs, adverbs, conjunctions, personal pronouns, impersonal pronouns, verbs, and negations. These characteristics are implemented by us in order to facilitate the identification of user comments for the radical survey.

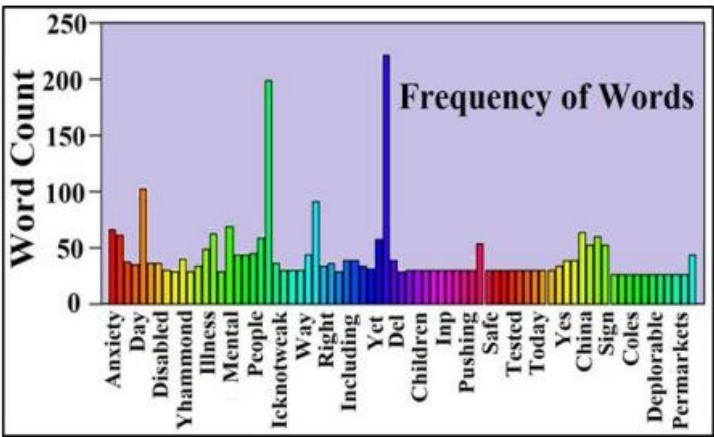


Figure 1: Frequency of depressive words in tweets

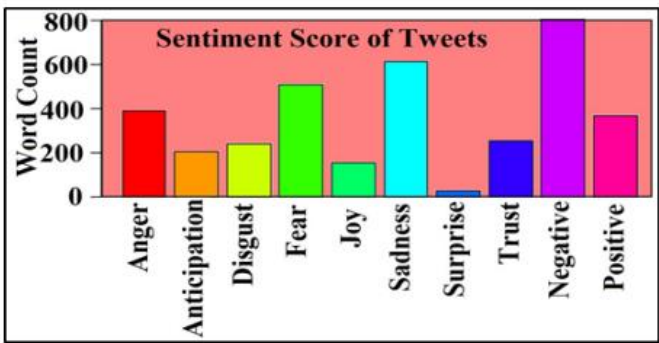


Figure 2: Sentiment score of tweets.

CONCLUSION

The online community network is a collection of separate or set of networks that share users' views and ideas. Through these platforms, users have the opportunity to meet new friends and join new online community groups by sharing their thoughts and ideas. Through the usage of online community networks, people are able to develop new acquaintances with one another. Both the way individuals work and the way they interact with one another in person have been altered as a result of the rise of social networking. In spite of the enormous danger posed by unfettered users, it is of the utmost importance in the business world; in recent years, an increasing number of hackers have been drawn to it. Today, a significant number of people make use of various online community networking sites, including Facebook, Twitter, Instagram, and others. Websites that facilitate online community networking provide the most recent technological advancements in terms of connectivity and communication, but they also present hurdles in terms of privacy and safety concerns.

REFERENCES

- [1] Devlin J, Chang MW, Lee K, Toutanova K. 2018. Bert: pre-training of deep bidirectional transformers for language understanding. ArXiv preprint DOI 10.48550/arXiv.1810.04805.
- [2] Di Cara NH, Maggio V, Davis OS, Haworth CM. 2023. Methodologies for monitoring mental health on Twitter: systematic review. *Journal of Medical Internet Research* 25:e42734 DOI 10.2196/42734.
- [3] Ekman P. 2004. What we become emotional about. In: Manstead ASR, Frijda N, Fischer A, eds. *Feelings and Emotions: The Amsterdam Symposium*. New York: Oxford University Press, 119–135.
- [4] Elbattah M, Arnaud É, Gignon M, Dequen G. 2021. The role of text analytics in healthcare: a review of recent developments and applications. In: *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021)—Volume 5: HEALTHINF*. 825–832 DOI 10.5220/0010414508250832.
- [5] Fleming MN. 2021. Considerations for the ethical implementation of psychological assessment through social media via machine learning. *Ethics & Behavior* 31(3):181–192 DOI 10.1080/10508422.2020.1817026.
- [6] Fodeh S, Li T, Menczynski K, Burgette T, Harris A, Ilita G, Rao S, Gemmell J, Raicu D. 2019. Using machine learning algorithms to detect suicide risk factors on Twitter. In: *IEEE International Conference on Data Mining Workshops, ICDMW 2019-November*. Piscataway: IEEE, 941–948 DOI 10.1109/ICDMW.2019.00137.
- [7] Francis DB. 2021. Twitter is really therapeutic at times: examination of Black men's Twitter conversations following hip-hop artist Kid Cudi's depression disclosure. *Health Communication* 36(4):448–456 DOI 10.1080/10410236.2019.1700436.
- [8] Fredrickson BL. 2013. Positive emotions broaden and build. In *advances in experimental social psychology*. Academic Press 47(6):1–53 DOI 10.1016/B978-0-12-407236-7.00001-2. GitHub.

2024. GCOdeveloper/mental-health-dataset. Available at <https://github.com/GCOdeveloper/Mental-Health-Dataset> (accessed 27 October 2023).
- [9] Google. 2024. One Click Comment Extractor for IG. Available at <https://chromewebstore.google.com/detail/one-click-comment-extract/cckachhlpdnnmhlhaepfcmhahdmpbgp?hl=es&pli=1> (accessed 18 October 2023).
- [10] Gronholm PC, Thornicroft G. 2022. Impact of celebrity disclosure on mental health-related stigma. *Epidemiology and Psychiatric Sciences* 31:e62 DOI 10.1017/S2045796022000488.
- [11] Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC. 2017. Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences* 18:43–49 DOI 10.1016/j.cobeha.2017.07.005.
- [12] Hallgren KA. 2012. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology* 8(1):23–34 DOI 10.20982/tqmp.08.1.p023.
- [13] Hasib KM, Islam MR, Sakib S, Akbar MA, Razzak I, Alam MS. 2023. Depression detection from social networks data based on machine learning and deep learning techniques: an interrogative survey. *IEEE Transactions on Computational Social Systems* 10(4):1568–1586 DOI 10.1109/TCSS.2023.3263128.
- [14] Iyortsuun NK, Kim SH, Jhon M, Yang HJ, Pant S. 2023. A review of machine learning and deep learning approaches on mental health diagnosis. *Healthcare* 11(3):285 MDPI DOI 10.3390/healthcare11030285.