Face News Detection Using Machine Learning Techniques

R. Sai Venkat[1], Ramu kuchipudi[1], K. Gangadhara rao[1], G. Srikanth[1], Palamakula Ramesh babu[1], T. Satyanarayana Murthy[1(✉)], and G. Venakata Kishore[2]

[1] Chaitanya Bharathi Institute of Technology, Hyderabad, India
{saivenkatr_it,ramukuchipudi_it,kgangadhar_it,srikanthg_it, prameshbabu_it,tsmurthy_it}@cbit.ac.in
[2] Srinidhi Institute of Technology, Hyderabad, India
kishore.g@sreenidhi.edu.in

Abstract. Nowadays, most people are shifting to online news reading rather than traditional methods because of its convenience and low cost. So, the number of online users is increasing day by day, which is also increasing fake news across the internet. This fake news must be detected and removed from the internet before causing damage to the nation's peace. Social media platforms use fake news recognition algorithms to detect and remove misinformation from their platforms. However, these algorithms can fail when they misidentify satirical or humorous content as fake news. News organizations use fake news recognition technology to fact-check articles and ensure the accuracy of their reporting. However, these algorithms may fail to detect more sophisticated forms of misinformation, such as deep fakes or highly persuasive disinformation campaigns, leading to the spread of false information. Existing framework contains three main modules: information retrieval, natural language processing, and machine learning. Also has two phases: the data collection phase and the machine learning model-building phase. In the data collection phase, we obtained a data set and analyzed the data using natural language processing techniques to extract good features from web data. A detailed survey on an automatic online fake news detection using machine learning techniques are elaborated.

Keywords: Machine Learning · Fake News · learning · optimization · deep learning

1 Introduction

These days fake news is creating different issues from sarcastic articles to fabricated news Fake news encompasses a range of issues in today's society, extending from sarcastic articles and fabricated news to planned government propaganda. These problems have significant ramifications, contributing to a growing lack of

O. Castillo et al. (Eds.): PerSOM 2023, LNICST 517, pp. 311–318, 2024.
https://doi.org/10.1007/978-3-031-66044-3_23

trust in the media and distorting public discourse. While the term "fake news" traditionally referred to purposely misleading stories, its definition has evolved, and some individuals now misuse it to dismiss facts that contradict their preferred viewpoints. The impact of disinformation in American political discourse gained considerable attention, particularly in the aftermath of the American presidential election. The phrase "fake news" became widely used to describe factually incorrect and misleading articles primarily created to generate revenue through page views. Consequently, researchers and experts have sought to develop models capable of accurately predicting the likelihood that a given article is fake news. Social media platforms like Facebook have faced significant criticism about fake news. Facebook has implemented features to flag potentially misleading content on its site when users encounter it. Furthermore, the company has publicly acknowledged its efforts to develop automated systems to distinguish fake news articles. However, achieving this goal is challenging. An effective algorithm must remain politically unbiased, as fake news exists across the political spectrum. Additionally, it should provide equal consideration to legitimate news sources regardless of their ideological leanings. Determining the legitimacy of news articles presents its difficulties. Addressing the problem of fake news requires a comprehensive understanding of its nature. Fake news encompasses intentionally deceptive information, sarcastic content, and misleading narratives that manipulate public opinion. It is essential to distinguish between legitimate journalism and fabricated or misleading stories. Initiatives to combat fake news involve a combination of technological advancements, media literacy programs, and responsible journalism practices. Developing algorithms that can accurately identify fake news is a complex task. It requires leveraging artificial intelligence and machine learning while incorporating human fact-checking to ensure accuracy and fairness. Educating individuals on media literacy and critical thinking equips them with the necessary tools to discern reliable sources of information and identify misinformation. Collaboration among technology companies, factchecking organizations, and policymakers is crucial. Transparent policies and practices should be implemented to promote accurate information sharing and discourage the dissemination of fake news. Additionally, addressing the underlying economic incentives that drive the creation and proliferation of fake news is essential for long-term solutions. Solving the fake news problem necessitates a collective effort and a commitment to promoting an informed and discerning society. By fostering media literacy, leveraging technology responsibly, and upholding ethical journalism practices, we can mitigate the negative impact of fake news and preserve the integrity of public discourse.

2 Literature Review

In this experimental study [1], the authors aim to unify misinformation detection across different domains, with a primary focus on recognizing fake news related to news articles. They use a limited set of machine learning algorithms to achieve this, bringing together multiple domains of misinformation detection

under a single setup. This paper [2] leverages blockchain techniques in machine learning to enhance the tracking of fake news on social media. It effectively utilizes a blockchain approach to improve the efficiency of tracking fake news; however, it has limitations in identifying fake images and deals primarily with text-based content. The authors present an empirical study on the evolution of fake news, providing insights into the transformation of fake news over time. This research [3] contributes to a better understanding of the evolution of fake news, making the identification process more efficient. It is important to note that this study primarily explains the evolution of fake news and does not delve into creating machine learning models using this dataset. This paper [4] conducts a systematic literature review on the topic of detecting fake news using machine learning techniques. It primarily focuses on supervised machine learning classifiers and their role in identifying fake news. These classifiers require labeled data for training, which can be a challenge in practice. The review underscores the importance of labeled data in training classifiers, but it also highlights the potential of unsupervised machine learning approaches to address this limitation . The authors present a novel stacking approach for accurate detection of fake news. In this paper [5], they evaluate the performance of various machine learning and deep learning models to enhance the accuracy of fake news detection. One notable advantage of this approach is its suitability for multiple languages, making it a versatile tool for different regions. However, it's important to mention that the paper primarily discusses a limited set of machine learning and deep learning languages, which might not cover all possibilities. In the realm of cybersecurity, this paper [6] introduces an efficient hybrid system for anomaly detection in social networks. The authors develop a hybrid anomaly detection method named DT-SVMNB that combines several machine learning algorithms to classify normal and abnormal users in social networks. The focus of this system is detecting fake users primarily based on the contents within the user profiles. An additional advantage lies in the potential to further enhance accuracy by exploring user interests and activities. This paper [7] explores construction site accident analysis using text mining and natural language processing techniques. It delves into various natural language processing and text mining methods to identify the causes of accidents on construction sites from their reports. The approach offers the potential for further accuracy improvements through the exploration and integration of new algorithms. This survey paper [8] provides insights into neural network language models, focusing on different architectures of classic neural network language models and their enhancements. Additionally, it introduces the related corpora and toolkits necessary for studying neural network language models. However, it's important to note that the paper primarily explains the tasks and advancements related to neural network language models without delving into their real-time applications. This paper [9] introduces an efficient convolutional neural network-based word segmenter for the Thai language, aiming to enhance word segmentation. The model utilizes character and syllable embeddings, resulting in a segmenter that is 5.6 times faster than previous state-of-the-art solutions. However, the approach faces challenges with

idiosyncratic datasets, particularly those in the domain of poetry. In the realm of fake news detection, this paper [10] reports an accuracy of 93percent through deep learning techniques. While achieving high accuracy, this approach highlights the importance of collecting more recent data, especially data relevant to the current period, to further improve performance. The paper demonstrates the effectiveness of deep learning techniques in fake news detection. This work [11] introduces the use of machine learning techniques to detect fake news. The model aims to identify fake news using machine learning methods, offering a potential solution to the issue of fake news proliferation. However, the paper primarily discusses the application of these techniques without extensively detailing their potential use cases. In this paper [12], the authors propose the TriFN model, which employs a tri-relationship embedding framework to model publisher-news relations and user-news interactions simultaneously for fake news classification. This approach enriches the understanding of the social context in which fake news spreads, contributing to improved detection. This paper [13] takes a data management and mining perspective to combat fake news. It introduces a warning system to reduce the spread of fake news by alerting users to potentially false information. While the warning system is a valuable addition, it does not entirely eliminate the challenge of fake news. This paper [14] focuses on fake news detection with generated comments for news articles. The model employs specific algorithms and methods to detect fake news within news articles. It is effective for this context but may not be applicable to fake news in other social media content Title Suppressed Due to Excessive Length 5 This work [15] concentrates on fake news detection using machine learning ensemble methods. The model excels in identifying fake news through the application of a limited set of machine learning algorithms. It is particularly wellsuited for text-based content, but it does not extend to fake news detection in images or other non-text formats. In the field of subsurface characterization, this paper [16] addresses noninvasive fracture characterization based on the classification of sonic wave travel times. The model's main advantage lies in its ability to classify fractures based on the waves generated, making it a valuable tool in this domain. However, it requires an experimental setup, involving external elements and additional work for implementation. This paper [17] explores bone cancer detection using machine learning techniques. The model efficiently detects diseases using machine learning algorithms, offering a swift diagnostic solution. However, one limitation is that it relies on a relatively limited dataset for identifying diseases, which may impact its accuracy in some cases. This paper [18] discusses Bayesian classifier combination as a method for classification tasks. The advantage of this method lies in its ability to classify items based on labels, contributing to effective categorization. However, it is important to note that this method primarily focuses on classification and does not offer detection capabilities. This paper [19] introduces a system for predicting power production of wind turbines using a fusion of multilayer perceptron (MLP) and adaptive neuro-fuzzy inference systems (ANFIS). The system is capable of detecting underperformance of turbines for various operating conditions. However, it relies on weather conditions to make

predictions and may not be as reliable in unfavorable weather conditions. The paper [20]. presents an LSTM-based predictive framework for literature-based knowledge discovery. The model features supportive attributes for predicting future literature-based discoveries and emerging trends. It is a valuable tool; however, it may face challenges related to accommodating domain experts in identifying semantic similarity of keywords and various forms of abbreviations.

3 Conclusion

In conclusion, the detection of fake news using machine learning techniques holds great potential in combating the proliferation of misinformation in today's digital age. Machine learning algorithms offer the ability to analyze vast amounts of data, extract meaningful patterns, and make accurate predictions about the authenticity of news articles. Through the utilization of supervised, unsupervised, and semi-supervised learning methods, machine learning models can learn from labeled and unlabeled datasets to identify distinguishing features and characteristics of fake news. By training on diverse and comprehensive datasets that encompass various types of fake news, these models can generalize well and improve their detection accuracy. Feature engineering plays a critical role in the effectiveness of machine learning-based fake news detection. By extracting relevant features from textual content, including linguistic cues, lexical patterns, sentiment analysis, and meta-data, models can capture the nuances that differentiate fake news from genuine information. Additionally, incorporating user-based features such as social network interactions, user behavior, and credibility of sources can enhance the detection process. However, challenges remain in the field of fake news detection using machine learning. The constantly evolving nature of fake news requires continuous updates and adaptations of the detection models to stay ahead of emerging techniques used by those spreading misinformation. Biases present in the training data, including both explicit and implicit biases, need to be addressed to ensure fair and unbiased detection. Furthermore, the interpretability and transparency of machine learning models are crucial to building trust and understanding of their decision-making process. To overcome these challenges, interdisciplinary research efforts are needed to combine expertise from fields such as natural language processing, information retrieval, social network analysis, and cognitive psychology. Collaboration between academia, industry, and policymakers is essential to develop robust frameworks, share datasets, and establish guidelines for responsible information dissemination and consumption. In conclusion, machine learning offers a promising avenue for the detection of fake news, leveraging its capabilities to process vast amounts of data and extract meaningful insights. By addressing challenges related to biases, interpretability, and the dynamic nature of fake news, machine learning-based detection systems can play a vital role in promoting accurate information, fostering media literacy, and safeguarding the integrity of public discourse.

References

1. Rahman, M.S., Halder, S., Uddin, M.A., Acharjee, U.K.: An efficient hybrid system for anomaly detection in social networks. Cybersecurity **4**(1), 10 (2021)
2. Jiang, T., Li, J.P., Haq, A.U., Saboor, A., Ali, A.: A novel stacking approach for accurate detection of fake news. IEEE Access **9**, 22626–39 (2021)
3. Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., Dyer, C.: Neural architectures for named entity recognition (2016)
4. Phatthiyaphaibun, W., et al.: Pythainlp v2.3.1 release! (2021)
5. Ahmad, I., Yousaf, M., Yousaf, S., Ahmad, M.O.: Fake news detection using machine learning ensemble methods. Complexity **2020**(1), 8885861 (2020)
6. Misra, S., Li, H.: Chapter 9-noninvasive fracture characterization based on the classification of sonic wave travel times, pp. 243–287. Gulf Professional Publishing (2020)
7. Shrivastava, D., Sanyal, S., Maji, A.K., Kandar, D.: Chapter 17 - Bone cancer detection using machine learning techniques, pp. 175–183. Academic Press (2020)
8. Kim, H.C., Ghahramani, Z.: Bayesian classifier combination. In: Artificial Intelligence and Statistics, pp. 619–627 (2012)
9. Choudhury, N., Faisal, F., Khushi, M.: Towards an LSTM-based predictive framework for literature-based knowledge discovery (2019)
10. Zhang, F., Fleyeh, H., Wang, X., Lu, M.: Construction site accident analysis using text mining and natural language processing techniques. Autom. Constr. **99**, 238–248 (2019)
11. Jing, K., Xu, J.: A survey on neural network language models (2019)
12. Parikh, S.B., Atrey, P.K.: Media-rich fake news detection: a survey. In: 2018 IEEE Conference on Multimedia Information Processing and Retrieval, pp. 436–441 (2018)
13. Guo, M., Chen, X., Li, J., Zhao, D., Yan, R.: How does truth evolve into fake news? an empirical study of fake news evolution (2021)
14. Ahmed, A.A.A., Aljabouh, A., Donepudi, P.K., Choi, M.S.: Detecting fake news using machine learning: a systematic literature review (2021)
15. Shu, K., Wang, S., Lee, D., Liu, H.: Mining disinformation and fake news: concepts, methods, and recent advancements (2020)
16. Lee, N., et al.: On unifying misinformation detection (2021)
17. Mookdarsanit, P., Mookdarsanit, L.: The COVID-19 fake news detection in Thai social texts. EAI Endorsed Trans. Energy Web **10**(2), 988–998 (2021)
18. Akhter, M.P., Zheng, J., Afzal, F., Lin, H., Riaz, S., Mehmood, A.: Supervised ensemble learning methods towards automatically filtering Urdu fake news within social media. PeerJ Comput. Sci. **7**, e425 (2021)
19. Umer, M., Imtiaz, Z., Ullah, S., Mehmood, A., Choi, G.S., On, B.W.: Fake news stance detection using deep learning architecture (CNN-LSTM). IEEE Access **8**, 156695–706 (2020)
20. Yanagi, Y., Orihara, R., Sei, Y., Tahara, Y., Ohsuga, A.: Fake news detection with generated comments for news articles. In: IEEE 24th International Conference on Intelligent Engineering Systems (INES), pp. 85–90 (2020)
21. Satyanarayana Murthy, T., Gopalan, N.P., Alla, D.S.K.: The power of anonymization and sensitive knowledge hiding using sanitization approach. Int. J. Modern Educ. Comput. Sci. (IJMECS), **10**(9), 26–32 (2018). https://doi.org/10.5815/ijmecs.2018.09.04

22. Satyanarayana Murthy, T., Gopalan, N.P.: A novel algorithm for association rule hiding. Int. J. Inf. Eng. Electron. Bus. (IJIEEB) **10**(3), 45–50 (2018). https://doi.org/10.5815/ijieeb.2018.03.06
23. Saibabu, A., Murthy, T.S.N.: Security provision in publicly auditable secure cloud data storage services using SHA-1 algorithm. IJCSIT Int. J. Comput. Sci. Inf. Technol. **3**(3), 4084–4088 (2012)
24. Sathyanarayana Murthy, T., Mohan Krishna Varma, N., Ravuri, D., Kishore Babu, D., Nazeer, S.: Classification of Precious and non-precious tweets using deep learning. In: Rout, R.R., Ghosh, S.K., Jana, P.K., Tripathy, A.K., Sahoo, J.P., Li, KC. (eds.) Advances in Distributed Computing and Machine Learning, vol. 427, pp. 393–399. Springer, Singapore (2022). https://doi.org/10.1007/978-981-19-1018-0_33
25. Satyanarayana Murthy, T., Mohan Krishna Varma, N., Roy, S., Nazeer, S.: Effective classification of Tweets using machine learning. In: Kumar, R., Ahn, C.W., Sharma, T.K., Verma, O.P., Agarwal, A. (eds.) Soft Computing: Theories and Applications. LNNS, vol. 425, pp. 439–446. Springer, Singapore (2022). https://doi.org/10.1007/978-981-19-0707-4_40
26. Murthy, T.S., Gopalan, N.P., Ramachandran, V.: A naive bayes classifier for detecting unusual customer consumption profiles in power distribution systems - APSPDCL. In: 2019 Third International Conference on Inventive Systems and Control (ICISC) at JCT College, Coimbatore, India, pp. 673–678 (2019)
27. Satyanarayana Murthy, T., Preethi, G., Gopalan, N.P.: An efficient way of anonymization without subjecting to attacks using secure matrix method. In: proceedings of the IEEE International Conference on Intelligent Computing and Control Systems at VAIGAI COLLEGE OF ENGG,MADURAI, pp. 1462–1465 (2018)
28. Satyanarayana Murthy, T., Gopalan, N.P.: An efficient meta-heuristic chemical reaction based algorithm for association rule hiding using an advanced perturbation approach. In: proceedings of the IEEE International Conference on Intelligent Computing and Control Systems, at VAIGAI COLLEGE OF ENGG,MADURAI. Indexed in IEEE (2018)
29. Gopalan, N.P., Satyanarayana Murthy, T.: Association rule hiding using chemical reaction optimization. In: Presented a paper at 7th International Conference on Soft Computing for Problem Solving - SocProS 2017, December 23-24, 2017, IIT Bhubaneswar, ORISSA (2017)
30. Devarajan, D., et al.: Cervical cancer diagnosis using intelligent living behavior of artificial jellyfish optimized with artificial neural network. IEEE Access **10**, 126957–126968 (2022)
31. Maheswari, V.U., Aluvalu, R., Kantipudi, M.P., Chennam, K.K., Kotecha, K., Saini, J.R.: Driver drowsiness prediction based on multiple aspects using image processing techniques. IEEE Access **10**, 54980–54990 (2022)
32. Satyanarayana Murthy, T., Gopalan, N.P., Balaji, B.: A modified un-realization approach for effective data perturbation. Int. J. Intell. Enterp. Inder Sci. 408–421 (2023). https://doi.org/10.1504/IJIE.2023.10054103
33. Satyanarayana Murthy, T., Udayakumar, P., Alenezi, F., Laxmi Lydia, E., Ishak, M.K.: Coot optimization with deep learning-based false data injection attack recognition. Comput. Syst. Sci. Eng. **46**(1), 255–271 (2023)
34. Yonbawi, S., Alahmari, S., Satyanarayana Murthy, T., Maddala, P., Laxmi Lydia, E., et al.: Harris hawks optimizer with graph convolutional network-based weed detection in precision agriculture. Comput. Syst. Sci. Eng. **46**(2), 1533–1547 (2023)

35. Yonbawi, S., Alahmari, S., Murthy, T.S., Daniel, R., Lydia, E.L., et al.: Modified metaheuristics with transfer learning based insect pest classification for agricultural crops. Comput. Syst. Sci. Eng. **46**(3), 3847–3864 (2023)
36. Ahmed, M.A., Murthy, T.S., Alenezi, F., Lydia, E.L., Kadry, S., et al.: Design of evolutionary algorithm based unequal clustering for energy aware wireless sensor networks. Comput. Syst. Sci. Eng. **47**(1), 1283–1297 (2023)
37. Devaraj, F.S., Satyanarayana Murthy, T., Alenezi, F., Laxmi Lydia, E., Md Zawawi, M.A., et al.: Enhanced metaheuristics with trust aware route selection for wireless sensor networks,". Comput. Syst. Sci. Eng. **46**(2), 1431–1445 (2023)
38. Kalyani, K., Parvathy, V.S., Abdeljaber, H.A.M., Murthy, T.S., Acharya, S., et al.: Effective return rate prediction of blockchain financial products using machine learning. Comput. Mater. Continua **74**(1), 2303–2316 (2023)
39. Satyanarayana Murthy, T.: An efficient diabetic prediction system for better diagnosis. Int. J. Intell. Enterp. 408–421 (2022). https://doi.org/10.1504/IJIE.2022.126397,
40. Satyanarayana Murthy, T., Gopalan, N.P., Athira, T.R.: Hiding critical transactions using modified un-realization approach. Int. J. Bus. Intell. **15**(3), 223–234 (2020)
41. Navaneetha Krishnan, S., Sundara Vadivel, P., Yuvaraj, D., Satyanarayana Murthy, T., Malla, S.J., et al.: Enhanced route optimization for wireless networks using meta-heuristic engineering. Comput. Syst. Sci. Eng. **43**(1), 17–26 (2022)
42. Shanmuga Priya, S., Yuvaraj, D., Satyanarayana Murthy, T., Chooralil, V.S., Navaneetha Krishnan, S., et al.: Secure key management based mobile authentication in cloud. Comput. Syst. Sci. Eng. **43**(3), 887–896 (2022)
43. Satyanarayana Murthy, T., Varma, M.K., Yadav, A.K.: A diabetic prediction system based on mean shift clustering. Nutrition **36**(2), 231–235 (2021)
44. Satyanarayana Murthy, T., Varma, M.K., Harsha.: Brain tumour segmentation using U-net based adversarial networks. Traitement du Signal, 36(4), 353–359 (2021). https://doi.org/10.18280/ts.360408
45. Satyanarayana Murthy, T., Banothu, B., Varma, M.K.: An un-realization algorithm for effective privacy preservation using classification and regression trees, Revue d'Intelligence Artificielle, vol. 33(4), 313–319 (2019). https://doi.org/10.18280/ria.330408.
46. Satyanarayana Murthy, T., Varma, M.K., Roy, S.: Improving the performance of association rules hiding using hybrid optimization algorithm. J. Appl. Secur. Res. **15**(3), 423–437 (2020). https://doi.org/10.1080/19361610.2020.1756155
47. Satyanarayana Murthy, T., Gopalan, N.P., Yakobu, D.: An efficient un-realization algorithm for privacy preserving decision tree learning using McDiarmid's bound. Int. J. Innovative Technol. Exploring Eng. (IJITEE) **8**(4S2), 499–502 (2019)
48. Satyanarayana Murthy, T., Gopalan, N.P., Gunturu, S.: A novel optimization based algorithm to hide sensitive item-sets through sanitization approach. Int. J. Modern Educ. Comput. Sci. (IJMECS) **10**(10), 48–55 (2018). https://doi.org/10.5815/ijmecs.2018.10.06