IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204119 |

A Review of Methodologies for Fake News Analysis

Mr.S. PRAGADEESWARAN, JANARTHANAN K, ANAND K, YUGAPRIYA S, GOWSHIKA S,

Assistant Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: Reviewing relevant literature for fake news analysis methods is the main focus of this paper. In a comprehensive review, we found that the overall methodology can be separated into two categories: study perspectives and fake news detection techniques. Fake news can be studied from four perspectives, and detecting fake news can be done manually or automatically. Manual fact-checks fall into two categories: expert-based and crowd-sourced. Many researchers have used data science techniques to detect fake news automatically. The main techniques used for detecting fake news are deep learning techniques such as CNN, RNN, LSTM, and Bi-LSTM. Fake news can be detected using SVM, NB, LR, DT, RFC, and BM, among other traditional machine learning techniques. As a future study method, Bayesian modelling could be the most promising. Due to the fact that it updates the prior distribution every time new data is received, Bayesian modelling is a very robust method. This makes it suitable for use in any field, including big data contexts.

Keywords: Bayesian modelling, classification, deep learning, fake news detection

I.INTRODUCTION

In today's world, we are inundated with information every second by a variety of media. Since data is easily accessible and loosely authenticated, much of it cannot be verified and can be fake. By posting, sharing, and commenting on fake news, people who have little time to verify it are contributing to propagating this false information. Consumers of this information are thus susceptible to fake details, which negatively affect society at all levels: local, national, and global. Fake information can be hazardous to the people who are the subject of the information. Especially when it comes to key world events like central elections, fake news propagation can have horrible results on the public, such as transforming their Fake news can be studied from four main perspectives, which have been discussed in Section III. Furthermore, fake news can be detected manually or automatically. Manual fake news detections are done by either expert-based techniques or crowed sourced based techniques. Manual-based methods were mainly used before automatic methods were invented.

This method was popular when technology was not so good and computation power was not very high. With technological developments and increasing computation power, contemporary methods such as automatic techniques were invented. Automatic fake news detections are done using deep learning and traditional machine learning techniques. A variety of deep learning approaches, including RNNs, LSTMs, Bi-LSTMs, CNNs, and others, have been proposed for detecting fake news. Moreover, researchers have proposed methods for detecting fake news using traditional machine learning techniques, including SVM, NBC, RF, DT, LR, and BM. Based on the comparison, we found that machine learning-based methods are more appealing and used by many researchers. Further detailed explanations of these methods are described in the next sections.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204119 |

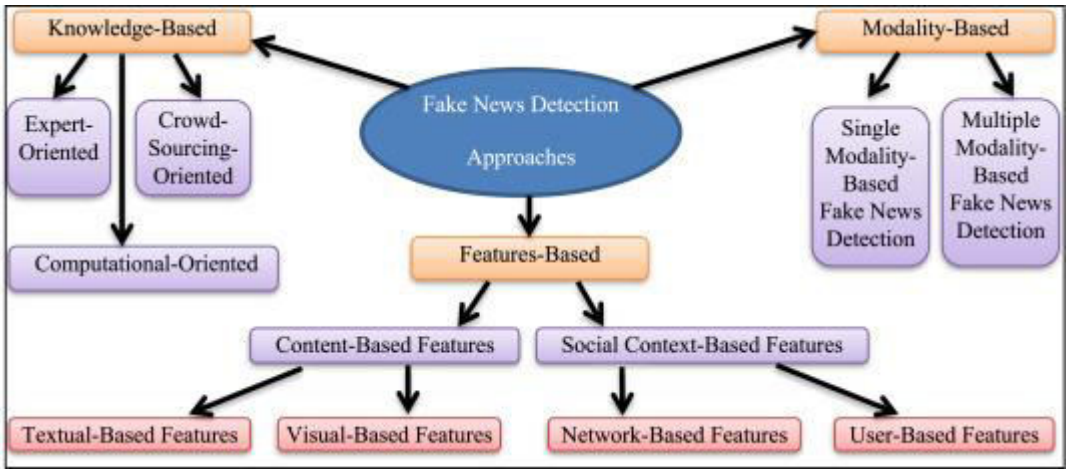

Fig 1: A review of fake news detection approaches

Knowledge-based approaches, also known as fact-checking, involve collecting raw facts from the open web to construct knowledge. However, this data must be further processed to address redundancy, invalidity, conflicts, unreliabilities, and incompleteness issues. Five tests are necessary to ensure the accuracy of the facts: entity resolution, time recording, consistency evaluation, completeness evaluation, and accuracy evaluation. Entity resolution is used to link records of similar facts, while time recording is used to measure the temporal validity of facts. Consistency evaluation ensures that facts are not contradictory, completeness evaluation ensures that all facts are accounted for, and accuracy evaluation ensures that the facts are correct [7], [8]. The data can be effectively cleansed by running these five tests, and knowledge-based approaches can be successfully implemented

II.LITERATURE SEARCH

This paper uses the Google Scholar database as a reliable source to assess recent trends in fake news detection research over the past five years. Using “fake news”, “fake news detection”, “multimodal fake news detection”, “multidisciplinary + fake news” and “explainable fake news detection” as relevant keywords, the database was queried. The search yielded a total of 18,100 references published in the fake news detection field between 2018 and 2023. The results obtained were carefully analyzed to identify the prominent research directions in the field. This review article provides a comprehensive overview of the current state of research in fake news detection and highlights the most promising avenues for future investigation. With the help of deep insights into the various concerns in the field of fake news detection research, we are able to better appreciate the potential of this technique to develop more efficient and credible fake news detection techniques.

The specific research on fake news detection includes fact verification, position detection, topic detection and other tasks involving text classification, text clustering, image understanding, speech recognition and other research directions. The related research on fake news detection uses technologies such as text mining, machine learning, deep learning, natural language processing machine vision and other technologies to extract and identify key information from subjective visual perceptions of text or news pages. According to the summary and classification of a large number of references, the entire fake news detection method can be divided into three stages: (1) Machine learning stage. (2) Compared with machine learning algorithms, deep learning is not limited by manual feature extraction.

It can extract text features from language texts through the self-learning ability of the network layer, which greatly improves the system performance of natural language processing tasks. Deep learning networks, including convolutional neural networks and recurrent neural networks, are applied to fake news detection tasks. They can effectively learn complex semantic features and high-level semantic representations from text and have been shown to improve the performance of fake news detection tasks. (3) However, the manual annotation of large text data is very complex, and the data used for natural language processing tasks is very limited. The application of deep learning models with strong dependency data in natural

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204119 |

language processing is also very challenging. In order to avoid the problems of overfitting and insufficient generalization ability caused by insufficient data volume, researchers began to explore the pre-trained model for semantic representation.

III.METHODS

The content-based fake news detection method aims to extract various semantic features from the news content and detect the authenticity of the news through these features. There are some linguistic differences between fake news and true news, and fake news can be detected by distinguishing the language style of true and fake news texts. Fake news is more subjective than real news. The study found that the first person and the second person are used more in fake news, and fake news contains more words that can be used for exaggeration (such as subject, transfinite words and modal adverbs), while real news often uses specific (such as numbers), objective (such as third person) and positive words. The author of fake news will be more extreme. The study analyzes the writing styles of left-wing news and right-wing news and finds that they all have extremist tendencies, political tendencies and hatred. However, not every linguistic feature has the same weight, and the importance of different lexical features is different. Song et al. extracted a complete set of content features from real and fake news, including the total number of words in the news, the length of the content, the number of capital words, special symbols, sentences at the beginning of the number, offensive words, etc. Through experiments, the importance ranking of the features is listed, and it is found that the total number of words, the length of the content and the number of capital words have a greater impact on the discrimination of real news. Abbreviations and the total number of words have a greater impact on discrimination against fake news.

The content-based approach can discover the linguistic features of true and fake news. However, sometimes fake news will mislead readers by deliberately imitating the writing techniques of real news. The content-based approach cannot distinguish the feature differences between such fake news and true news. In order to solve this problem, we can make full use of hidden information as auxiliary data, such as social background information and propagation paths in social networks. Social background information is one of the research directions. explored the relationship between user data and fake news on social media and used the user's social participation as auxiliary information for detection. Furthermore, proposed a framework to simulate the triadic relationship between news publishers, news articles and users, extract effective features from the participation behavior of news publishers and readers, and then capture the interaction between them. Studies have shown that the use of social background information can not only improve the effect of fake news detection but also effectively predict it early. Another research direction detects fake news by simulating the propagation path of fake news in the network. Through experiments, proposed a fake news detection framework based on the Transformer architecture, which includes encoder and decoder parts. The encoder part is used to learn the representation of fake news data, and the decoder part is used to predict future behavior based on past observations. The model uses the characteristics of news content and social background to improve classification accuracy.

With the rapid development and application of machine learning and artificial intelligence technology in various fields, it is very important to explain the results of the algorithm's output to the user. The interpretability of artificial intelligence means that people can understand the choices made by artificial intelligence models in their decision-making process, including the reasons, methods and content of decision making Simply put, interpretability is the ability to turn artificial intelligence from a black box into a white box. At present, explainable artificial intelligence methods are applied to different fields in different industries, including biomedical, financial applications, video payment, and media industries. The core of explainable artificial intelligence is to obtain human trust. From this, we can see that there are two important concepts that can explain artificial intelligence: trust and interpretation. For explainable artificial intelligence, the connotation of interpretation is that agents must communicate, exchange and run into different people repeatedly before they can gain human trust. Therefore, for the agent, when explaining, it is necessary to consider the different educational backgrounds, knowledge levels and other factors of the audience and then design the content and form of the explained information.

IV.RESULT ANALYSIS

In addition, research on multimodal fake news detection has gradually increased in recent years. proposed a hierarchical multi-modal context attention network for fake news detection, which includes two modules: a multi-modal context attention module and a hierarchical coding module. To model the multi-modal context of news posts, the multi-modal

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204119 |

context attention module uses pre-trained for text representation and pre-trained for image representation, ensuring a seamless integration of both textual and visual information. It combines inter-modal and intra-modal relationships to enhance fake news detection. The hierarchical coding module captures the rich hierarchical semantics of the text to improve the representation of multimodal news.

The MCAN model proposed by aims to learn multi-modal fusion representations by considering the dependencies between different modalities. The model includes three main steps: feature extraction, feature fusion and fake news detection. In the feature extraction step, three sub-models are used to extract features from the spatial domain, frequency domain and text. The VGG-19 network is used to extract visual features from the spatial domain, and the ACNN-based sub-network is designed to extract features from the frequency domain, especially for re-compressed or tampered images. Furthermore, the BERT model is used to obtain the text features of the text content. In the feature fusion step, the deep common attention model is used to fuse multimodal features. The fusion process simulates the way that humans first see the image and then read the text. The common attention model is composed of multiple common attention layers, which capture the interdependence between different features. Finally, the fusion feature is used to detect fake news, and the output of the common attention model is used to judge the authenticity of the input news.

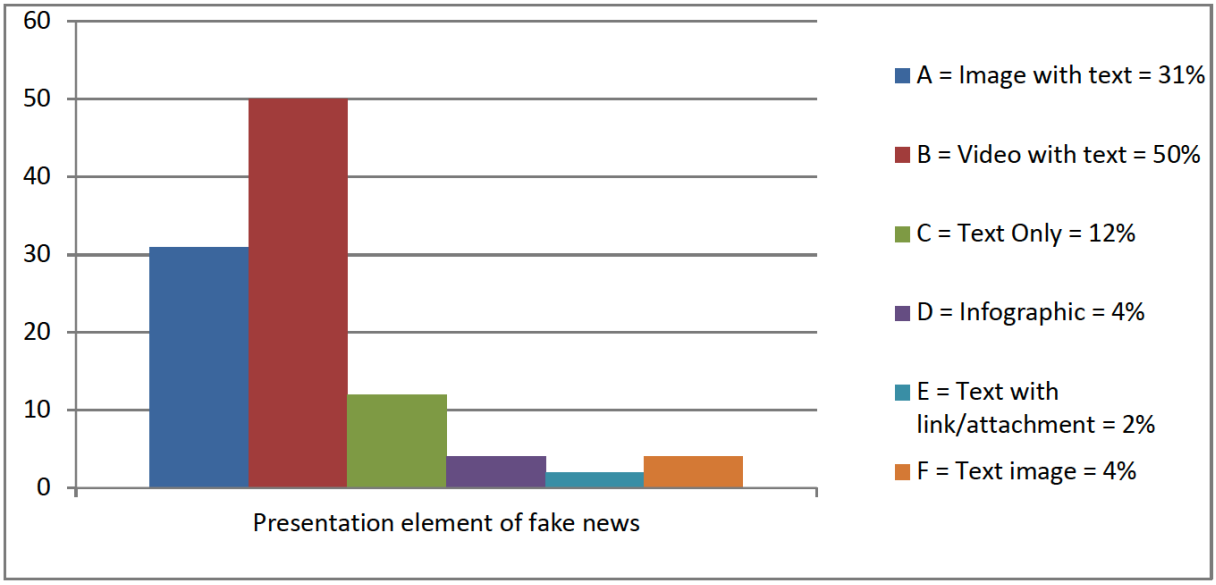

Fig 2: A Study on Fake News Subject Matter result analysis

proposed a cross-modal contrastive learning framework, COOLANT, for multimodal fake news detection. The framework consists of three main components: a cross-modal contrastive learning module for alignment, a cross-modal fusion module for learning cross-modal correction and a cross-modal aggregation module with an attention mechanism and guidance to improve the performance of multimodal fake news detection. The cross-modal contrast learning module aligns features by converting single-modal embedding into a shared space. It uses auxiliary cross-modal consistency learning tasks to measure the semantic similarity between images and texts and provides soft targets for the contrast learning module. The contrastive learning module uses the contrast loss to predict the actual image-text pairing in the batch.

V.CONCLUSION

The accuracy and reliability of information dissemination are of great help to the sustainable development of society and the economy. To embrace digital transformation, green information technology and responsible information production and consumption, we can reduce resource consumption and improve efficiency. These measures can achieve the long-term benefits of information dissemination. This paper focuses on investigating existing fake news detection technology and providing an overview of the research status of fake news detection methods. We collected almost all commonly used

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 4, April 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1204119 |

datasets, classified them from the perspectives of single-mode and multi-mode and summarized the research methods for fake news detection. This includes content-based detection methods, social network-based detection methods and knowledge-based detection methods. Considering the popularity of multimodal technology, we have also sorted out multimodal fake news detection methods. Additionally, this paper also discusses the research progress of fake news in multidisciplinary fields. Furthermore, we discuss the general fake news detection technology along with the explainable fake news detection method. Specifically, we propose a human-machine-theory explainable triangular communication framework. It is characterized by being people-centered, incorporating multidisciplinary knowledge and aiming to establish the sustainable development of a human–machine interaction information dissemination system.

REFERENCES

1. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* **2018**, *359*, 1146–1151. [Google Scholar] [CrossRef] [PubMed]
2. Apuke, O.D.; Omar, B. Fake news and COVID-19: Modelling the predictors of fake news sharing among social media users. *Telemat. Inform.* **2021**, *56*, 101475. [Google Scholar] [CrossRef] [PubMed]
3. Van Der Linden, S.; Roozenbeek, J.; Compton, J. Inoculating against fake news about COVID-19. *Front. Psychol.* **2020**, *2020*, 2928. [Google Scholar] [CrossRef]
4. Rocha, Y.M.; de Moura, G.A.; Desidério, G.A.; de Oliveira, C.H.; Lourenço, F.D.; de Figueiredo Nicolete, L.D. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: A systematic review. *J. Public Health* **2021**, *31*, 1007–1016. [Google Scholar] [CrossRef] [PubMed]
5. Moscadelli, A.; Albora, G.; Biamonte, M.A.; Giorgetti, D.; Innocenzio, M.; Paoli, S.; Lorini, C.; Bonanni, P.; Bonaccorsi, G. Fake news and Covid-19 in Italy: Results of a quantitative observational study. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5850. [Google Scholar] [CrossRef]
6. Zhou, X.; Zafarani, R. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Comput. Surv. (CSUR)* **2020**, *53*, 1–40. [Google Scholar] [CrossRef]
7. Zhang, X.; Ghorbani, A.A. An overview of online fake news: Characterization, detection, and discussion. *Inf. Process. Manag.* **2020**, *57*, 102025. [Google Scholar] [CrossRef]
8. Hu, L.; Wei, S.; Zhao, Z.; Wu, B. Deep learning for fake news detection: A comprehensive survey. *AI Open* **2022**, *3*, 133–155. [Google Scholar] [CrossRef]
9. Athira, A.B.; Kumar, S.M.; Chacko, A.M. A systematic survey on explainable AI applied to fake news detection. *Eng. Appl. Artif. Intell.* **2023**, *122*, 106087. [Google Scholar]
10. Hotho, A.; Nürnberger, A.; Paaß, G. A brief survey of text mining. *J. Lang. Technol. Comput. Linguist.* **2005**, *20*, 19–62. [Google Scholar] [CrossRef]

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details