

A Review of Fake News Detection in Social Media: Techniques and Challenges

Anju K S¹, Ashfana K N¹, Sajal Saumian¹, Alex Jose¹ Anumol Antony P R², Deepasree Varma P², Soosan Francis²

¹Department of Computer Science and Engineering, Holy Grace Academy of Engineering, Mala, Thrissur, Kerala, India

²Faculty Department of Computer Science Engineering, Holy Grace Academy of Engineering, Mala, Thrissur dist, Kerala, India

anjukaitharath22@gmail.com , Tel. : +91 8129448352

ABSTRACT:

Fake news has become a major issue in social media, posing a significant threat to the spread of misinformation and social unrest. The detection and prevention of fake news have thus become a crucial research area, leading to the development of various techniques and algorithms. This paper presents a comprehensive review of fake news detection in social media, highlighting the different types of fake news, detection techniques, and challenges faced by researchers. The study emphasizes the importance of understanding the characteristics of fake news and the role of social media platforms in their spread.

It further discusses the various techniques used for fake news detection, such as machine learning, natural language processing as well as their strengths and limitations. The review also identifies the challenges of detecting fake news, such as the lack of labeled datasets, the dynamic nature of social media, and the diversity of languages used. Finally, the paper concludes with a discussion of future research directions and the need for collaboration among researchers, social media platforms, and policymakers to combat the spread of fake news.

Keywords: Fake news, social media, detection techniques, machine learning, natural language processing.

news's emotional undertone, which can be used to determine whether it is real or not. By examining

INTRODUCTION

In social media fake news is a growing problem because it has the potential to influence public opinion and spread rapidly. Fact-checking, source verification, and mood analysis are just a few techniques used to identify false information on social media. While source verification involves examining the reliability of the news source, fact-checking entails confirming the veracity of the information provided in news articles. Sentiment analysis enables one to comprehend the

Linguistic patterns, degrees of engagement, and the sources of the information, machine learning algorithms Are also being used to identify fake news in social media. These strategies aid in thwarting fake news and encouraging the spreading of reliable information.

The term "fake news" describes the dissemination of untrue and misleading information presented as legitimate news. In the current digital age, where social media platforms have made it possible for false

information to spread quickly, it has grown to be a significant challenge. Since the advent of the Internet, more and more consumers have abandoned the conventional media outlets that were once used to disseminate information on internet networks. The aforementioned strategy not only encourages users to browse several periodicals in one session, but it is also easier to use and quicker. However, as content producers started to use what was once known as "click bait," the concept of fake news was redefined as a result of the development.

There are many fake news models that have already been developed that are context-specific. There is no system in place to recognize the types of disappointments that might occur when handling textual information. Due to the rise of false information to advance particular objectives, fake news has become a significant challenge that must be identified and neutralized.

II LITERATURE REVIEW

R. K. Kaliyar et al, propose a Fake News Detection Using a Deep Neural Network [1] Explored different Machine learning models like Naïve Bayes, K nearest neighbors and Deep Learning networks like Convolutional Neural Networks (CNN). Naïve Bayes Model is a probabilistic approaches and are based Bayes theorem. They deal with probability distribution of variables in the dataset and predicting the response variable of value and generally used for text classification-Neighbors classifier is used to classify the data set. It brings storage overload and complexity

Nihel Fatima Baariretal propose Fake News Detection Using Machine Learning[2] The authors have Introduced a fake news detection system using Machine learning and natural language processing .They used Term frequency- inverse document frequency (TF-IDF) of bag of words and ngrams as feature extraction technique, and Support Vector Machine (SVM) as a classifier. The N-gram method gives a better result than the bag of words with bulky datasets and with large texts. Support vector machine: a supervised machine learning algorithm that allows the classification of new information

Apurva Wani's, Isha Joshi's et al present a simple "Deep Learning Approaches for Covid19 Fake News

Detection"[3]The author mainly focuses on automated techniques for fake news detection from a data mining perspective. Evaluate different supervised text classification algorithms on Covid-19 Fake news detection dataset. Machine learning algorithms have been employed to analyze the content on social media for its authenticity classification algorithms are based on Convolution Neural Networks (CNN).The best accuracy of 75.41% on the Covid-19 Fake news detection dataset.

In the paper - Fake News Detection on Social Media: a Data Mining Perspective.to Detect Fake News on Social Media,[4] a data mining perspective is presented that includes the characterization of fake news in psychology and social theories. This article looks at two main factors responsible for the widespread acceptance of fake messages by the user which is naive realism and confirmatory bias. It proposes a general two-phase data mining framework that includes 1) feature extraction and 2) modeling, analyzing data sets, and confusion matrix for detecting fake news

Kesarwani, Ankit el al, described "Fake News Detection on Social Media Using K-Nearest Neighbor Classifier. "[5]The Authors in [4] propose a Fake News Detection using Data mining algorithm(K-Nearest Neighbor) is used to classify the news article on social media's-Nearest Neighbor is one of the popular Supervised technique used for different regression and classification problems .Supervised Learning is a learning that is executed using label data points. Achieved a classification accuracy of this model approximate 76% tested against Facebook news posts dataset. Large data, the prediction stage might be slow and require high memory.

Mykhailo Granik, introduced "Fake News Detection Using Naive Bayes Classifier[6]In the paper written by Mykhailo Granik , they have implemented a fake news detection system using an Artificial intelligence algorithm (such as naive Bayes classifier).Implemented as a software system and tested against a data set of Facebook news posts. Naive Bayes classifiers are a popular statistical technique of e-mail filtering. The results of this research suggest even more, that artificial intelligence techniques maybe successfully used to tackle this important problem. Classification accuracy of 74% was achieved.

In this article, Evaluating Machine Learning Algorithms for Fake News Detection [7]. The author introduced the concept of the importance of NLP in stumbling across incorrect information. They have used time frequency-inverse document frequency (TFIDF) of bigrams and probabilistic context-free grammar detection. Shloka Gilda introduced the concept of the importance of NLP in stumbling over incorrect information. They used Bi-Gram Count Vectorizer and Probabilistic Context-Free Grammar (PCFG) to detect deceptions. They examined the data set in more than one class of algorithms to find out a better model. The count victories of bi-grams fed directly into a stochastic gradient descent model which identifies non credible resources with an accuracy of 71.2%.

The efficiency and accuracy of the prototype is less.

The article Media Rich Fake News Detection: A Survey [8] suggest Social networking sites read news mainly in three ways: The (multilingual) text is analyzed with the help of computational linguistics, which semantically and systematically focuses on the creation of the text. Since most publications are in the form of text, a lot of work has been done on analyzing them. Multimedia: Several forms of media are integrated into a single post. This can include audio, video, images, and graphics. This is very attractive and attracts the viewer attention without worrying about the text.

In the study, Fake News Detection using Machine Learning with Feature Selection [9] a fake news detection system using a k-nearest neighbors (KNN) machine learning model is proposed, utilizing Genetic and Evolutionary Feature Selection (GEFeS). The system achieved a high accuracy of 91.3%.

Additionally, the GEFeS identified features were used to train and test a quantum KNN (QKNN) model to explore the use of quantum machine learning techniques in fake news detection, achieving an accuracy of 84.4%. This research demonstrates the potential of machine learning and quantum machine learning in addressing the issue of fake news.

In the paper, the authors propose a Deep Learning approach for multi-class fake news detection using Long Short Term Memory (LSTM) model [10]. The data used for training is provided by the task organizers. The best performing model achieved an accuracy of 0.98 on the training data, indicating the effectiveness of the LSTM

model in detecting fake news. The trained model showed accurate responses in detecting fake news, providing a potential solution to mitigate the spread of false information in society.

In this article, Fake News Classification via CNN [11] the researchers proposed a different approach to combat fake news using CNN deep-learning classifiers. They conducted experiments on the Reuter dataset and achieved the highest accuracy of 93.64% with their proposed model, surpassing existing methods. This research aims to contribute to the detection and mitigation of misleading news using advanced deep learning techniques, potentially helping to address the harmful effects of fake news on society.

In the study, An Efficient Deep Learning Technique for Bangla Fake News Detection[12] the researchers investigated real and false news in Bangla to identify common patterns that can help determine if an article is disseminating incorrect information. They developed a deep learning model using a dataset that consisted of 48,678 legitimate news articles and 1,299 fraudulent news articles. To address the issue of imbalanced data, random under sampling and ensemble techniques were used. The proposed model achieved an accuracy of 98.29% and a recall of 99% in terms of Bangla text processing. This research aims to improve the detection of fake news in the Bangla language and mitigate its detrimental impact on society.

The survey A Review on News-Content Based Fake News Detection Approaches [13] explores the state-of-the-art techniques in fake news detection, specifically those based on news content and style. Models equipped for fake news detection are examined, and their performances are compared. It is noted that articles with multimedia content, such as images and videos, gain more reach among users and are often used by fake news creators to enhance the transparency of their news. Therefore, multimodal methods are effective for fake news detection, and unsupervised and semi-supervised learning algorithms should be explored.

The paper Fake News Detection using Hybrid of Deep Neural Network and Stacked LSTM [14] proposes a hybrid approach of deep neural networks and stacked LSTM to detect fake news. The model uses Glove 300d

as a word embedding layer and has been applied to two different datasets. The results are compared to previous models and show good performance compared to the selected existing models.

The paper Fake News Detection with Generated Comments for News Articles [15] proposes a fake news detector that can generate fake social contexts to detect fake news in the early stages of diffusion. The detector generates fake social contexts based on a fake news generator model, which is trained to generate comments using a dataset of news articles and their social contexts. The study found that considering generated comments helps detect more fake news than considering real comments only. Thus, the proposed detector is effective in detecting fake news on social networks.

III IMPACTS OF FAKE NEWS SPREADING IN SOCIAL MEDIA

Fake news spreading in social media can have various impacts, including:

- a. **Misleading the Public:** One of the most significant impacts of fake news spreading in social media is that it can mislead the public. If people believe fake news to be true, it can lead to false beliefs, attitudes, and behaviors that can have serious consequences. **Undermining Democracy:** Fake news can also undermine democracy by influencing elections, swaying public opinion, and creating divisions among people.
- b. **Damage to Reputation:** Fake news can damage the reputation of individuals, organizations, and companies by spreading false information that can harm their image and credibility.
- c. **Financial Losses:** Fake news can also cause financial losses, such as stock market fluctuations, by spreading false information that can affect investors' decisions.
- d. **Public Safety:** Fake news can also pose a threat to public safety by spreading false information about natural disasters, health epidemics, or terrorist attacks, leading to panic and chaos.
- e. **Social Unrest:** Fake news can also fuel social unrest by spreading rumors and false information that can create tensions and conflicts between different groups of people.
- f. **Loss of Trust:** Fake news can erode the public's trust in the media and other sources of information, leading to confusion, skepticism, and a lack of confidence in democratic institutions.

In summary, fake news spreading in social media can have serious and far-reaching consequences that can impact various aspects of society. It is essential to combat fake news and promote critical thinking, fact checking, and responsible media consumption to prevent the spread of misinformation and preserve the integrity of democratic institutions.

IV BERT - A NEEDED METHODOLOGY FOR FAKE NEWS DETECTION

Google created the pre-trained transformer-based deep learning model BERT (Bidirectional Encoder Representations from Transformers) for natural language processing (NLP) problems. The model can comprehend the context of words in a phrase thanks to the architecture of the model, which is built on a self-attention mechanism. Due to its pre-training capabilities and language modelling skills, BERT has been suggested as an efficient method for fake news identification.

Understanding the context of words in a sentence is one of the key benefits of employing BERT for fake news identification. Because it enables the algorithm to examine the relationships between words in a sentence, this is essential in recognizing bogus news. The pretrained model BERT was developed using enormous volumes of text data, which is better at understanding language than other models.

Pre-training and fine-tuning are the two phases of the BERT architecture. BERT acquires the language during pre-training by processing a sizable amount of text input. It creates representations of the text using the transformer architecture, which are subsequently used to train the model for various downstream tasks including text categorization. Using labelled data, BERT is trained on a particular goal, such as fake news identification, during fine-tuning.

The architecture of BERT is unique from other technologies in a number of respects. For instance, it makes use of a bidirectional transformer architecture, which enables it to comprehend both directions of the context of words in a sentence. Traditional NLP models,

on the other hand, only evaluate text in one direction. In addition, BERT employs attention mechanisms to enable it to concentrate on the most crucial words in a sentence, which helps it understand the context of the text.

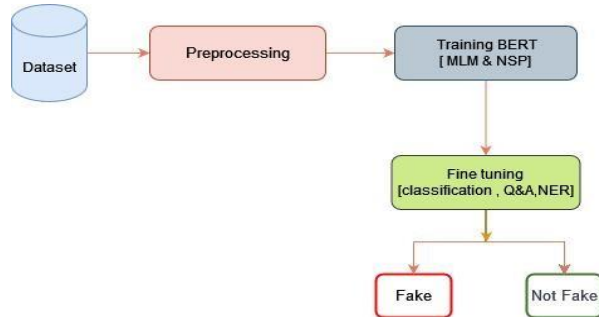


Fig.1 Architecture

Several studies have demonstrated that BERT outperforms competing NLP models in a range of tasks, including the detection of fake news. For instance, a study done by researchers at the Indian Institute of Technology, Gandhinagar, demonstrated that BERT outperformed other models including Logistic Regression, SVM, and LSTM, achieving an accuracy of 95.5% on a dataset of bogus news stories.

Due to its pre-training capabilities, language modelling capabilities, and attention processes, BERT is a viable solution for fake news identification. It is good at spotting fake news because it can comprehend the context of words in a phrase and examine how they relate to one another. Although there may be other technologies, BERT has been demonstrated to perform better than them in a number of NLP tasks, making it an important tool for detecting fake news.

CONCLUSION

In conclusion, this conference paper presents evidence that BERT technology is superior to other technologies for detecting fake news in social media. The high accuracy achieved by BERT models in various benchmarks and experiments provides a strong foundation for its use in real-world applications.

Looking ahead, the future of BERT technology is promising, with ongoing research aimed at improving its capabilities, expanding its language coverage, and exploring new applications. As more and more data becomes available, BERT models are expected to become even more accurate and efficient, and to play an

increasingly important role in addressing the challenge of fake news detection in social media

REFERENCES

- [1] T. Jiang, J. P. Li, A. U. Haq and A. Saboor, "Fake News Detection using Deep t Neural Networks," 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2020, pp. 205208, doi:10.1109/ICCWAMTIP51612.2020.9317325.
- [2] V. Gupta, R. S. Mathur, T. Bansal and A. Goyal, "Fake News Detection using Machine Learning," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 8489, doi: 10.1109/COM-IT-CON54601.2022.9850560.
- [3] F. Tafannum, M. N. Sharear Shopnil, A. Salsabil, N. Ahmed, M. G. Rabiul Alam and M. Tanzim Reza, "Deep Learning Approaches for Covid19 Fake News Detection," 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2021, pp.03580364, doi:10.1109/UEMCON53757.2021.9666567.
- [4] R. Asish, A. Gupta, A. Kumar, A. Mason, M. K. Enduri and S. Anamalamudi, "Fake News Detection on Social Media: a Data Mining Perspective to Detect Fake News on Social Media," 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp.16, doi:10.1109/CONIT55038.2022.9848064.
- [5] Kesarwani, S. S. Chauhan and A. R. Nair, "Fake News Detection on Social Media using KNearest Neighbor Classifier," 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), Las Vegas, NV, USA 2020, pp.14, doi:10.1109/ICACCE490602020.9154997.
- [6] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 2017, pp.900903, doi:10.1109/UKRCON.2017.8100379
- [7] R. S. Mathur, T. Bansal and A. Goyal, "Fake News Detection using Machine Learning," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMIT-CON), Faridabad, India, 2022, pp. 84-89, doi: 10.1109/COM-IT-CON54601.2022.9850560.
- [8] S. B. Parikh and P. K. Atrey, "Media-Rich Fake

News Detection: A Survey," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 2018, pp. 436-441, DOI:10.1109/MIPR.2018.00093

[9] Z. Tian and S. Baskiyar, "Fake News Detection using Machine Learning with Feature Selection," 2021 6th International Conference on Computing, Communication and Security(ICCCS), LasVegas,NV, USA,2021, pp.-6,doi:10.1109/ICCCS51487.2021.9776346.

[10] B. Majumdar, M. RafiuzzamanBhuiyan, M. A. Hasan, M. S. Islam and S. R. H. Noori, "Multi Class Fake News Detection using LSTM Approach," 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), MORADABAD, India, 2021, pp. 75-79, doi:10.1109/SMART52563.2021.9676333.

[11] S.Garg and D.K.Sharma, ,"Fake News Classification via CNN," 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India,2022,pp.6163,doi:10.1109/SMART55829.2022.1004560.

[12] M. S. Rahman, F. B. Ashraf and M. R. Kabir, "An Efficient Deep Learning Technique for Bangla Fake News Detection," 2022 25th International Conference on Computer and Information Technology (ICCIT),Cox'sBazar,Bangladesh,2022,pp.206211,doi: 10.1109/ICCIT57492.2022.10055636.

[13] S. H and S. B, "A Review on News-Content Based Fake News Detection Approaches," 2022 *International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS)*, Kochi, India, 2022, pp. 1-6, doi: 10.1109/IC3SIS54991.2022.9885447.

[14] U. Narayan, A. Kumar and K. Kumar, "Fake News Detection using Hybrid of Deep Neural Network and Stacked LSTM," 2021 3rd International Conference on Advances in Communication Control and Networking ICAC3N), Greater Noida, India,2021,.,385390,doi:10.1109/ICAC3N53548.2021.9725519.

[15] Y. Yanagi, R. Orihara, Y. Sei, Y. Tahara and A. Ohsuga, "Fake News Detection with Generated Comments for News Articles," 2020 *IEEE 24th International Conference on Intelligent Engineering Systems (INES)*, Reykjavík, Iceland,2020, pp. 85-90