International Journal of Enhanced Research in Educational Development (IJERED)
ISSN: 2320-8708, Vol. 12 Issue 5, Sept.-Oct, 2024 Impact Factor: 8.376

Al-Fueled Graph Neural Network (GNN)-Based Cyber Threat Forecasting

Ambika Kunj[1], Dr. Kushal Pal Singh[2]

[1]Research Scholar, Department of Mathematics, JS University, Shikohabad, Uttar Pradesh
[2]Supervisor, Department of Mathematics, JS University, Shikohabad, Uttar Pradesh

ABSTRACT

Businesses find it challenging to stay up with the latest advancements in internet dangers since these risks are always changing. In order to accurately and quickly detect dangers, modern techniques must be used. Most security solutions rely on rule-based or signature-based approaches, which struggle to detect complex, hidden, or ever-changing threats. This study shows that Graph Neural Networks (GNNs) can describe and evaluate deep links in data from networks, system logs, and patterns of user activity. Graph Neural Networks (GNNs) could be an excellent AI tool for cyber threat forecasting because of this. A node represents a device, person, IP address, or network flow in the GNN-based approach that has been detailed. Interactions between these nodes are shown by edges. This provides the ability to detect structural linkages as well as temporal trends, which are crucial for finding new dangers.

The algorithm uses the GNN's repeating message-passing characteristics and data from edges and nodes to create rich embeddings. Some of these features include the frequency of connections, patterns of user data access, details on previous security vulnerabilities, and behavioral aspects. This allows the system to detect potentially harmful actions, such as passing from one level of access control to another, or acquiring more rights.

Transferring power to another user or initiating DDoS assaults, all of which might be signs of system hacking. Additionally, with attention-based GNN variants, various relevance weights are assigned to the links under consideration. This allows the model to zero down on the most dangerous behaviors while also decreasing the number of false positives.

We test the suggested model using real-world network datasets and safety constraints, and compare it to popular machine learning methods like Support Vector Machines, Random Forests, and regular neural networks. The GNN-based system outperforms the competitors in terms of accuracy, precision, memory, and F1-score when it comes to detecting both new and existing dangers. Based on the results, we can say this. Additionally, the model maintains its effectiveness even when faced with busy or missing data, a common occurrence in real-world computer networks.

The results of this research show that graph-based deep learning is a viable and efficient method for predicting cybersecurity threats. The proposed method enhances discovery and decision-making capabilities by using interpretable embeddings and attention processes that highlight critical attack lines. After these factors are considered, combining AI with GNNs is a great way to build next-gen security systems that can handle complex threats in real-time network environments.

Keywords: Graph Neural Networks (GNNs), Cyber Threat Prediction, Network Security Analytics, Anomaly Detection

1. INTRODUCTION

Because information can be shared quickly on social media and similar platforms, the proliferation of false information is a major problem. Particularly in life-or-death domains like public health, this is the case. Approximately 4.8 billion individuals throughout the world were active on social media in May of 2024. There was a massive increase between January 2021 and January 2022, when the respective populations were 4.2 and 4.62 billion [1]. An indication that social media is still influencing how people throughout the globe communicate with one another and exchange information is its exponential growth. Misleading health information is still circulated widely. A large portion of the

population persists in using the word "misinformation" to describe factually incorrect or biased information that does not align with what physicians typically know.

Typically, the rate of dissemination of this misinformation exceeds that of actual scientific understanding. One common cause of this is the propagation of unproven claims or the lack of conclusive scientific evidence. Now that social media is used by everyone, the transmission and reception of information has changed. Ignorant information may spread more easily due to the accessibility of communication, which is a shame since this is particularly true when it comes to health-related subjects. Issues including public health crises, vaccination hesitancy, and potentially harmful or inefficient therapies could result from the spread of misinformation about health on social media.

Consequently, we must immediately devise efficient means of detecting and reducing the quantity of health-related disinformation propagated on these platforms. To prevent the broad distribution of harmful or incorrect statements, it is critical to identify cases of health misinformation on social media. Data analysis, machine learning, and NLP components are all part of the AI-based approaches that may fix this issue [2]. Some of the many ideas included under the umbrella word "misinformation" are: untrustworthy, spam, gossip, disinformation, information disorder, and many more.

Data that might be detrimental to one's health can take several forms, including uncertainty, contradictory outcomes, and evidence that emerges over time.

Moreover, even while communication is crucial, it is insufficient to address these issues, which include obviously incorrect information [3]. Various forms of misinformation, such as rumors, falsehoods, misleading information, and inaccurate information, are shown in Figure 1.Anybody may post anything on social media without external controls, therefore it could be hard to tell whether a source is credible [4, 5]. everybody may spread false information, and everybody that gets it or helps fix it will have their own personal goals in mind [6].

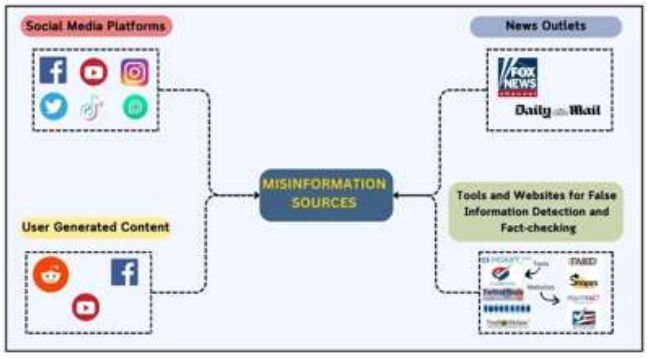


Figure 1.2: Sources of Misinformation

According to the findings of recent scientific investigations, the term "health misinformation" refers to any assertion made about health that is not shown by evidence. The dissemination of health information that is either inaccurate or misleading and that has the potential to put people's health and safety at jeopardy [8]. The graph in figure 1.3 illustrates various instances of health information that is deceptive. An investigation was conducted by the authors [9] into the accounts and tweets that are used by individuals who spread misleading information on cancer treatments.

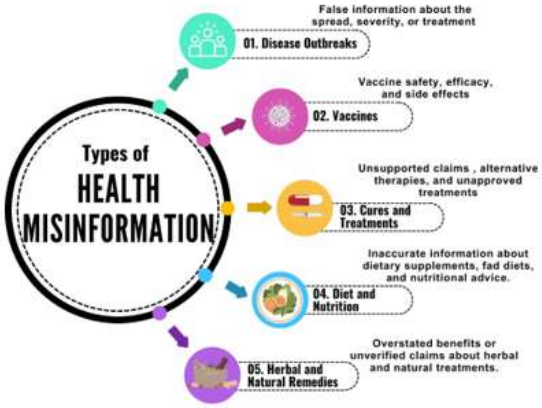


Figure 1.3: Types of Health Misinformation

A. Problems with Health Misinformation Detection

In real time, billions of individuals engage with content and share it with others via online social networks. This phenomena can be scaled very quickly. Controlling and preventing the fast spread of health-related misinformation is challenging since it may become viral in hours.

Retain in reserve. It is becoming increasingly difficult to protect the public's health and safety due to the quick spread of false information, which in turn increases the effect of this information.

The limitations of the detection systems that are now in use include: In order to detect health misinformation, many of the current programs rely on old-fashioned machine learning and rule-based approaches. It is difficult for these approaches to stay up with the ever-changing nature of disinformation. Given their inability to change, it's likely that they will become worse at spotting subtler forms of dishonesty over time. Furthermore, content-based analysis is the norm rather than the exception in many current approaches, which ignore the relational and contextual data present in social media [16].

People may become more likely to engage in harmful habits, make bad decisions, and lose faith in healthcare organizations if they hear misleading information about health. This is why it's critical to act swiftly on measures that effectively uncover and decrease health myths. Graphs, which are mathematical structures made up of nodes and edges, are used by Generalized Convolutional Neural Networks (GCNs), a kind of neural network, to identify health misinformation in textual data. Computer vision and social network research are only two of many fields that have found success using GCNs. Natural language processing (NLP) experts in fields including sentiment analysis, text categorization, and entity identification have recently taken notice of GCNs [17]. mentioned as [18]. The paper delves into a broad range of methodologies employed within graph machine learning on many levels, in addition to offering an overview of sickness prediction applications that use various Graph Machine Learning (GML) algorithms. By doing a thorough literature search and trend analysis, this study offers valuable insights into the possibilities of machine learning in disease prediction utilizing electronic health data[19]. The study's results corroborate these observations. Despite the many benefits of rapid information exchange, there is growing concern over the spread of inaccurate health information [20]. False or inaccurate health information has the ability to cause major harm, such as the promotion of unhealthy behaviors or the spread of incorrect information on medical procedures and treatments. There is an urgent need for effective tools to detect and counter health misinformation since this problem has grown substantially due to the rapid spread of information on social media [21].

II. LITERATURE REVIEW

Due to the widespread availability of inaccurate information on health, it is especially difficult for public health professionals to carry out their responsibilities and make decisions in this digital age. Identifying and preventing fake information is an essential step in the process of providing information that can be relied upon. In natural language processing (NLP) applications such as text categorization, sentiment analysis, and information discovery, there is a great deal of promise for machine learning and deep learning techniques [22]. By using these methods, we are able to search for wrong information by swiftly scanning and analyzing huge amounts of text using data-driven methodology. This allows us to find material that is not accurate. As a result of its capacity to show complicated patterns and relationships, GNNs have gained a significant amount of attention as prospective means for finding deceptive information in situations impacting health. In this literature review [23], we take a look at recent research and developments that make use of GNNs to uncover health myths.

The data determines which of the various AI-based tactics is used in order to get the best possible results. At the moment, there is a significant amount of false information that is being disseminated via the internet and on social media [40]. Due to the volume of information that is both correct and erroneous that is accessible online, it is difficult for users to quickly identify what they need. The process of determining whether or not data is real is referred to as identifying misinformation. Several different machine learning algorithms have been developed by researchers in an effort to detect bogus news; nevertheless, the outcomes of these attempts have been insufficient. RNNs, SVMs, KNNs, LRs, RFs, CNNs, sentiment-aware multimodal embedding, hierarchical attention networks, and dEFEND are some examples of fundamental machine learning models that are used to categorize data in accordance with their principles. Additional examples include CNNs. [41] [41] This is the [42].

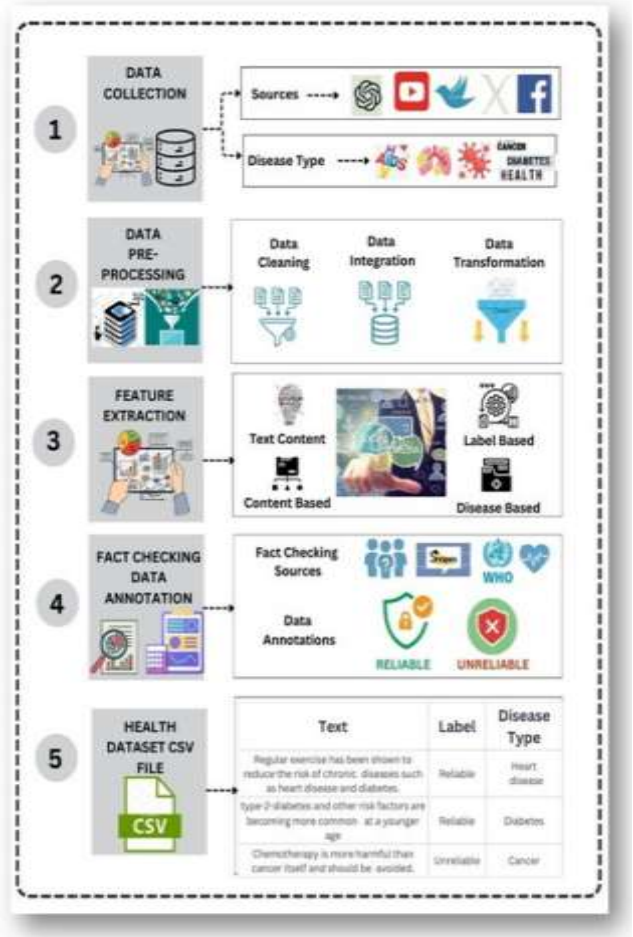
A viewpoint known as the Elaboration Likelihood Model (ELM) is used in order to analyze the two types of health fraud that occur online. These are referred to as the core and secondary. Propaganda, advertising, information that is not useful, and information that is wrong are the four primary types of misleading information that may be discovered in online health forums [43]. For the purpose of developing a model for the finding of health news, data on language, topics, feelings, and behaviors were merged. For the purpose of evaluating the features of the suggested model, a dataset from the actual world was used, and the results revealed that around 85 percent of health lies were accurate [44, 45].

When it comes to discovering health lying in online health communities, behavioral criteria, as opposed to linguistic ones, are more useful [46]. These individuals have developed three reliable methods for identifying stuff that is not true. A number of other models, including KNN, MC-CNN, and LSTM, which is a form of RNN, are now under consideration [37]. For the purpose of carrying out the simulations, the CoAID disinformation dataset was used. This dataset is comprised of a collection of misinformation about healthcare [47]. Along with the Naïve Bayes approach, gradient boosting algorithm, support vector machine, multi-layer perceptron, and decision tree, we use three main types of data. It was found that the SVM had the best level of accuracy among all of them [48]. We went with the LR model since it offers a simple cost function and an equation for categorizing problems into two or more categories. This was a major factor in our decision. A set of new choices for displaying documents is one of the most important changes that has been implemented.

III METHODOLOGY

One major problem with social media is the proliferation of false health claims. Several studies have shown that misinformation about health care has the potential to spread further and faster than the truth. Mistakes related to COVID-19, vaccination claims without evidence, and dubious therapies for several diseases are only a few examples of the major sectors that have been damaged. The massive user bases and easy user-to-user communication features of social media platforms make them perfect for disseminating false information like this. As an illustration: A large quantity of inaccurate information on the COVID-19 pandemic and its treatment was disseminated, according to a World Health Organization research.

This caused a "infodemic" that impeded the ability of public health workers to help those in need. Studies have shown that the spread of misinformation about vaccines on social media sites like Facebook and Twitter has led to fewer individuals being vaccinated or even refusing to get vaccinated at all. The overall population's health has suffered as a result. The healthcare system might end up costing the economy more money if consumers continue to wait for care or seek treatments that are not required because of misleading promises. Exposure to disturbing or false information on a regular basis may have a negative impact on people's mental health by making them anxious, confused, and upset. Vaccination rates are higher among the educated, but those with lower levels of education are more likely to forego vaccines, use therapies without enough research, and fail to take preventative actions. The worst case scenario is that more illnesses spread, leading to more deaths.



A. Feature Extraction Techniques

The Bidirectional Encoder Representations from Transformers (BERT) model is an example of a contextual language model that has the potential to create outputs in a safe manner by taking in information from both different sides. The information has been sent on by both parties. Transformers, as opposed to LSTM networks, are used in order to ascertain the direction of mood data by means of a cutoff. Transformers are able to increase understanding of word context via the use of procedures such as tokenization, filtering, and attention optimization algorithms. Because of this, the model is able to capture sensitive mood data using a more straightforward approach. When it comes to mood categorization, this method performs better than others since it makes use of transformer topologies like BERT or RoBERTa. In order to better manage complex language patterns, these topologies focus on the connections that exist between the words that are included inside a paragraph [24]. Within the study, the LSTM-Gate CNN model is dissected and examined. This technique combines LSTM and GCNN layers in order to improve the perception of context and emotions. By using attention and filtering processes [25], this hybrid model is able to extract more pertinent aspects from the input, hence improving the understanding of the context and the mood. Following the elimination of aspect and emotional components, this model employs a soft-max approach in order to properly categorize mood [26]. A innovative strategy to finding COVID-19 health lies is presented in this research. The approach makes use of ML, DL, and GCN software.

Through the examination of the language used in COVID-19-related tweets and news items, our strategy targets the extraction of meaning as well as information about the surrounding environment. Utilizing a number of different Tokenization, stemming, and feature extraction are examples of natural language processing (NLP) methods that are used in order to convert the textual input into numerical values that may be comprehended by machine learning models. The traditional RNN faces a challenge in the form of gradients that disappear [27]. To find a solution to this problem, the LSTM network, which is an RNN, was constructed. Specifically, they assert that content-based components are a part of the dynamic process that is responsible for the dissemination of incorrect health information on social media.

For the purpose of conducting study, they sifted through the Zika Twitter discourse that took place in 2016, selecting 264 significant messages that were incorrect while retaining 455 tweets that were equivalent and accurate. Through the use of a computer that retweeted each and every tweet, we were able to determine the information distribution network [28] and extract nine network identifiers. After then, an NHPP signal was used in order to manage the transmission of the information. Following that, forty different signal aspects were used in order to define each NHPP. It was language questions, document-to-vector, and word count that we used for the properties that were associated with the material. A total of fifty and sixty-three characteristics were collected from each tweet, respectively, by using these methodologies [29].

B. Methods for Combating False Health Claims Utilizing GNNs

In recent years, the approach of identifying deceptive information may have become more successful and accurate as a result of the coupling of graph-based learning with machine learning and deep learning algorithms [56], [59], [68]. In addition to shedding insight on the present status of GNN implementations in healthcare, the findings of the study also shed light on possible future problems that may arise with these systems. In this paper, a GNN-based text categorization technique as well as an MGTA (multi-granular topic-aware graph) are presented [69]. Using this method, which focuses on several levels of text semantic information, one may create a text graph that has joint relationships at the ego level and lower. The approach that was suggested performed better than three well-known text recognition algorithms on four different benchmark datasets [70]. 1. This diagram (Figure 2.2) illustrates how GNN may be used to identify misinformation, which includes the detection of fake news.

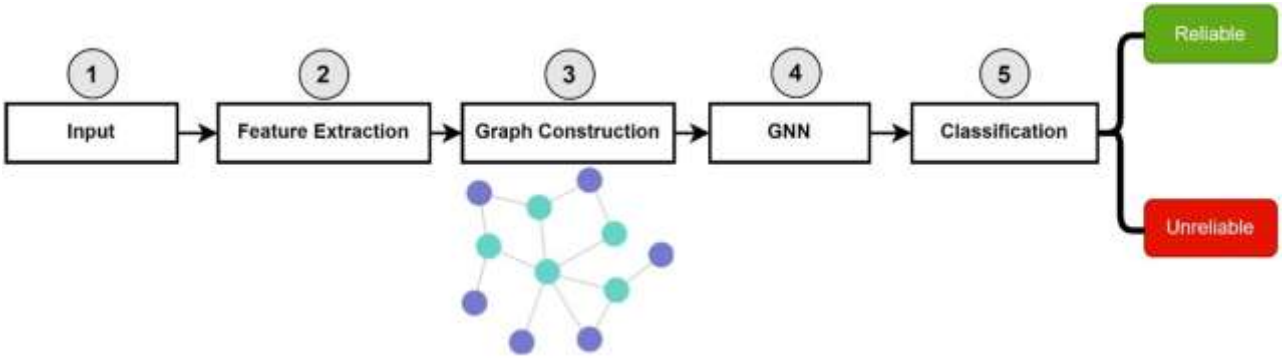


Figure 2.2: Structure for Identifying False Information Using a GNN Method

C.GNN Architecture

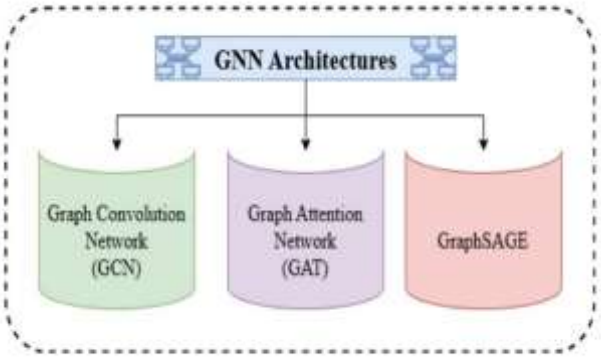


Figure 3.1: GNN Architecture

Patterns of Graph Neural Networks (GNNs) reflect the collection, movement, and change of information in graph-structured data. These patterns are used to generate neural networks. Throughout the years, a number of different designs have been created in order to address many different concerns, including expressiveness, scalability, and adaptability to different kinds of graphs. The manner in which these systems handle different kinds of graphs, in which they prioritize nodes, and in which they collect area data varies.

A well-known and well-established kind of neural network is known as the Graph Convolutional Network (GCN). The notion of convolution is transferred from grid-structured data to graphs via the use of GCNs. This is accomplished by adding together the weights of every node's neighbors, including the node itself. It is via the use of normalized adjacency matrices that they are able to avoid feature growth and keep the learning rate constant. When it comes to semi-supervised node classification tasks, GCNs perform very well, provided that the graph topology remains consistent and the graph size is relatively modest. It is possible that GCNs will have difficulty adequately portraying the surrounding nodes of complex networks since they operate on the premise that all nodes are of similar importance.

C. Graph Convolutional Networks (GCNs)

GCNs are a kind of neural network architecture that are used for the purpose of processing and evaluating data that is structured and organized in a graph. GCNs are well suited for tasks that include graph-structured data, such as those found in social networks, recommendation systems, reference networks, and other applications that are connected to these types of networks. The use of these networks might be a complement to the conventional CNNs. Examples of applications that may benefit from these include those that need data to be organized in a grid, such as photographs. CNNs exclusively use the convolution method on grids, despite the fact that GCNs also use it on graphs using the convolution method. GCNs, or graph convolutional networks, are convolutional algorithms that are used to handle graph data. By taking into consideration the distinctive characteristics of each node in addition to those of its neighbors, it is possible for data to be gathered and distributed across the nodes that make up a network.

In contrast to GNNs, CNNs are especially built to deal with ordered (Euclidean) data, such as grid-based photo collections. This is one of the most important differences between the two types of networks. GNNs, on the other hand, improve CNN's ability to deal with non-Euclidean data, which is often asymmetrical owing to the fact that the ordering of nodes and the kinds of connections involved are not uniform. A generalized convolutional neural network (GCN)

uses many layers of convolution and aggregation to enhance the representation of the graph's nodes. The ability of GCNs to infer complex connections and patterns within the structure of the graph is made possible by the technique of stacking layers.

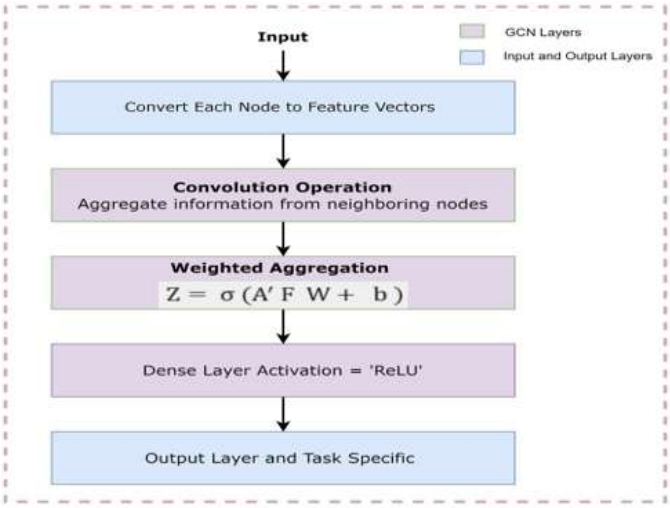


Figure 3.2: Working of GCN

IV DISCUSSION

About the discovery of health information that is not accurate. The primary objectives of this research are to investigate the efficacy of various modeling techniques and to determine the effects that mixed methods have on recognition performance. For each model, a comprehensive analysis of its accuracy, precision, recall, F1-score, and ROC-AUC was presented in the chapter that came before this one. In the sections that are to come, we will discuss the usefulness of graph-based techniques, the operation of mixed approaches, and the differences in performance that exist across different datasets.

A dataset having information about the COVID-19 pandemic and another dataset providing information about the original health are both used in this example to demonstrate how to apply health misinformation detection to two different datasets. There were 15,641 tweets included in the COVID-19 dataset, and only 57% of them were considered reliable, while 43% were considered to be untrustworthy. For the duration of the outbreak, computers are instructed to differentiate between genuine and fabricated material via the use of these tweets. In order to deliver accurate predictions for incoming text inputs, we reviewed the models that were deployed, which included GNN-based techniques, and illustrated how they functioned by using a live Streamlit interface. By comparing our mixed model against leading techniques, which obtained remarkable accuracy rates of 96.16% on the health dataset and 89.23% on the COVID-19 dataset, we were able to illustrate the effectiveness of our mixed model. Additional topics that were discussed in this part included the complexity of time and space, mathematical models, as well as the advantages and disadvantages of each combination strategy. The flexibility and reliability of a number of different techniques, as well as the computing requirements of graph-based models, were explored.

CONCLUSION

The inquiry into the project's goal of developing AI-driven systems to detect health-related disinformation on social media platforms yielded encouraging results. Text feature extraction techniques including TFIDF, Cosine Similarity, and PMI yielded substantially improved outcomes when combined with various graph neural network (GNN) models. This comparative research found that state-of-the-art methods, including Graph Neural Networks, may successfully detect cases of health fraud on social media. The effectiveness of these processes is greatly affected by the process of adjusting the parameters. Based on our research, we determined that GraphSAGE was the best tool currently available.

Tables 4.4 and 4.8 show that compared to deep learning and regular machine learning models, mixed models, especially those using Graph Neural Networks (GNNs) like GraphSAGE (GSAGE), perform much better. All things considered, this is correct. A few machine learning algorithms that excel in accuracy, precision, and F1-scores are Decision Trees (DTs), Random Forests (RFs), and Support Vector Machines (SVMs). For optimal results, use Random Forest. After analyzing ROC-AUC and other performance metrics, it becomes clear that the GNN-based mixed models are the winners. Since CNN and LSTM, two popular deep learning algorithms, have poor accuracy results on the COVID-19 lies dataset, it's clear that these models can't manage the dataset's complexity.

The other hybrid models make up GSAGE, however Hybrid Models 2 and 3 routinely outperform it in terms of performance. They have the best F1 scores and ROC-AUC values, which are almost flawless. These experiments demonstrate that GNNs, when paired with other embedding or similarity approaches, can effectively recognize complex text linkages. This makes them ideal candidates for careers that need them to spot misleading health information. The findings indicate that mixed models significantly improve performance, making them an excellent option for challenging text sorting challenges. This is particularly the case when using GSAGE modeling in conjunction with mixed models.

REFERENCES

[1]. DataReportal, "DataReportal," accessed Aug. 28, 2024. [Online]. Available: https://datareportal.com/.
[2]. D. Boovitha, M. Abirami, S. Gunavathi, N. Revathi, and S. Rubavarshini, "Fake Media Detection Based on Natural Language Processing and Blockchain Approaches," South Asian J. Eng. Technol., vol. 13, no. 1, pp. 69–82, 2023.
[3]. W. Y. S. Chou, A. Gaysynsky, and J. N. Cappella, "Where we go from here: Health misinformation on social media," Am. J. Public Health, vol. 110, pp. S273–S275, 2020, doi: 10.2105/AJPH.2020.305905.
[4]. B. Swire-Thompson and D. Lazer, "Public health and online misinformation: Chal- lenges and recommendations," Annu. Rev. Public Health, vol. 41, pp. 433–451, 2019, doi: 10.1146/annurev-publhealth-040119-094127.
[5]. S.-Y. C. C.-J. Y. &Fang Yu, "XFlag: Explainable Fake News Detec- tion Model on Social Media," Int. J. Hum. Comput. Interact., 2022, doi: https://doi.org/10.1080/10447318.2022.2062113.
[6]. Vraga, Emily K., and Leticia Bode. "Correction as a solution for health misinformation on social media." American Journal of Public Health 110, no. S3 (2020): S278-S280.
[7]. S. Zhang, F. Ma, Y. Liu, and W. Pian, "Identifying features of health misinformation on social media sites: an exploratory analysis," Libr. Hi Tech, no. 71420107026, 2021, doi: 10.1108/LHT-09-2020-0242.
[8]. Suarez-Lledo, Victor, and Javier Alvarez-Galvez. "Prevalence of health misinformation on social media: systematic review." Journal of medical Internet research 23, no. 1 (2021): e17187.
[9]. L. Cui, H. Seo, M. Tabar, F. Ma, S. Wang, and D. Lee, "DETERRENT: Knowl- edge Guided Graph Attention Network for Detecting Healthcare Misinformation," Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 492–502, 2020, doi: 10.1145/3394486.3403092.
[10]. S. G. Paul, A. Saha, M. Z. Hasan, S. R. H. Noori, and A. Moustafa, "A systematic review of graph neural network in healthcare-based applications: recent advances, trends, and future directions," IEEE Access, 2024.
[11]. H. T. Phan, N. T. Nguyen, and D. Hwang, "Fake news detection: A survey of graph neural network methods," Appl. Soft Comput., p. 110235, 2023.
[12]. T. Liu et al., "Rumor Detection with a novel graph neural network approach," arXiv Prepr. arXiv2403.16206, 2024.
[13]. B. Khemani, S. Patil, K. Kotecha, and S. Tanwar, "A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future di- rections," J. Big Data, vol. 11, no. 1, p. 18, 2024.
[14]. J. Zhou et al., "Graph neural networks: A review of methods and applications," AI Open, vol. 1, no. January, pp. 57–81, 2020, doi: 10.1016/j.aiopen.2021.01.001.
[15]. E. Min and S. Ananiadou, "PESTO: A Post-User Fusion Network for Rumour De- tection on Social Media," in Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, 2023, pp. 1–10.
[16]. A. D'Ulizia, M. C. Caschera, F. Ferri, and P. Grifoni, "Fake news detection: a survey of evaluation datasets," PeerJ Comput. Sci., vol. 7, p. e518, 2021.
[17]. M. N. Alenezi and Z. M. Alqenaei, "Machine learning in detecting covid-19 misinfor- mation on twitter," Futur. Internet, vol. 13, no. 10, p. 244, 2021.
[18]. S. V. Mahadevkar et al., "A Review on Machine Learning Styles in Computer Vision - Techniques and Future Directions," IEEE Access, vol. 10, no. September, pp. 107293– 107329, 2022, doi: 10.1109/ACCESS.2022.3209825.
[19]. H. O. Boll et al., "Graph neural networks for clinical risk prediction based on electronic health records: A survey," J. Biomed. Inform., p. 104616, 2024.
[20]. H. Gao, Z. Wang, and S. Ji, "Large-scale learnable graph convolutional networks," in Proceedings of the 24th ACM SIGKDD international conference on knowledge discov- ery & data mining, 2018, pp. 1416–1424.
[21]. P. Vyas, G. Vyas, and J. Liu, Proliferation of health misinformation on social media platforms: a systematic literature review., Issues Inf. Syst., vol. 22, no. 3, 2021.
[22]. B. Khemani, S. Patil, D. K. Kotecha, and D. Vora, "Detecting Health Misinformation: A Comparative Analysis of Machine Learning and Graph Convolutional Networks in