IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 3, June 2025, pp. 1951∼1959
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i3.pp1951-1959

Comparative analysis of machine learning models for fake news detection in social media

Bahaa Eddine Elbaghazaoui[1,2], Mohamed Amnai[2], Youssef Fakhri[2], Ali Choukri[2], Noreddine Gherabi[3]
[1]National School of Applied Sciences of Sultan Moulay Slimane University, Beni Mellal, Morocco
[2]Laboratory of Computer Sciences Research, Faculty of Sciences, Ibn Tofail University Kenitra, Kenitra, Morocco
[3]National School of Applied Sciences of Sultan Moulay Slimane University, Khouribga, Morocco

Article Info

Article history:

Received Sep 2, 2024
Revised Nov 22, 2024
Accepted Jan 27, 2025

Keywords:

Data profiling
Fake news
Information diffusion
Machine learning
Social media

ABSTRACT

The rapid rise of information sharing on social media has amplified the spread of fake news, making its detection increasingly critical. As fake news continues to proliferate, the need for efficient detection mechanisms has become more urgent to protect users from misinformation and disinformation. This paper presents a comparative analysis of multiple machine learning models for detecting text-based fake news on social media platforms. Using models such as gradient boosting, XGBoost, and linear support vector classifier (SVC) on the Information Security and Object Technology (ISOT) fake news dataset, the study demonstrates that gradient boosting achieves the highest accuracy of 99.61%, while XGBoost provides a strong balance with 99.59% accuracy and a significantly lower execution time, making it more suitable for real-time applications. These results offer valuable insights into the trade-offs between accuracy and computational efficiency, contributing to the development of more practical detection systems and future research in the field.

Corresponding Author:

Bahaa Eddine Elbaghazaoui
National School of Applied Sciences in Beni Mellal, Sultan Moulay Slimane University
Beni Mellal, Morocco
Email: b.elbaghazaoui@usms.ac.ma

1. INTRODUCTION

Social media users frequently share, like, or repost content, which leads to the widespread dissemination of information, including news from independent authors [1]. This uncontrolled flow of information has contributed to the rise of fake news, a growing global issue [2]. While the concept of fake news is not new, the pervasive use of social media has significantly amplified its impact, facilitating the rapid spread of misinformation [3].

The urgency of addressing fake news lies in its potential to harm both individuals and society [4]. Fake news, defined as intentionally misleading information disseminated through news outlets or social media, can influence elections, deepen political divides, and shape public perception in dangerous ways [5]. The pervasive spread of fake news undermines the truth within the information ecosystem and misleads readers into accepting biased or inaccurate information, thereby affecting how people interpret real news [6]. In response to this growing problem, this paper addresses the challenge of detecting fake news by applying advanced machine learning algorithms to analyze and classify false information shared on social media platforms. Identifying fake news accurately is essential for preserving the integrity of online discourse and protecting users from misinformation.

Traditional computational methods, such as relying on satirical news sources or fact-checking websites, have limitations, including a narrow focus on specific topics like politics and the requirement for human expertise [7]. These constraints hinder the development of models that can generalize across multiple domains, which is necessary for an effective and scalable fake news detection system. The primary objective of this study is to identify the most effective machine learning model for detecting fake news in social media posts. We achieve this by collecting and cleaning a dataset, applying various machine learning models, and comparing their performance to determine the optimal approach for fake news detection.

This study contributes to the field of fake news detection by providing a comparative analysis of multiple machine learning models, specifically gradient boosting, XGBoost, and linear linear support vector classifier (SVC), applied to a large social media dataset. Unlike previous research that often focuses on a single detection method or narrow use cases, this study evaluates both the accuracy and execution time of various models, offering insights into their trade-offs in real-world applications. Additionally, the study demonstrates the practical relevance of balancing computational efficiency with detection precision, making it a significant contribution for the development of scalable, real-time fake news detection systems. The findings provide a valuable foundation for future research in improving the detection of misinformation, with a particular emphasis on real-time detection capabilities and broader generalization across domains.

This paper is structured as follows: section 2 reviews the existing literature on fake news detection, highlighting gaps in current approaches and outlining the objectives and methodology of this study. Section 3 examines the spread of misinformation on social media and explains how it can be profiled using data, including details on data collection, cleaning, and preprocessing techniques. Section 4 describes the application of various machine learning models for detecting fake news, and compares their performance based on accuracy and execution time. Finally, section 5 evaluates the results, interprets the findings in comparison to previous studies, discusses the limitations of the study, and offers suggestions for future research, including the potential for real-time detection systems and the integration of multimedia data.

2. RELATED WORKS

This section outlines significant early research on the concept of fake news and provides an overview of the growing body of studies examining its dissemination [8]. The advent of the world wide web (WWW) and the rapid adoption of social media platforms like Facebook and Twitter have revolutionized the way information is shared, enabling unprecedented levels of content dissemination [9]. Social media has allowed consumers to produce and share more information than ever before, some of which is inaccurate or misleading, contributing significantly to the spread of fake new [10].

Scholars have conceptualized fake news in various ways, but the core definition remains consistent across different interpretations [11]. Fake news typically refers to content that is deliberately created and disseminated with the intention of deceiving people into believing falsehoods [12]. This includes myths, rumors, conspiracy theories, hoaxes, and other forms of misleading or incorrect information, which are frequently shared on social media platforms either intentionally or inadvertently [13].

A pertinent example of the impact of fake news is the use of artificial intelligence (AI)-generated content involving prominent figures like Elon Musk [14]. For instance, a fabricated image circulated online depicts Musk endorsing a financial scheme called "Quantum AI," falsely claiming that it allows Australians to "get rich quick." In reality, the following Figure 1, designed to mimic a CNN news broadcast, is part of a broader scam that exploits AI to deceive viewers into investing in a fraudulent venture. This case illustrates how sophisticated technologies, such as AI, can be manipulated to create convincing fake content, complicating efforts to discern truth from deception on social media platforms. The use of well-known figures like Elon Musk, combined with credible-looking news formats, can easily trick unsuspecting viewers into believing false claims. This underscores the importance of media literacy and the ability to critically assess the authenticity of online content.

Moreover, the COVID-19 pandemic has heightened the relevance of fake news research, with studies exploring the link between social media and the spread of misinformation related to public health [15]. During the pandemic, fake news about the coronavirus, such as exaggerated death tolls and unfounded conspiracy theories, spread rapidly on social media, posing significant challenges to public health efforts [16]. This misinformation has fueled public anxiety, created confusion, and hindered the efforts of healthcare workers and public health systems in managing the crisis effectively.

The primary objective of this article is to develop methods for identifying fake news across various databases. By leveraging appropriate machine learning models, we aim to accurately predict and detect fake news and misinformation. The outcomes of this study will aid in identifying false information and help users avoid being misled by deceptive content, thereby reducing the impact of fake news on public discourse.

While previous studies have focused on either individual machine learning models or specific domains such as politics or public health, the novelty of this work lies in its comprehensive comparative analysis of multiple machine learning models for detecting fake news across a broader domain. Unlike existing research that often emphasizes accuracy alone, our study evaluates both the accuracy and execution time of models like gradient boosting, XGBoost, and linear SVC, making it one of the few works to consider computational efficiency alongside detection performance. This dual focus on precision and speed is critical for developing real-time fake news detection systems that can be applied across diverse datasets and social media platforms. Furthermore, by exploring the trade-offs between model effectiveness and resource efficiency, our work offers practical insights that can be applied to the real-world deployment of scalable detection systems.

Figure 1. AI-generated fake CNN broadcast featuring Elon Musk promoting fraudulent 'Quantum AI' scheme

3. CHALLENGES IN FAKE NEWS DETECTION

Once news is released, a large number of users may engage in its spread across online social media networks, where users can share their thoughts and opinions [17]. For each news item, numerous posts can be observed and collected on these platforms. However, many of these unverified posts are false [18]. Previous attempts to utilize news propagation for fake news detection have employed various features. Most research in this area has focused on supervised learning approaches, while fewer studies have explored unsupervised models. Additionally, many of these methods rely on a single machine learning model to build a specialized framework for detection. In contrast, this paper aims to tackle the challenge of fake news detection by evaluating multiple models to identify the one that offers the best balance of speed and accuracy.

4. SPEADING FAKE NEWS

Fake news is disseminated through social and mainstream electronic media in various forms, including comments, political agendas, news articles, rumors, and satire. It is widely used to spread misleading information, false persuasion, and confusion, posing a significant threat to public trust in online activities such as social interactions, e-commerce, and media consumption [19]. The challenge of detecting and mitigating fake news is complex due to its dynamic and heterogeneous nature. Researchers in natural language processing (NLP) have developed numerous solutions to address this issue [20]. The key challenges faced by academic researchers in this field include the diversity, speed, quantity, and persistence of fake news, as illustrated in Figure 2.

Fake news often starts as disinformation, which is deliberately fabricated for a specific purpose, and later spreads as misinformation, when people unknowingly share incorrect information. The primary motivations behind the creation of fake news include:

- Misinterpretation or distortion of true news
- Profit generation through clicks or website traffic
- Promotion of individuals, political parties, or specific viewpoints
- Misunderstandings from jokes or parody content being taken seriously

Due to confirmation bias, people are more likely to accept information as true if it aligns with their existing beliefs [21]. Even those who do not typically believe in such content may be influenced by its shocking nature, leading them to share it.

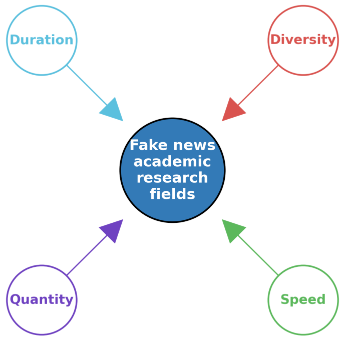

Figure 2. Core research areas in fake news studies

5. DATA PROFILING AND FAKE NEWS

The core business of many technology companies involves creating identity systems based on data collection and profiling [22]. This has significant implications for news organizations, which often struggle to deliver content effectively on platforms like Facebook and Twitter [23]. These technology platforms have the power to shape audiences and influence their demands, holding much of the data necessary to fully understand how information reaches its intended audience. This situation creates modern data blind spots and contributes to a growing mistrust between news organizations and the public. Without accurate and comprehensive communication of vital facts, the reality of information can become distorted.

Data profiling is crucial in understanding the value of information and identifying tools that can detect fake news on social media [24]. It supports the development of systems that automatically detect misinformation in news. Metadata features such as account names, images, followers, number of likes, tags, and release dates can provide signals that indicate suspicious activities and help identify fake news.

The internet hosts data in various formats, including text, images, video, and audio. Detecting and classifying unstructured news published online can be challenging, often requiring human expertise. However, computational techniques, including data profiling, can be employed to detect anomalies that differentiate deceptive text articles from fact-based ones. Additionally, data profiling can be used to analyze the spread of fake news compared to real news, providing insights into how the dissemination of false information on the internet differs from that of accurate reporting.

6. MACHINE LEARNING AND FAKE NEWS

AI, machine learning, and data science are valuable tools for addressing interdisciplinary problems, including the challenge of fake news [25]. These fields are often distributed across multiple sub-disciplines within the data analysis community. The spread of fake news has led to widespread distrust in media, politics, and established institutions globally. While emerging technologies like AI have the potential to exacerbate the problem, they can also be harnessed to combat misinformation. For example, deepfake technology, which uses AI to manipulate audio, images, and videos, can cause individuals to appear to say or do things they never actually did [26]. This innovation could usher in a new wave of misinformation, easily distributed via social media.

The creation and dissemination of false information are often driven by economic and ideological motives [27]. Unfortunately, efforts to combat fake news are unlikely to deter those with vested interests from developing new methods to spread misinformation [28]. Research in this area often focuses on understanding why people are susceptible to false information, as these vulnerabilities are frequently exploited by those who create and spread fake news.

Identifying fake news requires a deep understanding of the issue, which is challenging because many people tend to believe misinformation without verifying it, simply because they are not aware of the full context. Numerous studies have focused on identifying and categorizing fake news on social media platforms such as Facebook and Twitter. Fake news is conceptually classified into various forms, and this classification is then used to develop machine learning models that can generalize across different domains.

Machine learning offers a powerful and automated way to detect fake news. When fake news is posted, a machine learning system can analyze the content and determine its authenticity. Researchers are actively working on developing the most effective machine learning classifiers to identify fake news. The accuracy of these classifiers is crucial; if a classifier fails to correctly identify false news, it could have harmful consequences for many individuals. The accuracy of a classifier depends on the quality of its training. A well-trained model can achieve higher accuracy in detecting fake news, and various machine learning classifiers can be employed for this purpose.

7. EXPERIMENTAL EVALUATION

In our study, we utilized the ISOT Fake News Dataset, which was sourced from Kaggle. This dataset is widely used for fake news detection research and contains a comprehensive collection of real and fake news articles. The ISOT Fake News Dataset, created by the Information Security and Object Technology (ISOT) research group at the University of Victoria, is widely used for fake news detection tasks [29]. The dataset is structured into two categories: real news and fake news. The following are the key statistics:

a) Total number of articles: the dataset contains a total of 44,898 articles.
b) Real news: 21,417 real news articles were collected from credible news sources such as Reuters.
c) Fake news: 23,481 fake news articles were gathered from unreliable websites flagged by Politifact and Wikipedia.
d) Time period: the dataset covers articles published mainly in 2016 and 2017, particularly during significant political events.
e) Data format: the dataset is provided in CSV format with the following columns:
   - Title: the title or headline of the article.
   - Text: the full text or body content of the article.
   - Label: a binary label indicating whether the article is real (1) or fake (0).

This dataset is frequently used for training and evaluating machine learning models to detect fake news [30]. Common metrics for evaluation include accuracy, precision, recall, and F1-score. From this dataset, we obtained a file containing the title, content, creation date, and the truth label (real or fake) for each article. Using this data, we processed all inputs according to the workflow shown in Figure 3.

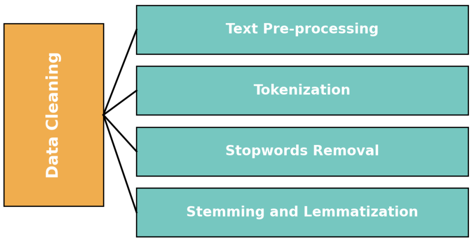

Figure 3. Data cleaning process: steps involved in preparing text data for analysis

The methodology begins with preprocessing the collected data, where all non-essential elements like links, emojis, numbers, and special characters are removed. The remaining text is then standardized to consist only of the alphabet. Following this, we tokenize the text into individual words and remove common stopwords that do not contribute meaningfully to fake news detection. Advanced techniques like stemming and lemmatization are applied to group words by their root forms. For example, ”boating” and ”boats” are reduced to their base form ”boat.” We also remove numbers and links (both HTTP and HTTPS).

Next, we tokenize each text by converting it into tokens. For example, the sentence ”The earth is round” would be tokenized into ”the,” ”earth,” ”is,” and ”round.” However, common words like ”the” and ”is” do not significantly contribute to the understanding of the sentence’s meaning. To reduce unnecessary complexity, we remove stopwords—words like ”the” and ”is”—leaving us with meaningful tokens such as ”earth” and ”round.”

We also apply stemming, which involves reducing words to their root form. For instance, a search for ”boat” will also return results for ”boats,” ”boating,” and ”boater,” as they all share the root ”boat.” In contrast, lemmatization takes into account the broader vocabulary of the language, converting words like ”was” to ”be” and ”mice” to ”mouse.”

After stopword removal and stemming, we proceed with vectorizing the text using term frequency-inverse document frequency (TF-IDF). The TF-IDF method assigns a weight to each word in the document based on how frequently it appears within the document term frequency (TF) and how rare the word is across the entire corpus inverse document frequency (IDF). This approach helps emphasize words that are important to a particular document but not common across many documents. The result is a numerical representation of the text, which can then be fed into machine learning algorithms for classification. For example, in the sentence ”earth is round,” after stopword removal and stemming, the important word ”earth” would have a higher TF-IDF score if it appears infrequently in the dataset, making it a strong feature for distinguishing fake news from real news.

After cleaning the dataset, we moved on to applying machine learning models to the processed data. The input to our models consists of the extracted word features from the news articles, with the output being the news validity (fake or real). We implemented various machine learning algorithms using Python and ran them on the same machine, producing the results shown in Table 1.

Table 1. Model comparison for fake news detection

Model	Accuracy (%)	Recall (%)	F1-score (%)	Duration (s)
Linear SVC	99.43	99.40	99.42	21.2639
Logistic regression	98.89	98.85	98.87	38.6611
Multinomial naive Bayes	93.66	93.50	93.58	22.0979
Bernoulli naive Bayes	95.84	95.70	95.77	22.7570
Gradient boost	99.61	99.60	99.60	159.8904
XGBoost	99.59	99.55	99.57	64.7849
Decision tree	99.55	99.50	99.52	41.5180
K nearest neighbour	90.43	90.00	90.21	101.2192
Aggressive classifier	99.38	99.30	99.34	20.4679

The results show that gradient boosting achieved the highest accuracy and F1-score, making it the top-performing model in terms of precision. However, when considering execution time, XGBoost emerges as a more practical choice, providing nearly the same accuracy but with significantly reduced processing time, as shown in Figure 4 and Table 1. This trade-off between speed and accuracy is critical for applications requiring real-time analysis. Additionally, the Linear SVC model, though slightly less accurate, offers the fastest execution, which can be beneficial in resource-constrained environments. The main challenges encountered were optimizing these models for both speed and accuracy, particularly when working with large datasets.

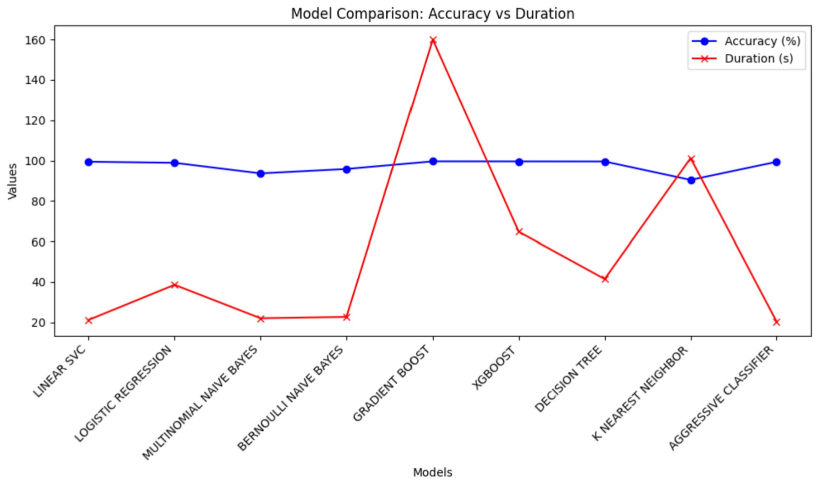

Figure 4. Comparison of machine learning models: accuracy vs. execution time for fake news detection

We cannot definitively say that gradient boosting is optimal model, as other models achieved similar accuracy with shorter execution times. It's also worth noting that execution time might not be the best metric for model comparison, given differences in system architectures. Performance can be improved by using GPUs or adding more cores to the CPU, potentially reducing execution time and further optimizing the models.

8. CONCLUSION

This research primarily focused on addressing the four fundamental challenges associated with big data: data collection, storage, analysis, and prediction. In the context of processing text data, we implemented various machine learning models to detect fake news on social media platforms. Through a comparative analysis of several models, we determined which model best met our requirements based on accuracy and performance. Our findings indicate that the gradient boosting model is the most effective, achieving an accuracy of 99.61% in predicting fake news. Additionally, the XGBoost model also demonstrated high performance, with an accuracy of 99.59%. XGBoost provided a strong balance between accuracy and execution time, making it an excellent choice for applications where both precision and speed are critical. The consistent performance of both models highlights their robustness in handling complex data and delivering reliable predictions. This study is limited by its focus on text-based fake news detection, which does not account for the multimedia elements often involved in misinformation, such as images or videos. Additionally, the dataset used, while comprehensive, is constrained to specific time periods and domains. Future research should explore integrating image and video analysis into fake news detection systems to improve accuracy further. Moreover, expanding the scope to real-time detection could provide more dynamic and responsive solutions for combating misinformation on social media platforms.

FUNDING INFORMATION

The authors declare that no specific funding was received for this research.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Bahaa Eddine Elbaghazaoui	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Mohamed Amnai		✓				✓		✓	✓	✓	✓	✓		
Youssef Fakhri	✓		✓	✓		✓			✓		✓		✓	✓
Ali Choukri	✓		✓	✓	✓				✓		✓		✓	
Noreddine Gherabi	✓		✓	✓		✓			✓		✓		✓	✓

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject Administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding Acquisition
Fo : **Fo**rmal Analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

The data used in this study is the ISOT fake news dataset, created by the Information Security and Object Technology (ISOT) research group at the University of Victoria. This dataset is publicly available and widely used for fake news detection tasks. It consists of two categories: real news and fake news. The dataset can be accessed at: https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets.

REFERENCES

[1] M. Cai, H. Luo, X. Meng, Y. Cui, and W. Wang, "Influence of information attributes on information dissemination in public health emergencies," *Humanities and Social Sciences Communications*, vol. 9, no. 1, pp. 1-22, 2022, doi: 10.1057/s41599-022-01078-3.

[2] E. C. Tandoc, "Fake news," in *The Routledge Companion to Media Disinformation and Populism*, London, United Kingdom: Routledge, 2021, pp. 110-117.

[3] S. V. D. Linden, "Misinformation: susceptibility, spread, and interventions to immunize the public," *Nature Medicine*, vol. 28, no. 3, pp. 460-467, 2022, doi: 10.1038/s41591-022-01713-6.

[4] J. George, N. Gerhart, and R. Torres, "Uncovering the truth about fake news: A research model grounded in multi-disciplinary literature," *Fake News on the Internet*, pp. 175-202, 2023.

[5] R. Rodríguez-Ferrándiz, "An overview of the fake news phenomenon: From untruth-driven to post-truth-driven approaches," *Media and Communication*, vol. 11, no. 2, pp. 15-29, 2023, doi: 10.17645/mac.v11i2.6209.

[6] P. N. Petratos, "Misinformation, disinformation, and fake news: Cyber risks to business," *Business Horizons*, vol. 64, no. 6, pp. 763-774, 2021, doi: 10.1016/j.bushor.2021.07.010.

[7] E. Aïmeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: a review," *Social Network Analysis and Mining*, vol. 13, no. 1, 2023, doi: 10.1007/s13278-023-00951-3.

[8] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Information Processing & Management*, vol. 57, no. 2, 2020, doi: 10.1016/j.ipm.2019.102025.

[9] D. Evans, S. Bratton, and J. McKee, *Social Media Marketing: the next generation of business engagement*, Hoboken, United States: A G Printing & Publishing, 2021.

[10] G. D. Domenico, J. Sit, A. Ishizaka, and D. Nunan, "Fake news, social media and marketing: A systematic review," *Journal of Business Research*, vol. 124, pp. 329-341, 2021, doi: 10.1016/j.jbusres.2020.11.037.

[11] J. P. Baptista and A. Gradim, "Understanding fake news consumption: A review," *Social Sciences*, vol. 9, no. 10, 2020, doi: 10.3390/socsci9100185.

[12] G. Pennycook and D. G. Rand, "The psychology of fake news," *Trends in Cognitive Sciences*, vol. 25, no. 5, pp. 388-402, 2021, doi: 10.1016/j.tics.2021.02.007.

[13] R. Armitage and C. Vaccari, "Misinformation and disinformation," in *The Routledge Companion to Media Disinformation and Populism*, London, United Kingdom: Routledge, 2021.

[14] I. P. Idoko, O. M. Ijiga, L. A. Enyejo, S. I. Ugbane, O. Akoh, and M. O. Odeyemi, "Exploring the potential of Elon musk's proposed quantum AI: A comprehensive analysis and implications," *Global Journal of Engineering and Technology Advances*, vol. 18, no. 3, pp. 48-65, 2024, doi: 10.30574/gjeta.2024.18.3.0462.

[15] O. D. Apuke and B. Omar, "Fake news and COVID-19: Modelling the predictors of fake news sharing among social media users," *Telematics and Informatics*, vol. 56, 2021, doi: 10.1016/j.tele.2020.101475.

[16] Y. M. Rocha, G. A. de Moura, G. A. Desidério, C. H. de Oliveira, F. D. Lourenço, and L. D. F. Nicolete, "The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review," *Journal of Public Health*, vol. 31, pp. 1007–1016, 2021, doi: 10.1007/s10389-021-01565-3.

[17] A. Mishra and H. Sadia, "A comprehensive analysis of fake news detection models: A systematic literature review and current challenges," *Engineering Proceedings*, vol. 59, no. 1, 2023, doi: 10.3390/engproc2023059028.

[18] E. Min *et al.*, "Divide-and-conquer: Post-user interaction network for fake news detection on social media," in *Proceedings of the ACM Web Conference*, 2022, pp. 1148-1158, doi: 10.1145/3485447.3512163.

[19] P. Meel and D. K. Vishwakarma, "Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities," *Expert Systems with Applications*, vol. 153, 2020, doi: 10.1016/j.eswa.2020.112986.

[20] C. Busioc, S. Ruseti, and M. Dascalu, "A literature review of NLP approaches to fake news detection and their applicability to Romanian language news analysis," *Revista Transilvania*, vol. 10, pp. 65-71, 2020, doi: 10.51391/trva.2020.10.07.

[21] K. Bryanov and V. Vziatysheva, "Determinants of individuals' belief in fake news: A scoping review," *PLOS ONE*, vol. 16, no. 6, 2021, doi: 10.1371/journal.pone.0253717.

[22] B. E. Elbaghazaoui, M. Amnai, and Y. Fakhri, "Predicting user behavior using data profiling and hidden Markov model," *International Journal of Electrical & Computer Engineering*, vol. 13, no. 5, pp. 4819-4827, 2023, doi: 10.11591/ijece.v13i5.pp5444-5453

[23] B. E. Elbaghazaoui, M. Amnai, Y. Fakhri, A. Choukri and N. Gherabi, "Event profiling in social media: A survey with an application," in *2024 4th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, FEZ, Morocco, 2024, pp. 1-5, doi: 10.1109/IRASET60544.2024.10548943.

[24] B. E. Elbaghzaoui, M. Amnai, and Y. Fakhri, "Predicting the next word using the Markov chain model according to profiling personality," *The Journal of Supercomputing*, vol. 79, no. 11, pp. 12126-12141, 2023, doi: 10.1007/s11227-023-04964-7.

[25] A. A. A. Ahmed, A. Aljabouh, P. K. Donepudi, and M. S. Choi, "Detecting fake news using machine learning: A systematic literature review," *arXiv-Computer Science*, pp. 1-10, 2021.

[26] R. K. Kaliyar, A. Goswami, and P. Narang, "DeepFakE: Improving fake news detection using tensor decomposition-based deep neural network," *The Journal of Supercomputing*, vol. 77, no. 2, pp. 1015-1037, 2021, doi: 10.1007/s11227-020-03294-y.

[27] S. Rathje, J. Roozenbeek, J. J. V. Bavel, and S. V. D. Linden, "Accuracy and social motivations shape judgements of (mis)information," *Nature Human Behaviour*, vol. 7, no. 6, pp. 892-903, 2023, doi: 10.1038/s41562-023-01594-9.

[28] M. T. Sadiq and S. K. Mathew, "The disaster of misinformation: A review of research in social media," *International Journal of Data Science and Analytics*, vol. 13, no. 4, pp. 271-285, 2022, doi: 10.1007/s41060-022-00321-3.

[29] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," *Security and Privacy*, vol. 1, no. 1, 2017, doi: 10.1002/spy2.9.

[30] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using N-Gram analysis and machine learning techniques," in *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, Springer, Cham, 2017, pp. 127-138, doi: 10.1007/978-3-319-69155-8_9.

BIOGRAPHIES OF AUTHORS

Bahaa Eddine Elbaghazaoui began his academic journey in mathematical sciences, earning his baccalaureate in this field. After completing preparatory classes in 2013, he advanced to the National School of Applied Sciences of Khouribga, where he pursued computer engineering, ultimately earning an engineering diploma in software engineering in 2019. In December 2023, he obtained his Ph.D. in computer sciences & artificial intelligence. With over 5 years of experience as a software engineer and full-stack developer, he continues to contribute to the field through his work and research. He is now an Assistant Professor at ENSA Beni Mellal. He can be contacted at email: elbaghazaoui.bahaa@gmail.com, bahaaeddine.elbaghazaoui@uit.ac.ma, or b.elbaghazaoui@usms.ac.ma.

Mohamed Amnai completed his undergraduate studies in computers, electronics, electrical, and automation at Molay Ismail University in Errachidia in 2000. He later obtained his master's degree in the same field from Ibn Tofail University in Kenitra in 2007. In 2011, he earned a Ph.D. in computer science and telecommunications from Ibn Tofail University. Since 2014, he has been serving as an Assistant Professor at the National School of Applied Sciences Khouribga of Settat University. In 2018, he was appointed as an Associate Professor in the Department of Computer Science and Mathematics at the Faculty of Sciences in Kenitra, Ibn Tofail University. He is also a member of the Networks and Telecommunications Team at the Kenitra Faculty of Science and a researcher in the LaRIT Laboratory. He can be contacted at email: mohamed.amnai@uit.ac.ma.

Youssef Fakhri obtained his Bachelor of Science degree in electronic physics from the University Mohammed V's Faculty of Sciences in Rabat in 2001. He then pursued a master's degree in computer and telecommunication, where he completed a project with the Moroccan ICI Company in 2003. In 2007, he received his Ph.D. from the University Mohammed V-Agdal in Rabat, Morocco, in collaboration with the Polytechnic University of Catalonia in Spain. Following his academic achievements, he joined Ibn Tofail University's Faculty of Sciences in Kenitra as an Associate Professor in 2009, where he currently teaches in the Department of Computer Science and Mathematics. He also serves as an associate researcher at the Rabat Faculty of Sciences and is the Laboratory Head at LaRIT. He can be contacted at email: fakhri@uit.ac.ma.

Ali Choukri is an Assistant Professor at the National Academy of Applied Sciences. He received his master's degree in computer science and telecommunications from the University of Ibn Tofail, Kenitra, Morocco, in 2008, and in 1992, he earned a degree from ENSET (Higher Normal School of Technical Teaching). He holds a Ph.D. from the School of Computer Science and Systems Analysis (ENSIAS). He is part of the MIS team in the SIME Laboratory, where his research focuses on mobile intelligent ad hoc communication systems and wireless sensor networks. His research interests span a wide range of topics, including ubiquitous computing, the internet of things, delay/fault-tolerant networks, wireless networks, QoS routing, mathematical modeling and performance analysis of networks, control and decision theory, game theory, trust and reputation management, distributed algorithms, meta-heuristics and optimization, and genetic algorithms. He can be contacted at email: choukriali@gmail.com.

Noreddine Gherabi is a Professor of computer science with extensive experience in both academia and industry. He holds a doctorate in computer science and has been teaching at Mohamed Ben Abdellah University since 2013, and at Sultan Moulay Slimane University in Morocco as a research professor since 2015. A member of the International Association of Engineers, he has made significant contributions to information systems, focusing on big data, semantic web, pattern recognition, and intelligent systems. He has numerous publications in computer science, including book chapters, journal articles, and conference papers, and has edited several books. He has also chaired and convened over 52 conferences and workshops. His recent books published by Springer include intelligent systems in big data, semantic web, and machine learning, advances in information, communication, cybersecurity, and information technology and communication systems. He can be contacted at email: gherabi@gmail.com or n.gherabi@usms.ma.