International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Review Paper on Analyzing Machine Learning Enabled Fake News DetectionTechniques

Harsh Dubey(0225CS20MT08)

(Department of CSE, Global Nature Care Sangathan's Group of Institutions, Jabalpur)

Abstract

Fake news, or fabric that appears to be untrue with the issue of open deception, has long evolved into ubiquity. Dissemination of such data breeds social cohesion and well-divided political divisions and doubts in government. With so much news being disseminated through social media, human confirmation has become incomprehensible, leading to the refinement and adjustment of robotized strategies for identifiable evidence of false news. Fake news publishers use various stylistic techniques to increase the popularity of their works, one of which is to arouse the emotions of the readers. Because of this, sentiment analysis of text analytics, which determines the polarity and intensity of emotions expressed in text, is now being used as a basis for systems or as a complementary component in false news detection methods. This assessment analyzes the full explanation of false news identification. The study also emphasizes the characteristics, characteristics, classification, different types of data in the news, categories of false news and methods for detecting fake news.

1. Introduction

In order to describe fake news as any material capable of making readers believe in information that is not real, one must first define what is false news [1]. Spreading false news causes widespread harm to society and the individual. Initially, this type of false news has the potential to change or destroy the balance of authenticity in the news ecosystem. Due to the nature of fake news, people are forced to accept false or skewed ideas that they would otherwise reject [2]. Political messages or influences are often communicated through the use of false news and propaganda [3]. Fake news has a lasting effect on how people interact and respond to true news. To minimize the harmful effects of false news, it is important to develop a system that can automatically detect when it appears on social media [4]. However, there are many difficult research issues with the discovery of fake news on various social platforms. The various research objectives observed in this regard include identifying the source of origin or uploading specific news or data on social networks, understanding the real purpose or meaning of the uploaded data and determining the extent of authenticity. And validate it for decision making so that it can be considered genuine or counterfeit. News features make it a difficult task to automatically detect fake news. Initially, readers are deceived by false news, which makes it impossible to tell the difference between real and fake information [5]. When it comes to fake news, there is a wide range of formats and topics to choose from. These false reports attempted to distort the facts using different language approaches [6]. Current knowledge bases fail to effectively validate false news when it is linked to time-critical events because there are not enough supporting claims or facts to support them [7]. Data generated by false news (i.e., unstructured, noisy, incomplete and big data) is also on social media [8]. Researchers have tried in recent years to expose problems with their credibility on false news, social media, especially Twitter, YouTube, Facebook and television [9]. The data is used to evaluate political / product views, user sentiments, natural phenomena in the process, global events, and healthcare service customer satisfaction [10].

2. Fake News Detection

2.1. Definition. Fake news is deliberately written misleading material meant to deceive the public. Authenticity and purpose are the two most important aspects of this concept. Fake news has two characteristics: firstly, it contains incorrect material that could be confirmed as such, but secondly, it is produced with the dishonest goal of misleading readers. The distribution of false material through social media may have important implications, such weakening public faith in the news ecosystem, hurting an individual's or organization's reputation, or causing fear among the general public, all of which can affect society's stability

The methods used to manipulate information differentiatereal news from fake news. Alternatively, news material may use deceptive tactics such as fabricating facts to make the customer believe something they do not want to believe. It is also possible to impose material that seems to be from reputable sources, but the sources are not. Additionally, fraudulent features of fake news include the use of altered material, such as headlines and pictures that do not match the information delivered or the contextualization of fake news using real components and information but in a misleading context.

3. *Types.* Fake news detection may be split into three categories as follows:

(a) Fabrication: fabricated news is a deliberate omission of information that typically only comes from a single source. The source is likely aware of the story's inaccuracy. Clickbait is critical to the success of fake news stories

(b) Hoax: this kind of reporting employs more complex deception techniques to mislead the public. Multiple outlets disseminate fake news. It is possible that some people consider the tale to be real. This kind of news may be found on a variety of sites, such as the fake news about Donald Trump that circulated throughout the election on different social media platforms, including Twitter, Facebook, and blogs, so the general public is more likely to believe it

(c) Satire: faked news that is presented as humorous by the source. In the case of sharing satire with individuals who are not acquainted with the material's origins. Some people may mistakenly believe it to be true

4. Fake News Approaches

4.1. Approaches. The following are the results of a false news detection study carried out by a variety of researchers using avariety of approaches:

(a) Knowledge-based: approaches based on knowledge, such as fact-checking in news reports with the assistance of additional sources owing to fact-checking, an article's statements may ascribe a correct value in light of the mitigating circumstances. Approachesto fact-checking may be divided into three types: expert-driven, crowd sourcing-driven, and computationally driven.

(b) Style-based: based on the writing style, style-based methods identify false news. For the most part, there are two major types of style-based methods: one that is deception-oriented and the other of which is objectivity-oriented. Deception-oriented is concerned with assertions or claims in news material that are false or misleading. Approaches that focus on objectivity look for style cues that suggest that news reporting have skewed toward sensationalism and bias

(c) Stance-based: stance-based methods leverage customers' views from appropriate post contents to verify the authenticity of the actual news reports. When representing the stance of customers, one has the option of using explicit or implicit methods. Approaches based on stance in a social setting

(d) Propagation-based: based on propagation, these methods looked at how misinformation propagated on social media and the relationships between postings to estimate news credibility. Approaches focused on propagation in a social environment [13].

5. Features for Fake News Detection

Numerous studies have used feature-based classification to better identify false news stories. False information may be detected with ease using textual characteristics. The following sections go through a few of the features [15].

(a) Semantic features: semantic features capture the semantic (meaning) aspect of the text. These features derive a meaningful pattern from the data

(b) Lexical features: lexical features are mainly used in tf-idf vectorization for summarizing the total num- ber of unique words and the frequency of the word. Lexical features include pronouns, verbs, hash tags, and punctuation

(a) Sentence-level features: these features include a bag-of-word approach, part-of-speech, and n-gram approach. Sentence level features are the language feature which is mostly used in text classification

(b) Psycholinguistic features: these features and word count is based on dictionary-based text mining software.

6. Taxonomy of Fake News Detection

Fake news is exposed to use a variety of detecting methods drawn from different networks and databases. First, a breakdown of the different networks is provided:

6.1. Platforms. Carrier platforms are used to offer fresh material to end consumers, and a list of the most prominent carriers is given in this section. The most popular operating systems are shown as follows:

(a) Standalone website: any website can submit new tales, and everyone will have its URL. These URLs are being used directly by users if they wish to share or publish a social media post. Most websites fall into one of the three categories: blogs, media, or prominent news websites. Famous new sites, with their very own social media presence, are the ones that provide genuine material. Blog sites, that heavily rely on user-generated material and unsupervised content, are prime locations for spreading misinfor- mation. In accordance with media-rich content, media sites enable customers to create their websitesby creating material depending on style and cus- tomer interests

(b) Social media: a most popular method for disseminating the information on these websites is via the spreading of it. A daily news source is shared by almost 70% of its consumers. Consumers distribute the data using the most popular social networking platforms, like Twitter, Facebook, and WhatsApp. You may reach a bigger audience using false information if you produce sponsored advertisements for every Facebook post in which users submit mes-sages of restricted character length and then distrib- ute them with the other Twitter users by tweeting them back to them

(c) Emails: users trust email as a reliable medium for receiving news; however, verifying the legitimacy of news emails is a difficult problem

(d) Broadcast networks (Podcast): just a tiny percent- age of people use podcasting for anything other than news, making it a specialized subset of sound multimedia Radio service: the verification of sound veracity is asignificant challenge for radio services since these services are efficient news sources for the community.

7. Types of Fake News

Researchers in the social sciences examine fake news from avariety of angles before coming up with a broad classification of different kinds of fake news. This classification was given in the following statement:

(a) Visual-based: the types of fake news are described in the material using a graphical depiction of video or photo shopped pictures or a mix of both [10]

(b) User-based: using this method, the intended audience could be attracted by establishing fictitious accounts that reflect certain demographics such as gender, age, and culture [16]

(c) Post-based: social media sites like Facebook posts with video or picture captions, memes, tweets, and so on are the common places for this kind of fake news to emerge [17]

(d) Network-based: there are some people of an organization who are linked to this type of fake news, where this concept is primarily used to groups of linked persons on LinkedIn and friends-of-friends on Facebook [18]

Knowledge-based: these new articles will be created using articles that provide plausible explanations or scientific knowledge about an unsolved problem to disseminate false info [19]

(e) Style-based: false news may be produced by anybody with the ability to write in a variety of styles, but this style-based news was only concerned with how the false info was presented to end consumers [20]

8. Detection Methods for Identifying Fake News

There are many current methods created depending on feature extraction for characterizing false news and also different kinds of news data in the preceding section. The next paragraphs describe several feature-based approaches.

8.1. Linguistic Feature-Based Methods. Mishra et al. [21] utilizing linguistic-based methods, the essential language characteristics may be gleaned. In addition to syntax and grams, there are a variety of other characteristics like punctuation and readability, with the most significant of these being represented as follows:

(ii) n-grams: for the purpose of finding unigrams and bigrams in a narrative, writers gather words. TFIDF (term frequency inverse document frequency) is used to save and retrieve the extracted characteristics

(iii) Punctuation: the FND methods use punctuation to show the distinction between honest and fraudu- lent texts

(iv) Syntax: context-free grammar (CFG) states thatthis approach extracts several characteristics. Inline with the lexicalized production rules, which

Deception Modeling-Based Methods. When it comes to grouping true vs misleading tales, two theoretical methods come into play: vector space modeling (VSM) and Rhetori- cal Structure Theory (RST). Each text in a hierarchical tree can be analyzed utilizing RST to uncover rhetorical connec-tions. The VSM is used to identify the RS relations' out- comes, and that is all. RST-VSM technique offers curating data edge depending on the distance among samples as con-trasted using similarity-based cluster analysis [22].

Clustering-Based Methods. Using the Graphical Cluster-ing Toolkit clustering package, writers have utilized cluster-ing to assist distinguish among reports that employ the same clustering methodology [23]. When utilizing this technique,a large no. of sets are processed, and a limited number of clusters are formed/sorted employing hierarchical clustering and k-nearest neighbor technique, grouping comparable news items depending on a normalized frequency of connections. To determine if a new narrative is deceptive, researchersuse the concept of computed coordinate distances. It appearsthat this technique, depending on the author's assertion of a success rate of 63 percent, is particularly effective on big datasets. There is a chance that if this method is used on current false news, it may not offer reliable results since comparable news story sets may not be accessible, i.e., similar data may not be accessible [24].

Nontext Cue-Based Methods. The nontext substance of news is utilized to persuade readers to place their trust in tainted information, which is the primary goal of this method. Several analyses are employed here, including two which are classified as:

Image analysis: by employing a well-known keytechnique, or strategic use of pictures, the aim is to influence viewers' emotions

User behavior analysis: a content-independent tech- nique known as user behavior analysis is used to evaluate reader behavior (e.g., how they interacted with news). The method's primary goal is to better understand social media customer behavior and the pictures they post as teasers.

8.2. Content Cue-Based Methods. Mishra et al. [25] this technique will be created in accordance with the ideologies of news readers' choices and the manner in which journalists write the news for them as readers. There are many different methods to write these news articles, but they all use the same basic information. This technique presents 2 distinct analyses, as described in the following:

(i) Lexical and semantic level of analysis: as a result of the author's use of persuasive language, readers would take the false news as fact as in a narrative. The stylometric characteristics of text may be extracted utilizing automated techniques to tell between two journalistic genres

(ii) Syntactic and pragmatic level analysis: the pragmatic function is used in the discourse to identify the reference for future sections. By creating catchy titles, you may ensure that messages are full of meaningless rambling.

TABLE 1: Literature Review

Authors (year)	Aim	Issues	Methods	Datasets	Findings	Shortcomings/future work
Sansonetti et al. [38]	Detecting fake news and unreliable users	Figuring out whether or not a piece of social media news is trustworthy by looking for trustworthy sources	Deep learning techniques, CNN, LSTM, and neural network classifier	Twitter (user-profiles and shared news)	Recognizing reliable profile. Check for distributed material and social information's trustworthiness. The average accuracy of 90%.	Can be used of new datasets of news and users with further reliability prediction features.
Antoun et al. [39]	To address the risk for fake news spreading	Detecting fake news, identifying domains, and identifying bots in tweets	Bi-LSTM and voting classifier	QICC contest dataset	Superior performances and high impact features	Enhancing the detection of fake news by using elements from fact-checking sites in addition to Google searches.
Lin et al. [40]	To properly determine the position of news	Only one inference direction was utilized for categorizing the stance, which may have resulted in some crucial info being lost	BERT language model	Fake news challenge stage 1(FNC-1)	Detect the stance more accurately	The misprediction example of the proposed model is the stance relation of claim and article is "disagree" class.
Konkobo *et al.* [41]	The massive quantity of unlabeled data on social media must be dealt with	Unwilling to cope with fake news' massive quantity of unlabeled data	CredRank algorithm, CNN	Politifact and GossipCop	Politifact's precision is 71.10 percent, whereas Gossipcop's accuracy is 68.07 percent	There will be improvements to a greater number of prototypes that have been suggested. We will create a multilingual dataset and look at how different languages affect news categorization.
Huang and Chen [42]	So that fake news may be detected with greater precision	Cross-domain intractability issue, the required information is often unavailable or inadequate at the early stage	LSTM, depth LSTM, LIWC CNN, and N-gram CNN	FOR dataset, Buzzfeed corpus, SFL dataset, FND dataset satire, and political dataset	Increase the precision, optimize the weights, and look at the intractability problem cross domains	The grammatical analysis would be investigated in-depth, and the preprocessing will extract more useful info, resulting in improved precision in detecting fake news.
Yanagi et al. [43]	So that others may provide remarks and assist classify documents to assess	Early detection of fake news	Neural network model	FakeNewsNet dataset	Achieved the best recall score	Creating remarks may assist fact-checkers in determining if something is genuine or not.

Table 2: Continued.

Authors (year)	Aim	Issues	Methods	Datasets	Findings	Shortcomings/future work
Paixão et al. [44]	To the identification of news depending on several kinds of characteristics	Automatic detection of false news	Supervised and unsupervised learning	Fake.Br corpus	Acquired F1 scores up to 96%	Topic modeling.
Song et al. [45]	To recognize characteristics of fake news	Maintaining the modality-specific characteristics has an impact on the model's performances	Multichannel convolutional neural networks with residual cross-modal attention	Four real-world datasets	Learns more discriminable feature representations	—
Ren and Zhang [46]	To learn the node description in HIN	Intentional rumors may conceal a writing style that has not been well modeled and used to match the additional and distracting multimodal info	Hierarchical graph attention network	Two real-world fake news datasets	Expandability and generalizability	Other node classification related applications.
Ying et al. [47]	To make use of the relationship between segments at the same time	Falsification of multimedia data is a problem that is hardly handled.	Multimodal topic memory network (MTMN)	WEIBO and PHEME	Best performance	Discover innovative ways to utilize deep neural networks' prior knowledge for detecting fake news and other valuable complementary info.
Meel and Vishwakarma [48]	To create and falsify news stories that include both text and graphic elements to check error level analysis	As a result of its capability to have catastrophic effects by focusing on a particular kind of news, social networking became a significant issue.	Hierarchical attention network, bidirectional GRU, and ensemble learning	All Data[1], fake news Sample[2], and fake news detection[3]	The dataset's best precision was 95.90 percent while using fake news samples	Fake news identification utilizing multimodal data is still a challenging and unknown area that needs further study.
Khan et al. [49]	The effectiveness of several machine learning techniques on three distinct datasets will be evaluated	Maintaining the modality-specific characteristics has an impact on the model's performances	Pretrained language models, deep learning, and machine learning approaches	Liar[4], fake or real news dataset[5], and corpus	With limited datasets, pretrained algorithms like BERT and others do the greatest detecting fake news	On social media during the COVID-19 epidemic, could identify misleading and health-related fake news?

Table 1: Continued.

Authors (year)	Aim	Issues	Methods	Datasets	Findings	Shortcomings/future work
Shim *et al*. [50]	To avoid the distribution of fake news	Only propagation inside a single social media may be tracked when looking at the distributors' network.	link2vec	The fake news dataset in English and the dataset in Korean	A new source of background details for identifying false news that is successful in identifying short-form fake news	Further study is needed to determine how many top connections are ideal for maximizing the effectiveness of a fake news detecting system.
Samadi *et al*. [51]	To dive into comparison research regarding utilizing various classifiers and embedding frameworks	It is not clear which contextualized embedding will provide the classifier with the most useful characteristics.	CNN, single-layer perceptron (SLP), multilayer perceptron (MLP), BERT, RoBERTa, GPT2, and funnel transformer	LIAR, ISOT, and COVID-19	Achieve promising results without using additional features and network information increase the capacity of learning create dense contextualized embeddings for input texts	Has the ability to mine client profiles for additional information. Fake news may be detected using a variety of textual analyses in NLP.
Mhatre and Masurkar [52]	To determine the truthfulness of the news by scraping it from the internet	Increase the precision of categorization	Methods for NLP, like naive Bayes, KNN, decision tree, logistic regression, and the passive-aggressive classifier, or the SVM ensemble learning methodology	Web-scrapped data	Fake news identification accuracy has increased; thus, the results may now be classified as trustworthy or untrustworthy	—

Conclusion

Despite the high success rate in latent semantic analysis, the potential to detect false news and postings,Still, the ever-changing qualities and characteristics of false news on social media networks make it difficult to categorize. DL, on the other hand, is characterized by the ability to calculate hierarchical characteristics. With the deployment of DL research and application in the current past, many research tasks will apply DL technologies, including CNN, Deep Boltzmann machines, DNN, and Deep autoencoder models, including audio and voice processing, NLP, but modeling, Data retrieval, objective identification, and computer vision, as well as DNN implementation. In this comprehensive study, the basic concept of counterfeit news detection is described in detail along with their types, features and characteristics and also the classification for counterfeit news search model. Various fake search methods have been implemented to identify user behavior by spreading rumors or fake news. Three lies, fake news and numerous traditional machine learning and deep learning techniques have been compared on corpus datasets. This comparison shows that deep learning techniques have lagged behind traditional machine learning techniques. This study will be helpful for further research to identify fake news and develop new models or tools for early detection.

References

[1] K. S. Adewole, T. Han, W. Wu, H. Song, and A. K. Sangaiah, "Twitter spam account detection based on clustering and classification methods," *The Journal of Supercomputing*, vol. 76, no. 7, pp. 4802–4837, 2020.

[2] M. Z. Asghar, A. Ullah, S. Ahmad, and A. Khan, "Opinion spam detection framework using hybrid classification scheme," *Soft Computing*, vol. 24, no. 5, pp. 3475–3498, 2020.

[3] C. Boididou, S. Papadopoulos, M. Zampoglou, L. Apostolidis, O. Papadopoulou, and Y. Kompatsiaris, "Detection and visualization of misleading content on Twitter," *International Journal of Multimedia Information Retrieval*, vol. 7, no. 1, pp. 71–86, 2018.

K. Dhingra and S. K. Yadav, "Spam analysis of big reviews dataset using Fuzzy Ranking Evaluation Algorithm and Hadoop," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2143–2162, 2019.

[4] Y. Boshmaf, Logothetis, Siganos et al., *Íntegro: veveraging victim prediction for robust fake account detection in large scale OSNs*, vol. 61, Computers & Security, 2016.

[5] B. Jang, S. Jeong, and C. K. Kim, "Distance-based customer detection in fake follower markets," *Information Systems*, vol. 81, pp. 104–116, 2019.

[6] E. Kauffmann, J. Peral, D. Gil, A. Ferrández, R. Sellers, and H. Mora, "A framework for big data analytics in commercial social networks: a case study on sentiment analysis and fake review detection for marketing decision-making," *Industrial Marketing Management*, vol. 90, pp. 523–537, 2020.

[7] D. Plotkina, A. Munzel, and J. Pallud, "Illusions of truth—experimental insights into human and algorithmic detections of fake online reviews," *Journal of Business Research*, vol. 109, pp. 511–523, 2020.

[8] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211–236, 2017.

[9] C. S. Atodiresei, A. Tănăselea, and A. Iftene, "Identifying fake news and fake users on Twitter," *Procedia Computer Science*, vol. 126, pp. 451–461, 2018.

M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. Dipierro, and L. De Alfaro, "Automatic online fake news detection combining content and social signals," in *2018 22nd Conference of Open Innovations Association (FRUCT)*, Jyvaskyla, Finland, 2018.

[10] A. Pathak and R. K. Srihari, "BREAKING! Presenting fake news corpus for automated fact checking," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pp. 357–362, Florence, Italy, 2019.

[11] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social Media," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[12] S. Mishra, "Sentiment analysis of movie reviews," in *35th MP Young Scientist Congress*, p. 34, MPCST, 2020.

[13] P. B. Petkar, "Fake news detection: a survey of techniques," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 9, pp. 383–386, 2020.

[14] S. Mishra, "Analysis of user sentiments using machine learning algorithms," in *34th MP Young Scientist Congress*, p. 45, MPCST, 2019.

[15] D. Bright, R. Brewer, and C. Morselli, "Using social network analysis to study crime: navigating the challenges of crimi- nal justice records," *Social Networks*, vol. 66, pp. 50–64, 2021.

[16] T. T. Aurpa, R. Sadik, and M. S. Ahmed, "Abusive Bangla comments detection on Facebook using transformer-based deep learning models," *Social Network Analysis and Mining*, vol. 12, no. 1, p. 24, 2022.

[17] J. F. de Oliveira, H. T. Marques-Neto, and M. Karsai, "Measuring the effects of repeated and diversified influence mechanism for information adoption on twitter," *Social Network Analysis and Mining*, vol. 12, no. 1, p. 16, 2022.

[18] H. Alzahrani, S. Acharya, P. Duverger, and N. P. Nguyen, "Contextual polarity and influence mining in online social networks," *Computational Social Networks*, vol. 8, no. 1, p. 21, 2021.

[19] S. Mishra, P. K. Shukla, and R. Agarwal, "Location wise opinion mining of real time Twitter data using Hadoop to reduce cyber crimes," in *2nd International Conference on Data, Engineering and Applications (IDEA)*, pp. 1–6, Bhopal,India, 2020.

[20] G. Shrivastava, P. Kumar, R. P. Ojha, P. K. Srivastava, S. Mohan, and G. Srivastava, "Defensive modeling of fake news through online social networks," *IEEE Transactionson Computational Social Systems*, vol. 7, no. 5, pp. 1159– 1167, 2020.

[21] Y. Qin, D. Wurzer, and C. Tang, "Predicting future rumours,"

Chinese Journal of Electronics, vol. 27, no. 3, pp. 514–520, 2018.

[22] K. Zaamout, T. Arjannikov, and K. Barker, "Corrections to "a social crowdsourcing community case study: interaction patterns, evolution, and factors that affect them" [Jun 20 659-671]," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 5, pp. 1317–1317, 2020.

[23] S. Mishra, P. Shukla, and R. Agrawal, "A survey paper on classification of Sybil activities on online social network using machine learning algorithms," in *International Conference on Recent Advances in Smart Innovative ideas with Multidisciplinery Research (IC-RASIMR)*, pp. 133–136, Bhopal, Madhya Pradesh, India, 2019.

[24] C. Chen, F. Shen, J. Xu, and R. Yan, "Probabilistic latent semantic analysis-based gear fault diagnosis under variable working conditions," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 2845–2857, 2020.

[25] A. Farahat and F. Chen, "Improving probabilistic latent semantic analysis with principal component analysis," in *11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, 2006.

[26] C. Hong, W. Chen, W. Zheng, J. Shan, Y. Chen, and Y. Zhang, "Parallelization and characterization of probabilistic latent semantic analysis," in *2008 37th International Conference on Parallel Processing*, Portland, OR, USA, 2008.

[27] T. Hofmann, "Probabilistic latent semantic analysis," 2013, http://arxiv.org/abs/1301.6705.

[28] I. E. Faculty and M. Engineering, *Analysis of Different Machine Learning Master thesis*, International Journal of Computer and Information Technology, 2020, https://www.uib.no/en/rg/ml/128703/available-masters-thesis-topics-machine-learning.

[29] L. Singh, "Fake news detection using machine learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 9, no. 5, pp. 980–983, 2021.

[30] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *COLT '92: Proceedings of the fifth annual workshop on Computational learning theory*, Pittsburgh Pennsylvania USA, 1992.

[31] G. Revathya, R. M. Ariethb, D. Kalaiabiramic, and R. Arunad, "Revelation of diabetics by inadequate balanced SVM," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 2, pp. 2482–2486, 2021.

[32] M. R. Islam, S. Liu, X. Wang, and G. Xu, "Deep learning for misinformation detection on online social networks: a survey and new perspectives," *Social Network Analysis and Mining*, vol. 10, no. 1, p. 82, 2020.

[33] L. Alekya, L. Lakshmi, G. Susmitha, and S. Hemanth, "A survey on fake news detection in social media using deep neural networks," *International Journal of Scientific & Technology Research*, vol. 9, no. 3, 2020.

[34] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: ahybrid CNN-RNN based deep learning approach," *International Journal of Information Management Data Insights*, vol. 1, no. 1, article 100007, 2021.

[35] D. H. Lee, Y. R. Kim, H. J. Kim, S. M. Park, and Y. J. Yang, "Fake news detection using deep learning," *Journal of Information Processing Systems*, vol. 15, no. 5, pp. 1119–1130, 2019.

[36] G. Sansonetti, F. Gasparetti, G. D'aniello, and A. Micarelli, "Unreliable users detection in social media: deep learning techniques for automatic detection," *IEEE Access*, vol. 8, pp. 213154–213167, 2020.

[37] W. Antoun, F. Baly, R. Achour, A. Hussein, and H. Hajj, "State of the art models for fake news detection tasks," in *2020 IEEE international conference on informatics, IoT, and enabling technologies (ICIoT)*, pp. 519–524, Doha, Qatar, 2020.

[38] S.-X. Lin, B.-Y. Wu, T.-H. Chou, Y.-J. Lin, and H.-Y. Kao, "Bidirectional perspective with topic information for stance detection," in *2020 International Conference on Pervasive Artificial Intelligence (ICPAI)*, pp. 1–8, Taipei, Taiwan, 2020.

[39] P. M. Konkobo, R. Zhang, S. Huang, T. T. Minoungou, J. A. Ouedraogo, and L. Li, "A deep learning model for early detection of fake news on social media∗," in *2020 7th International Conference on Behavioural and Social Computing (BESC)*, pp. 1–6, Bournemouth, United Kingdom, 2020.

[40] Y.-F. Huang and P.-H. Chen, "Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms," *Expert Systems with Applications*, vol. 159, article 113584, 2020.

[41] Y. Yanagi, R. Orihara, Y. Sei, Y. Tahara, and A. Ohsuga, "Fake news detection with generated comments for news articles," in *2020 IEEE 24th International Conference on Intelligent Engineering Systems (INES)*, Reykjavík, Iceland, 2020.

[42] M. Paixão, R. Lima, and B. Espinasse, "Fake news classification and topic modeling in Brazilian Portuguese," in *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 427–432, Melbourne, Australia, 2020.

[43] C. Song, N. Ning, Y. Zhang, and B. Wu, "A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks," *Information Processing & Management*, vol. 58, no. 1, article 102437, 2021.

[44] Y. Ren and J. Zhang, "Fake news detection on news- oriented heterogeneous information networks through hier- archical graph attention," in *2021 International Joint Con- ference on Neural Networks (IJCNN)*, pp. 1–8, Shenzhen, China, 2021.

[45] L. Ying, H. Yu, J. Wang, Y. Ji, and S. Qian, "Fake news detection via multi-modal topic memory network," *IEEE Access*, vol. 9, pp. 132818–132829, 2021.

[46] P. Meel and D. K. Vishwakarma, "HAN, image captioning, and forensics ensemble multimodal fake news detection," *Information Sciences*, vol. 567, pp. 23–41, 2021.

[47] J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," *Machine Learning with Applications*, vol. 4, article 100032, 2021.

[48] J.-S. Shim, Y. Lee, and H. Ahn, "A link2vec-based fake news detection model using web search results," *Expert Systems with Applications*, vol. 184, article 115491, 2021.

[49] J.-S. Shim, Y. Lee, and H. Ahn, "A link2vec-based fake news detection model using web search results," *Expert Systems with Applications*, vol. 184, article 115491, 2021.

[50] M. Samadi, M. Mousavian, and S. Momtazi, "Deep contextualized text representation and learning for fake news detection," *Information Processing & Management*, vol. 58, no. 6, article 102723, 2021

[51] S. Mhatre and A. Masurkar, "A hybrid method for fake news detection using cosine similarity scores," in *2021 International*

Conference on Communication information and Computing Technology (ICCICT), pp. 1–6, Mumbai, India, 2021