International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Health Misinformation on Social Media: A Review Analysis

Shagufta Khan [a], Dr. Rupal Snehkunj [b]

[1]*Ph.D. Research Scholar, Department Of Computer Science, Sarvajanik University, Surat – 395001, India*

[2]*Assistant Professor, Department Of Computer Science, Sarvajanik University, Surat – 395001, India*

A B S T R A C T

The dissemination of Health misinformation on social media is increasing day by day. With the advent of the internet, individuals these days have the liberty to post any thought that generates in their minds. At times, a lot of social media users tend to post misinformation, intentionally or unintentionally. Whatever the intention of posting is, the matter of concern is that such information when goes unnoticed becomes a significant threat to the individual's life and ultimately to society. At times, the loss is substantial and irreversible. The impact of fabricated medical information surged during the Covid-19 pandemic and created havoc around the globe. Individuals, even those from non-medical backgrounds, would share cures for the virus on their social media accounts. What's more shocking is that such cures were embraced by many individuals. Over the past few years, medical information on social media has attracted the attention of many researchers, and hence tools and techniques have been developed by many researchers to alleviate the impact of online health misinformation**.** The present review aims to find out what datasets are accessible and what methodologies researchers have used so far, to limit the spread of medical misinformation on social media. Further, gaps in the existing research have been identified which can help new researchers to explore new pathways to help society fall from the trap of medical misinformation.

Keywords: Health misinformation, social media, Fake health information, health misinformation datasets, Machine learning, Natural language processing, Deep learning

1. Introduction

In today's world, social media has evolved into one of the most powerful tools for quicker dissemination of any piece of information to a wide range of individuals around the globe, that too, within no time. As good as it sounds, it has downsides, too.

With each passing day, we are becoming more and more reliant and dependent on online sources especially, "Social Media Tools", for entertainment, communicating, forming relationships, and even making decisions. Sharing and exchanging health information is no exception. What if, on the other hand, the information is deceptive? To begin with, mobile computing technologies have given birth to a plethora of social media mobile applications. As a result, individuals in large numbers are joining social media. A study reveals that currently, more than 3.2 billion users across the globe are leveraging the services of social media apps. Among the apps, Twitter, Facebook, WhatsApp, Instagram, and YouTube top the list and engage users daily in unimaginable numbers. Further, the users visit these apps several times a day and spend a healthy amount of time (in hours per day) (Gu & Hong, 2020). What's more, the number of individuals seeking health-related information on the web is increasing at an alarming rate. It would not be wrong to quote that online media is currently considered one of the most important sources when it comes to seeking medical advice. According to research, approximately 81.5% of the population of the U.S. searches for medical or health-related information online. Moreover, 68.9% of adults in the U.S. prefer browsing for seeking health-related information before consulting an expert (Dai, Sun, & Wang, 2020). The matter of concern here is the information which appears on the web comes from sources that are not always legitimate and hence, such information when consumed by someone can have dreadful consequences. Not every individual on social media is a medical professional, and not every individual is concerned about the veracity of the information he is consuming on the web. Many such reasons are the answers to why a user fall into the trap of False information.

1.1 What is Health misinformation?

"Health misinformation can be defined as a claim of fact regarding health that is currently inaccurate owing to the lack of scientific data" (Chou, Oh, &Klein, 2018). Not to mention, social media platforms have made it quite easier to disseminate a piece of information, (be it legitimate or not) within no time. Individuals these days are highly influenced by social media and hence, a majority of them believe what they read. Further, 'Share what you read' has become a notion. While it may not seem to be dangerous to many, researchers suggest otherwise. Individuals when consuming ill information relevant to the medical domain does serious harm to themselves and others as well.

"Infodemic", term gained a lot of popularity during the Coronavirus health crisis (Fernández-Torres, Almansa-Martínez, & Chamizo-Sánchez, 2021). As per WHO, "An Infodemic is too much information including false or misleading information in digital and physical environments during a disease outbreak". To help stop the Infodemic, WHO united with the government of the United Kingdom to combat the misinformation related to Covid 19 (Bradd, 2020 & World Health Organization, 2021). The overabundance of information made it quite difficult for people to seek reliable resources during times of need. According to research conducted by Ofcom, the UK's communication regulator, over half of UK internet adults encountered inaccurate or misleading material on the Coronavirus (Covid- 19), just in a week. They further stated that almost 40% of the people were confused about what is factual and what is misleading concerning the virus (Ofcom, 2022). A few misconceptions cleared by them are:

- Drinking more water can flush out the infection (seen by 33% of adults online) (ofCom, 2022).
- It can be treated by gargling with salt water or avoiding cold food and beverages (seen by 24% of adults online) (ofCom, 2022).

Unfortunately, the diffusion of fake health information has seen a surge, especially after the "COVID-19" pandemic. It could pose a serious threat to individuals and ultimately to society. As a result, it necessitates prompt attention. Even though this issue has attracted the eyes of many health organizations, governments, medical institutions as well as individual users, only a few attempts have been made to counteract the issue effectively, and a lot is yet to be explored.

2. Review

For the reviewing process, a systematic approach has been adopted. To commence with, standard databases like Springer, Scopus, Web of Science, ACM, IEEE, and other sources were searched to access articles, Journals, and other relevant works. The search was conducted by using phrases like "Health misinformation AND Social media", "Fake medical information detection approaches", and keywords like "Fake health information, Natural language processing, Deep learning, machine learning". Research articles relevant to the medical misinformation domain, and within the period of the year, 2017 to 2022 were selected for review. The purpose of the review was to identify the available datasets and the approaches that can be utilized by the new researchers to prevent the dissemination of medical misinformation on social media. For the mentioned purpose, datasets, and detection/fighting approaches have been briefly discussed in this section.

2.1 Health misinformation Datasets

For any research to be conducted, data is incredibly significant. However, despite the plethora of information related to health available on social media, as compared to other domains, datasets for fake medical information are in lesser numbers. Recently, a few researchers have not only created datasets after collecting data from several social networking sites but also released them for other researchers to proceed with their research work. Some of them are discussed below:

a. CoAID: Covid-19 heAlthcare mIsinformation Dataset

CoAID is a healthcare misinformation dataset related to Covid19 that includes 5216 news, and 958 social platforms posts corresponding to 296,752 users' engagements from December 1, 2019, to September 1, 2020. It includes automatic annotations for tweets, responses, and Covid 19 false information claims. Topics that were included are Coronavirus, Covid 19, flu9, Pneumonia, lockdown, quarantine, stay home, and ventilator.
The dataset includes both correct and incorrect information related to Covid 19. To collect articles stating correct information, the authors selected nine reliable media outlets. Similarly, they collected URLs from a few fact-checking websites to collect incorrect information. Later, they categorized and labeled false and true claims. In addition to this, the dataset also contains includes data on user engagements (Tweets and replies). This was done using Twitter API. Next, they collected both, fake and true social media posts from platforms like Twitter, Facebook, YouTube, Instagram, and TikTok (Cui & Lee, 2020). The attributes of the dataset contents differ which are listed in Table 1.

Table 1 - CoAID dataset attribute description

Sr.no.	Dataset content	Attributes
1	Fact and misinformation on websites	ID, title, URL (information and fact-checking), article title, content, publish date, keywords
2	User engagements	For tweets: Id, Tweet Id, For replies: Id, Tweet Id, and Reply Id
3	Social Media platform posts	Id, URL (Post and fact-checking), title

b. Labeled dataset for medical misinformation

The authors of the work (Kinsora, Barron, Mei, & Vydiswaran, 2017) made use of techniques like information retrieval, coding, and labelling for creating a labeled dataset. The dataset was created after collecting both misinformative and non-misinformative comments from MedHelp, an online health forum. Nine features were considered, and the Recursive Feature Elimination technique was implemented to classify non-misinformative and misinformative comments. They achieved an accuracy of 90.1% using the Random Forest classifier. In addition to this, they even presented a feature analysis of their labeled dataset. It can be useful in building automated approaches for detecting misinformative posts on online health forums.

c. Fake health

A few researchers constructed a comprehensive data repository; Fake Health, which contains two datasets, namely HealthStory and HealthRelease. The repository contains information on various health topics. The two datasets; HealthStory and HealthRelease contain four sets of information: news content, news reviews, user networks, and social engagements. Each of them has multiple attributes which are listed in Table 2. They also evaluated the performance of their repository using several classification algorithms. The dataset uses ten criteria to review a post and can be utilized in developing several Health fake information detection applications based on an explainable model, knowledge graphs, and multi-modal. Further, it can also aid in developing an application for the early detection of health information Dai, (E., Sun, Y., & Wang, S. ,2020).

Table 2: Attribute description of FakeHealth datasets

Sr.no.	Category	Attributes
1	News reviews	Title, Category, Summary of the review, Images, Description, News Sources, News rating, ground truth, labels of ten criteria, explanation of criteria judgment.
2	News content	Title, URL, keywords, Image URL, Tags, Author, Publishing Date
3	Social engagements	Tweets, Re-tweets, Replies
4	User Network	Profile, Followers and Followings, Timelines

d. ReCovery

(Zhou, Mulay, Zafarani, & Ferrara, 2020) constructed 'ReCovery', a multi-modal data repository to combat misinformation about Covid 19. Apart from the textual information, it also provides visual, temporal, and network information. Further, it explains the dissemination of news on social media. It contains two files namely, News data and social media data which have data with different attributes which are listed in Table 3.

Table 3: Attribute description of ReCovery datasets

Sr.no.	Files	Attributes
1	News data	News Id, URL, Publisher, Publish Date, Author, Title, Image, Body text, Political Bias, Country, Reliability
2	Social Media Data	News Id, tweet Id

Table 4 summarizes the information of various health misinformation datasets that are mentioned in this paper.

Table 4: Summary of Health misinformation Datasets

Dataset	Details	Author	Access link
CoAID	Covid-19 heAlthcare mIsinformation Dataset	Limeng Cui, Dongwon Lee	**https://github.com/cuilimeng/CoAID**
Labeled dataset for medical misinformation	Medical misinformation dataset	Alexander Kinsor, Kate Barron, Qiaozhu Mei, V.G. Vinod Vydiswaran	**N/A**
FakeHealth	Data repository for fake health information	Enyan Dai, Yiwei Sun, Suhang Wang	**https://github.com/EnyanDai/FakeHealth**
ReCovery	Multimodal repository for Covid 19	Xinyi Zhou, Emilio Ferrara, Apurva Mulay, Reza Zafarani	**https://github.com/apurvamulay/ReCOVery**

2.2 Detection Approaches and Gaps

The open and unlimited access to information on the internet, today, while empowering the knowledge on one side, has downsides, too. To elaborate, not every piece of information that is posted online is monitored and hence, it is likely one has come across unhealthy or fabricated information. In the case of social media, every user is the king of his social media account and hence, is free to post whatever he wants to. From fake cures to myths, social media is full of health information pollution. Health misinformation when it reaches the public domain can cause panic. Not to mention, it also leads to psychological distress among people, especially during the era of the Pandemic. To counteract this issue several researchers have tightened their waist belts to limit the spread of misinformation on social sites. This section throws some light on various approaches that have been developed by researchers to combat medical misinformation on several social media platforms. In addition to this, it also includes research gaps that can be useful for scholars to make a start.

To begin with, (Wang, Yin, & Argyris, 2021) proposed a deep learning model based on the multi-modality concept. They developed a mechanism that processes both textual and visual information to fight against anti-vaccine messages on Instagram. The model, further, for independent extraction of features, consists of three branches (images, caption, and hashtags) which were fused to predict results. Their model leverages techniques like attention mechanisms, neural networks, etc. However, the model fails to deliver accurate predictions under circumstances where human domain knowledge is required. In addition to this, in the absence of sufficient information, the models fail to predict correct results. Their model has an accuracy of 97%. They've made their work available publicly on https://github.com/wzhings/antivaccine_detection.

Some researchers presented a simple Natural Language Processing methodology for detecting Covid-19 related misinformation spread through YouTube videos (Serrano, Papakyriakopoulos, & Hegelich, 2020). For prediction, rather than training the models on the video data, they collected the comments posted by YouTube users on the misinformative videos. Later, features were extracted for detecting misinformation. Comments for their datasets were collected from both factual and non-factual (misinformative) videos. Their dataset, in total, includes 113 misinformative videos with 32,273 comments, and 67 factual videos with 119,294 comments. However, for their work, they selected only ten percent of the total comments in the datasets. The selection was done randomly. Based on transfer learning, they created a multi-label classifier to detect ill comments. Initially, their accuracy was 82.2% which jumped to 89.4% once they included the tf-idf feature in their classifier. Nevertheless, the model-generated results are based on the percentage of the comments that were made on the videos, and not the videos themselves. In addition to this, the research only included the first hundred comments to increase the accuracy.

A group of researchers (Gundapu & Mamidi, 2021) developed a methodology to verify the information relevant to Covid 19 on social media. They relied on ensemble concepts for dealing with medical misinformation. They ensemble three transformer models, namely, XLNET, ALBERT, and BERT to train their model on the Constraint AI 2021 Fake News Detection dataset for evaluating Covid-19 fake news. The f1-score for their system was 0.9855.

A few authors conducted a study for revealing cancer fake cures on Twitter. They adopted a user-centric approach in which they monitored the social media users posting dubious health-related information. For the mentioned purpose, features like user attributes, sentiments, and writing style were taken into consideration to build a classifier for identifying the users who were posting unverified health misinformation. This research, nonetheless, was limited to identifying the users propagating false information (Ghenai & Mejova, 2018).

A study was conducted by a group of scholars, to identify structural, topical, and semantic variations between information about health from credible and unreliable media on Twitter. To elaborate, they used a large-scale database to discriminate the features of reliable information from unreliable ones. They also stated that these features hold extreme importance when it comes to understanding and developing systems to help fight the spread of false medical information. Using machine learning techniques, they built classification models which were capable of predicting the article source with an F1 score of 96%. However, their study has some limitations. For instance, they exclude information like user comments, cited experts, and videos from their work (Dhoju, Rony, Kabir, & Hassan, 2019).

Table 5: Summary of Health Misinformation Detection Approaches

Title	Author (s)	Platform	Topic	Method (s)	Test Accuracy / F1-score
Detecting medical misinformation on social media using Multimodal Deep Learning	Zuhui Wang, Zhaozheng Yin	Instagram	Anti-Vaccine	Deep learning, Attention mechanism, Neural network, Multi-modal	97%
NLP-based feature extraction for the detection of covid-19 misinformation through YouTube videos.	Juan Carlos Medina Serrano Orestis Papakyriakopoulos , Simon Hegelic	YouTube	Covid 19	Natural Language Processing, Machine Learning	89.4%
Transformer based automatic Covid-19 Fake News Detection System	Sunil Gundapu, Radhika Mamidi	Twitte, Facebook, and Instagram	Covid 19	Ensemble model: BERT, ALBERT, and XLNET	98.55% (f1-score)
Fake cures:User-centric modeling of health Misinformation in social media	Amira Ghenai, Yelena Mejova	Twitter	Cancer	Machine learning	Over 90%
Differences in health news from reliable and unreliable media	Sameer Dhoju, Md Main Uddin Rony, Muhammad Ashad Kabir, Naeemul Hassan	Twitter	Cancer	Machine Learning	96% (f1-score)

Conclusion

To understand the depth of the dissemination of health misinformation on social media platforms, a systematic literature review was conducted in the medical domain. Relevant Articles and journals were selected and reviewed. Datasets were listed down which can be the first step toward combating medical misinformation. Further, a literature review has been conducted on the works of the scholars and also, and research gaps have been identified where future researchers can work to contribute their knowledge.

References

- Gu, R., & Hong, Y. K. (2019). Addressing health misinformation dissemination on mobile social media.
- Dai, E., Sun, Y., & Wang, S. (2020, March 30). *Ginger cannot cure cancer: Battling Fake Health News with a comprehensive Data Repository*.
- Chou, W. S., Oh, A., & Klein, W. M. (2018). Addressing health-related misinformation on social media. *JAMA, 320*(23), 2417
- M. J. Fernández-Torres, A. Almansa-Martínez, and R. Chamizo-Sánchez, “Infodemic and fake news in Spain during the COVID-19 pandemic,” International Journal of Environmental Research and Public Health, vol. 18, no. 4, p. 1781, 2021.
- Bradd, S. (n.d.). Infodemic. Retrieved November 19, 20[1222, from https://www.who.int/health-topics/infodemic#tab=tab_1
- World Health Organization. (2021, April 27). Fighting misinformation in the time of covid-19, one click at a time.
- ofCom (2022, February 21). Half of UK adults exposed to false claims about coronavirus.
- Cui, L., & Lee, D. (2020, November 03). CoAID: Covid-19 healthcare misinformation dataset.
- Kinsora, A., Barron, K., Mei, Q., & Vydiswaran, V. (2017). Creating a labeled dataset for medical misinformation in Health Forums.
- Zhou, X., Mulay, A., Zafarani, R., & Ferrara, E. (2020, October 01). Recovery: Proceedings of the 29th ACM International Conference on Information & Knowledge Management.
- Wang, Z., Yin, Z., & Argyris, Y. (2021, June). Detecting medical misinformation on social media using Multimodal Deep Learning.
- Serrano, J., Papakyriakopoulos, O., & Hegelich, S. (2020, June). NLP-based feature extraction for the detection of covid-19 misinformation videos on YouTube. Gundapu, S., & Mamidi, R. (2021, January 21). Transformer based automatic COVID-19 Fake News Detection System.
- Ghenai, A., & Mejova, Y. (2018, November 01). Fake cures: User-centric modeling of health misinformation in social media: Proceedings of the ACM on Human-Computer Interaction: Vol 2, NO CSCW.
- Dhoju, S., Rony, M., Kabir, A., & Hassan, N. (2019). Differences in health news from reliable and unreliable media.