ff ff

5-4-2025 Shubanshu.docx

My Files

My Files

University

Document Details

Submission ID

trn:oid:::17268:89787930

Submission Date

Apr 5, 2025, 7:28 PM GMT+5:30

Download Date

Apr 5, 2025, 7:30 PM GMT+5:30

File Name

5-4-2025 Shubanshu.docx

File Size

268.8 KB

6 Pages

3,043 Words

18,725 Characters

0% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Detection Groups

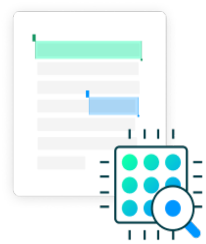
1 AI-generated only 0%
Likely AI-generated text from a large-language model.

2 AI-generated text that was AI-paraphrased 0%
Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer
Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?
The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.



False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

Leveraging Artificial Intelligence for Social Media Analytics in Public Health Surveillance

Shubhanshu Pandey
AIT-CSE
Chandigarh University
Mohali, Punjab, India
pandeyshubhanshu12@gmail.com

Jatin
AIT-CSE
Chandigarh University
Mohali, Punjab, India
sharmajatin10064@gmail.com

Sushil
AIT-CSE
Chandigarh University
Mohali, Punjab, India
21BCS4228@cuchd.in

Shameem Ahmad
AIT-CSE
Chandigarh University
Mohali, Punjab, India
21BCS3807@cuchd.in

Vijay Mohan Shrimal
Department of AIT-CSE
Chandigarh University
Mohali, Punjab, India
drvijaymshrimal@gmail.com

***Abstract* — The study of social media data has significantly increased the use of AI in public health research. Social media environments, where millions of users provide health content in real time, are crucial for monitoring public health issues, disease outbreaks, mental health trends, and attitudes toward vaccines. From a natural language processing and machine learning standpoint, this work explores AI-driven methods that enable the extraction of pertinent patterns from large amounts of unstructured social media data. Techniques such as sentiment analysis, topic modelling, and even deep learning can be used to spot new health risks and trends in populations. Moreover, AI-powered models forecast epidemiological patterns, guaranteeing that decision-makers and healthcare professionals act sooner. However, to guarantee that the analyses are correct and moral, some concerns about bias, disinformation, and data privacy in AI systems must be addressed. Using social media analytics for crisis management, public health surveillance, and evidence-based policymaking, this study seeks to demonstrate AI's capabilities.**

Keywords- Social media usage, Mental health, Emotional well-being, AI analysis, depression detection.

I. INTRODUCTION

Historically, public health surveillance conducted by health agencies and research organizations has relied on epidemiological studies, clinical surveillance systems, and population surveys to investigate disease outbreaks and to assess health trends.[1] Although these types of studies and sources of information are still useful, social media has rapidly emerged as a new, worldwide facility to communicate.[2] Today, Twitter, Facebook, and Reddit offer a unique new dimension to health monitoring, as millions of users frequently post to express their symptoms, health experiences, or health concerns.[3] This information is generated by users who are posting their health and caring experiences based on their own values, beliefs, and perspective. This user-generated content offers a new opportunity for researchers to track the experience of health and public health trends sooner and in more extensive detail than what has historically been possible.[4,5]

The growing power of AI, with NLP and ML, then allows this huge power of analyzing social networking sites millions of users generating billions of unstructured data every hour.[6] These technologies offer a means for the automated analysis of large-scale public content to identify health-related mental issues quickly, track disease outbreaks, survey public sentiment around vaccines, and ultimately identify newly emerging health trends and topics. Thus, for example, sentiment analysis programming can facilitate interpreting attitudes of the public surrounding their anxieties and concerns (e.g., vaccine safety and hesitation) concerning health-related issues.[7,8] Among all health-related topics discussed, healthcare policy is consistently at the forefront of the conversation.

The AI social media content analytics is especially important for detecting diseases and predicting potential epidemics. Normal methods of epidemiological detection often time-lags due to reporting, but social media platforms provide near-immediate access to symptoms and public health

concerns as they occur. AI developments based on prior historical health data will be indispensable for providing public health decision-making.[9] In this way, health policy makers and administrators would have the opportunity to take timely and preemptive steps.
Even so, there are several deterrents to a broad application of AI in public health surveillance. The use of publicly available social media content for study raises questions of data privacy and ethics. Furthermore, misinformation and bias within AI models may lead to erroneous outcomes, which make strong validation methods a requisite. Addressing these challenges is vital to building trust in AI-based health insights.[10]

II. LITERATURE REVIEW

Numerous studies have investigated using artificial intelligence (AI) to examine social media data to gain insights on public health. The literature review explores the technique used by such studies and compares their limitations and effectiveness to achieve this goal. To measure vaccination reluctance, for example, NLP-based sentiment analysis has proven to be quite helpful. However, it may not work as well when contextual complexity is present. Although deep learning has greatly increased the accuracy of disinformation detection, it has come at the cost of enormous computer resources.

This paper explored multimodal AI methods that integrate audio, image, and video data to predict mental health states from social media material. The authors call for larger datasets based on precise diagnostic dates to increase prediction accuracy, particularly for severe disorders like schizophrenia. They proposed the idea that asking users for their social network information with their consent could allay privacy concerns in their discussion of ethical difficulties.[11]

How AI and machine learning (ML) can enhance the utility of social media in public health initiatives. The authors discuss the overwhelming volume of data generated on social platforms and propose AI and ML as solutions to manage this information effectively, thereby improving telehealth services, remote patient monitoring, and overall community well-being.[12]

It creates and implements an AI-based method to examine public opinions about COVID-19 vaccines on social media sites in the US and the UK. By better understanding public views and concerns, the study hopes to offer insightful information for health communication tactics. The development and implementation of a COVID-19 vaccination is progressing quickly on a global scale. The public must cooperate significantly to administer vaccines widely, which is necessary to create herd immunity.[13]

This study explores the requirements of users using social media analysis tools for public health initiatives. To determine these demands and provide information for the creation of more potent instruments for public health monitoring and intervention, the authors carry out a cross-sectional study. "This study addresses a difficult issue: how to relate the information flow across social media and into the realm of health outcomes.[14]

By providing preliminary insights about enhancing surveillance in public health practice, this qualitative study explores specialist views on the likely influence of AI on public health practice. The authors discuss the cautious optimism regarding the synergetic incorporation of AI technology in public health systems. The successful uptake of AI requires initiatives to bolster public health knowledge of AI and increase financing for public health. Public health policy innovations should increase access, integration, and standardization of relevant high-quality data.[15]

Given the research on public perceptions of AI, the authors examine the various public concerns, attitudes, and strong beliefs that back up the claim that AI can take the position of human physicians in the healthcare industry. The researchers intend to use content analysis based on social media data with a few chosen keyword searches to investigate public opinions on artificial intelligence (AI) in healthcare, including concerns raised by the public, attitudes toward AI in healthcare, and the rationale behind these opinions, as well as whether AI can take the place of human doctors.[16]

In this paper, the ALEX framework is proposed to improve the performance of large language models (LLMs) in public health studies by addressing data imbalance in social media datasets. To address data imbalance, the authors present an augmentation pipeline. To enhance model interpretability, they present an LLM explanation method. In public health tasks, their method performs better than others.[17]

With a particular focus on COVID-19-related tweets, this work creates a deep learning-based system to speed up selecting colloquial medical dictionaries from social media data. The authors propose a pipeline that includes named entity recognition, entity normalization, and entity mapping to standard medical concepts to support useful information retrieval in public health research.[18]

Breast cancer (BC) is the most prevalent invasive cancer among women worldwide, and BC medications are known to have significant side effects, yet large-scale quantitative data regarding these effects is limited and underutilized. Using a multi-layer rule-based model to create a vocabulary of medication side effects, they show that social media does hold potential as a health-related sensor in that they were able to infer patterns of medication usage and associated side effects in breast cancer patients.[19]

Together, this research shows how AI may be used in a variety of ways to analyses social media information to improve disease monitoring, public health surveillance, and comprehension of how the general population views health-related topics.

III. METHODOLOGY

This section outlines the study's methodology, which covers data collecting, preprocessing, analytical techniques, and assessment frameworks, to glean insightful information from social media data for public health analysis.

Start – The process begins with initiating AI-based public health surveillance using social media data.

Data Gathering – Social media channels (Twitter, Facebook, Reddit, etc.) create large amounts of user-generated data, such as discussions, symptoms, and opinions relating to health. This step involves collecting relevant posts, comments, and interactions for analysis.

Preparing Data - The raw data collected from the social media site will often be noisy and unstructured. This part consists of cleaning up the text data by removing stop words, emojis, irrelevant data, language processing, and making the data usable for subsequent analyses.

AI & Machine Learning - Text data can reveal relevant information based on advanced ML and natural language processing (NLP) algorithms. AI models can conduct text classification; attitude analysis (positive, negative, or neutral) identification; identification of new health problems; and tracking public reaction to health policy changes.

Trend and network analysis - At this stage, the main objective is to identify trends in public health, but this may include identifying trends in misinformation, or vaccine hesitancy, or new outbreaks of disease. Network analysis assists in understanding how information spreads through social media.

Visualizing & Reporting - The data from the analysis will be summarized in the form of reports, dashboards, and visual charts to display and communicate results. This information can help public health officials and decision-makers in making decisions based on the evidence.

End - The process allows for insight to action, which can lead to prevention, improved health care policies, and improved disease surveillance.

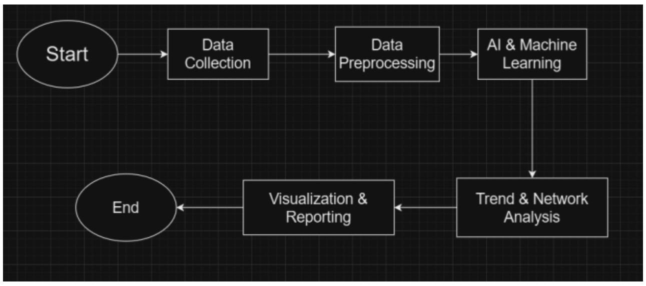


Fig:1 Process Flowchart

IV. DATASET OVERVIEW

The dataset consists of 481 survey samples responses exploring the relationship among mental health and social media use. Coupled with demographic predictors (e.g., age, gender, relationship status, and profession); social media habits (e.g, platforms of decision-making, daily time spent using social media, habitual mindless usage); and mental health outcomes (e.g., social media self-comparison, social media validation-seeking behavior, depressive symptoms, disturbance to sleep, and distraction).

Most of the responses provided participants with Likert scale ratings (1-5) that allow opportunities for quantitatively examining behavioral and emotional trends. This dataset could be enhanced to allow for AI-based sentiment analysis, trend spot detection, and public health monitoring to inform public health officials and others about the role that social media plays in influencing mental health fog.

Key Columns & Their Meanings:

1. Timestamp – Refers to the date and time of survey response submission.
2. Age (1. What is your age?) – The respondent's age.
3. Gender – The gender of the respondent (e.g., male, female, etc.).
4. Relationship Status – If the respondent identifies as single, in a relationship, etc.
5. Occupation Status – The respondent's employment or education.
6. Organization Affiliation – The organization the respondent is affiliated with (e.g., university, job).
7. Social Media Usage – If the respondent uses social media, yes or no.
8. Common Social Media Platforms Used – Identifies the platforms the respondent utilizes regularly (e.g., Facebook, Twitter, Instagram).
9. Daily Time on social media – Breaking up response categories like "more than 5," or "between two and three."
10. Social Media Behavior & Psychological State

- Usage Without Purpose: Frequency of aimlessly using social media.
- Distraction Level: How frequently is social media a distraction for respondents.
- Restlessness Without social media: Measures dependency on social media.
- Self-Comparison: How often a respondent compares themselves to others on social media.
- Validation Seeking: How often a respondent seeks approval on social media.
- Mental Health Indicators: Feelings of depression, distraction, sleep disruption, and fluctuating interest in daily tasks.

V. RESULT ANALYSIS

The heatmap indicates relationships between behaviour's related to social media usage and mental health issues. Indicators of restlessness resulting from not being on social media correspond positively to distraction with respect to social media use (0.61). The correlations for self-comparison and validation-seeking behaviours correspond to lower positive correlations (0.51), showing that, at least in nonclinical settings, online engagement effects human's emotional state. However, feeling down significantly correlates with symptoms of sleeping issues (0.52), a different affective state indicative of possible psychological distress. Furthermore, correlating concentration difficulties such as restlessness were strongly correlated to distractions (0.56), which are relevant not to just social media but to many digital engagements claims about attention spans. The correlations indicated use without purpose or intention yielded lower direct impact. Collectively these correlations are indicating to inform AI models about mental health risk assessment predictions based on behaviour's related to social media usage, who could then effectively demonstrate to inform actionable intervention early in the development of clinical readiness.

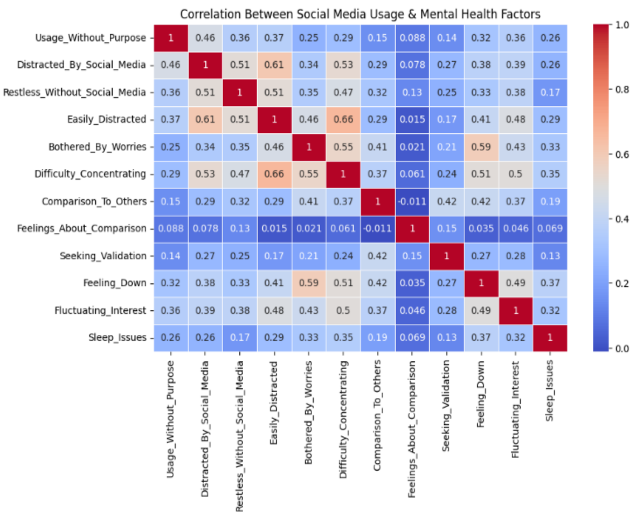


Fig – 2: Correlation between social mediausage and mental health factors

Submission ID trn:oid:::17268:89787930

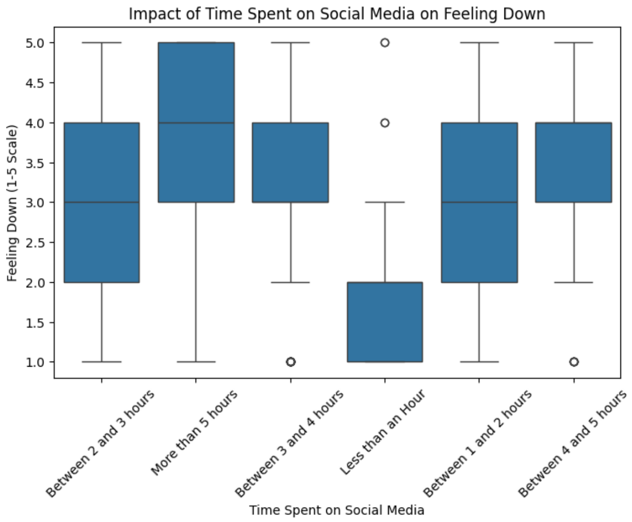


Fig – 3: Impact of time spent on social media on feeling down

The box plot depicts the connection between hours spent each week on social media and the likelihood of feeling down (reported on a scale of 1-5). Social media users reporting >5 hours per week are associated with a higher median feeling down score and greater variability. Users reporting <1 hour of social media usage had the lowest median feeling down score, which can suggest less emotional distress, as their social media use (and associated emotional distress) were much lower. Therefore, those who spent more time on social media had a higher association with mental/emotional distress and more firmly supports what we know about the association between social media and mental

health issues. Outliers do range quite widely in that some individuals indicated severe reductions in mood in both duration groups implying extreme mood drops can occur regardless of social media usage duration. This knowledge helps support the potential for AI to monitor mood and mental health-related constructs and identify individuals for early interventions.

VI. CONCLUSION & FUTURE SCOPE

AI-based social media analysis for public health has a bright future because to developments in real-time analytics, deep learning, disinformation detection, and personalized treatment. Global health crisis monitoring, prediction, and mitigation will be made easier for the public health sector by combining AI with cutting-edge technologies like blockchain, IoT, and reinforcement learning.

A revolutionary method to public health monitoring is provided by AI-based social media data analysis, which provides real-time insights into trends in mental health, disease outbreaks, and disinformation patterns. Public health analytics will become even more accurate and secure with the integration of AI with deep learning, IoT, and
blockchain. For AI-driven insights to be fair and reliable, however, ethical issues, data protection, and bias reduction must be given top priority.

REFERENCE

1. AI for Analyzing Mental Health Disorders Among Social Media Users: Quarter-Century Narrative Review of Progress and Challenges [J Med Internet Res 2024 Nov 15:26:e59225. doi: 10.2196/59225.]

2. Using AI–ML to Augment the Capabilities of Social Media for Telehealth and Remote Patient Monitoring[Healthcare(Basel)2023Jun 10;11(12):1704.doi: 10.3390/healthcare11121704]

3. Artificial Intelligence–Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study[Published on 05.04.2021 in Vol 23, No 4 (2021): April]

4. Leveraging AI to connect social media discourse with public health trends[Ding, X., et al. (2024). Leveraging Prompt-Based Large Language Models: Predicting Pandemic Health Decisions and Outcomes Through Social Media Language. CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems. doi.org/10.1145/3613904.3642117.]

5. Informing social media analysis for public health: a cross-sectional survey of professionals[Published: 02 January 2024].

6. Public Perception of Artificial Intelligence in Medical Care: Content Analysis of Social Media [Published on 13.07.2020 in Vol 22, No 7 (2020): July]

7. Balanced and Explainable Social Media Analysis for Public Health with Large Language Models[Submitted on 12 Sep 2023].

8.Streamlining Social Media Information Retrieval for COVID-19 Research with Deep Learning [Submitted on 28 Jun 2023 (v1), last revised 18 Mar 2024 (this version, v3)]

9. Social Media as a Sensor: Analyzing Twitter Data for Breast Cancer Medication Effects Using Natural Language Processing [Submitted on 26 Feb 2024]

10. AI's gonna have an impact on everything in society, so it has to have an impact on public health": a fundamental qualitative descriptive study of the implications of artificial intelligence for public health[Published: 06 January 2021]

11. Al-Ghadir AI, Azmi AM, Hussain A. A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments. Information Fusion 2021 Mar;67:29-40. [doi: 10.1016/j.inffus.2020.10.003]

12. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: Cross-country survey study. JMIR Mhealth Uhealth

2020 Aug 28;8(8):e19857 [FREE Full text] [doi:10.2196/19857] [Medline: 32759102]

13. COVID-19 AI-powered dashboard for sentiment and opinion mining in social media platforms. COVID Tracker Cloud.:Edinburgh Napier University URL: http://covidtracker.cloud [accessed 2020-11-09]

14. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of UK and US public attitudes on Facebook and Twitter towards COVID-19 vaccinations. J Med Internet Res 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: 10.2196/26627] [Medline: 33724919]

15. Yoo BH, Kim J, Lee B, Hoogenboom G, Kim KS. A surrogate weighted mean ensemble method to reduce the uncertainty at a regional scale for the calculation of potential evapotranspiration. Sci Rep 2020 Jan 21;10(1):870 [FREE Full text] [doi: 10.1038/s41598-020-57466-0][Medline: 31964919]

16. Hutto CJ, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media. 2014 May 16 Presented at: Eighth International AAAI Conference on Weblogs and Social Media; June 1-4, 2014; Ann Arbor, MI URL: http://comp.social.gatech.edu/papers/ icwsm14.vader.hutto.pdf

17. Loria S, Keen P, Honnibal M, Yankovsky R, Karesh D, Dempsey E. TextBlob: simplified text processing. Secondary TextBlob: Simplified Text Processing. 2014. URL: [accessed 2020-08-10]

18. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. : Association for Computational Linguistics; 2019 Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1(Long and Short Papers); June 2019; Minneapolis, MN p. 4171-4186 URL: https://www.aclweb.org/anthology/N19-1423 [doi: 10.18653/v1/N19-1423]

19. Edinburgh Napier University - covid-app: Data and code for using artificial intelligence-enabled analysis of social media data to understand public perceptions of COVID-19 contact tracing apps. GitLab. URL: https://gitlab.com/edinburgh-napier-university/covid-app [accessed 2021-05-04]