ISSN: 2581-3404 (Online) IF: 5.445 (SJIF)
International Journal of Innovative Research in Technology and Management, Volume-6, Issue-2, 2022.

A Review on Fake News Detection using Machine Learning Algorithm

Shreya priyadarshi[1], Chetan Agrawal[2], Priyanka Parihar[3]
Dept. of CSE, Radharaman Institute of Technology & Science, Bhopal, India[1, 2, 3]
priyadarshi.shreya49@gmail.com[1], chetan.agrawal12@gmail.com[2], priyanka.parihar28113@gmail.com[3]

Abstract: *Now days the usage of Internet and online marketing has become very popular. Millions of products and services are available in online marketing that generate huge amount of information. Social media sites can have a major influence in expanding the span of this kind of story. Fake news is news created to intentionally misguide or mislead readers. Fake news is spread mainly for gaining political or financial incentives. There has been a large surge of fake news in recent times due to the immense use of social media and online news media. It has become much easier to spread fake news then how it used to be earlier. This kind of fake news when spread may have a severe effect. Hence it is extremely essential that certain measures should be taken in order to reduce or distinguish between real and fake news. This paper presents a survey on fake news detection based on various supervised, unsupervised and semi supervised datamining and machine learning techniques.*

Keywords: Fake News, Deep Learning**,** Machine Learning, Social Media, Rumor, Misinformation.

Introduction

News is basic part of our life. In regular daily existence current news are valuable to improve data what happen all over. So most of society favor watching news by far most of the society all around slant toward examining paper quickly in the initial segment of the day getting an accuse out of cup of tea. If news is phony that will trick social orders occasionally counterfeit word used to get out gossipy goodies about things or it will impact some political pioneer positions taking into account counterfeit news. By far most of the serious cell phone customers need to scrutinize the news by methods for online media over web. The news destinations are circulating the news and give the wellspring of affirmation. The request is the best approach to check the news and articles which are flowed among electronic media like WhatsApp social occasions, Facebook Pages, Twitter and other scaled down scale online diaries and individual to individual correspondence objections. It is hazardous for the overall population to acknowledge on the pieces of chatter and affirm to be news. So it's crucial to find the phony news. Online news organizes phenomenally sway our overall population and culture in both positive andnegative ways.

As online media ends up being more dependent for wellspring of information, a lot of phony news is posted on the web, that wide with people following it with no prior or complete information of event validity. Such misdirection can control famous sentiments. The exponential improvements of phony news inciting have become a remarkable peril to open for news steadfastness. It has become a persuading issue for which finding, inspecting and overseeing counterfeit news has extended famous. Spreading of deception on the web these days addresses a troublesome issue, as their impact on people'sappraisals may be immense.

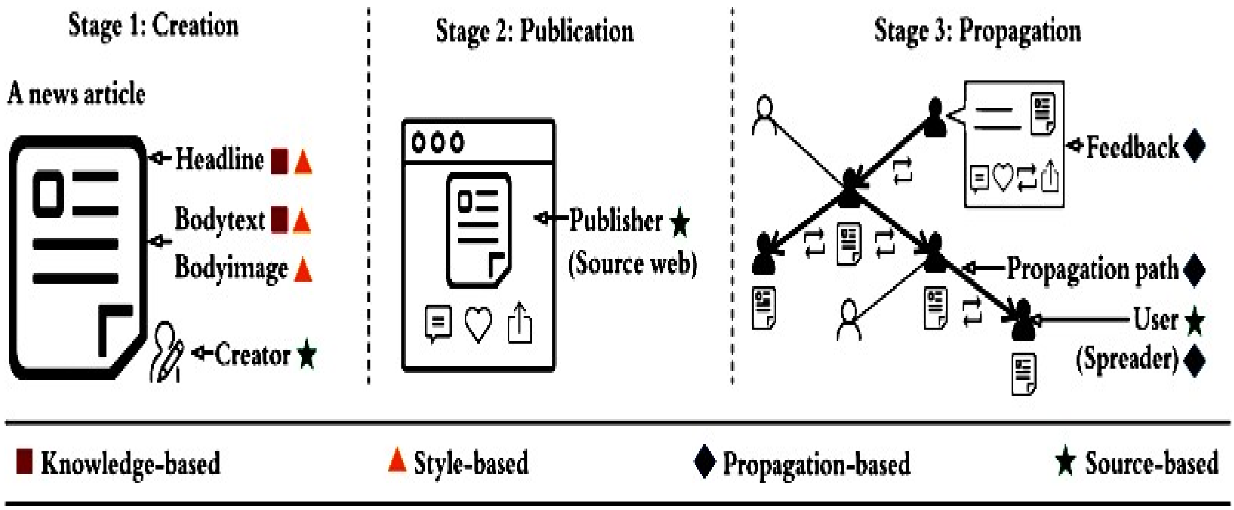

Figure 1: Fake News Life Cycle.

Fake news addresses a specific kind of double dealing. While its location was by and large being performed genuinely previously, automated procedures using AI and related fields ended up being more fundamental. On the other hand, significant learning methods ended up being standard and a great part of the time used techniques in the field of information examination starting late.

This audit paper presents a survey of the phony news recognition and discussion of the promising assessment heading. The key motivations of this examination are summarized as follows:

- Fake news through online media has been going on for a serious drawn-out period of time; nevertheless, there is no unending gracefully of the term counterfeit news". To all the almost certain guide the future headings of phony news recognition research, appropriate clarification are fundamental.
- Social media has wind up being a mind boggling hotspot for counterfeit news spread. There are some rising models that can be utilized for counterfeit news identification in online media. A review on existing phony news recognition methods under various online media circumstances can give a basic understanding on the top tier counterfeit news discovery procedures.

Fake news discovery by means of online media is still in the early season of headway, and there are up 'til now many testing issues that need further assessments. It is critical to discuss potential investigation heading that can improve counterfeit news discovery and easing limits.

II. Literature Review

Fake news detection has recently received a lot of research attention. Early attempts on fake news detection were mainly focused on information that was extracted from text to capture the different linguistic patterns used in fake and real news. One of the early works was presented by Castillo et al. [1] who explored the effectiveness of various statistical text features, such as count of word and punctuation on information credibility. More recently, Rashkin et al. [2] incorporated various linguistic features extracted with the Linguistic Inquiry and Word Count

(LIWC) dictionary [3] such as personal pronouns and swear words into an LSTM network in order to differentiate between credible and not credible claims, whereas Wang [4] proposed a hybrid convolutional neural network to combine user metadata with text for fake news detection. Based on the intuition that fake news triggers different emotions compared to real news to the users, some researchers proposed extracting the emotions expressed in the text and they explored their effectiveness on the task of fake news detection. Vosoughi et al. [5] investigated true and false rumors on Twitter and found that false rumors triggered fear, disgust and surprise in their replies, whereas the true rumors triggered joy, sadness, trust and anticipation. Giachanou et al. [6] pro- posed an LSTM-based neural network that leveraged emotions expressed in the text. They explored three different ways to extract the emotions, two of them were lexicon-based and one was based on a neural network. In their study, Giachanou et al. showed the effectiveness of the emotions expressed in the text on credibility detection. Another work that explored the impact of emotions on fake news detection was presented by Ghanem et al. [7]. Ghanem et al. who proposed to extract the emotions expressed in the text and incorporated them into an LSTM network showed that emotions are useful for the classification of the different types of fake news. Users can also play an important role in the propagation of fake news since they are the ones that decide to share the fake information intentionally or unintentionally. To this end, some researchers explored the role of users in the detection and propagation of fake news. Shu and Wang [8] performed an analysis of user profiles that share fake or real news. The analysis showed that there are features (e.g., registration time) that are different between users that share fake news and those that share real news. In addition, they examined the effectiveness of those features on fake news detection and showed that combining user profile features with the psycho- linguistic characteristics of the document can be very effective for fake news detection. Vo and Lee [9] analyzed linguistic characteristics of fact-checking tweets (i.e., tweets that confirm that an article is fake) and also proposed a deep learning framework to generate responses with fact-checking intention. Their analysis showed that the fact-checkers tend to refute fake news and use formal language. Multimodal fake news detection has also received research attention since the majority of the articles contain one or more images. Fake news usually contains images that are manipulated in sophisticated ways to deceive the users, attract their attention and convince them to share them. Visual information extracted from the images, such as image tags and color histogram can complement the textual one. Different studies have focused on that and showed that visual features can be an important indicator for fake news detection. Jin et al. [13] proposed several visual and statistical features to characterize different patterns used in fake and real news in order to detect fake news. However, their work was based on hand-crafted features that cannot capture complex distributions of visual content. More recently, the multimodal approaches that were pro- posed exploited the advances of deep learning area. Wang et al. [10] proposed the Event Adversarial Neural Networks (EANN) model that consists of two components: the textual and the visual. The textual component was represented by word embeddings generated using Convolutional Neural Network (CNN), whereas the visual one was represented by features that were extracted using the VGG-19 model pre- trained on ImageNet [11]. The two representations were then concatenated and fed to two fully connected neural networks, one network was used for event discriminator and the second for fake news classification. Khattar et al. [12] proposed an end-to-end network, Multimodal Variational Autoencoder (MVAE) model based on bi-directional LSTMs and VGG-19 for the text and image representation respectively. The model consists of three main components, an encoder, a decoder and a fake news detector module. The variational autoencoder is capable of learning probabilistic latent variable models whereas the fake news detector utilizes the multimodal representations obtained from the variational autoencoder to classify posts as fake or not. Singhal et al. [13] focused also on multimodal fake news detection and proposed the SpotFake system. SpotFake is based on the textual and visual features of an article. For the textual representation, Singhal et al. used BERT to incorporate contextual information, whereas for the image features, they used the VGG-19 pre-trained on ImageNet dataset. The rep- resentations from both the modalities are then concatenated together to produce the desired news vector. Their results showed the importance of combining contextual information and visual features for fake news detection. Regarding the image-text similarity, there are few works that have explored its effectiveness on fake news detection. Zlatkova et al. [14] explored the effectiveness of text-image similarity in addition to other visual information. However, Zlatkova et al. focused on claim factuality prediction with respect to an image that is a different problem to the one of fake news detection. Zhou et al. [15] proposed the Similarity- Aware FakE news detection method (SAFE) that consisted of three components, the multimodal one, the within modal and the cross-modal similarity extraction. Zhou et al. used neural networks to automatically obtain the latent representation of the textual and visual information based on which a similarity measure was defined between them. Unlike previous works, we propose visual features that are extracted from multiple images and which we pass then into an LSTM layer to model the sequence information. In addition, we explore the effectiveness of the similarity between text and image that can better capture different patterns used in fake and real news. For the textual component we use BERT that can learn the context of a word based on all of its surroundings, whereas the similarity is calculated using the embeddings of the post's text and the image tags.

III. Evaluation Metrics

To evaluate the performance of algorithms for fake news detection problem, various evaluation metrics have been used. In this subsection, we review the most widely used metrics for fake news detection. Most existing approaches consider the fake news problem as a clarification problem that predicts whether a news article is fake or not:

- True Positive (TP): when predicted fake news pieces are actually annotated as fake news
- True Negative (TN): when predicted true news pieces are actually annotated as true news
- False Negative (FN): when predicted true news pieces are actually annotated as fake news
- False Positive (FP): when predicted fake news pieces are actually annotated as

  true news

By formulating this as a clarification problem, we can define following metrics,

$$Precision = \frac{|TP|}{|TP| + |FP|}$$

$$Recall = \frac{|TP|}{|TP| + |FN|}$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

$$Accuracy = \frac{|TP| + |TN|}{|TP| + |TN| + |FP| + |FN|}$$

1. DETECTION TECHNIQUES

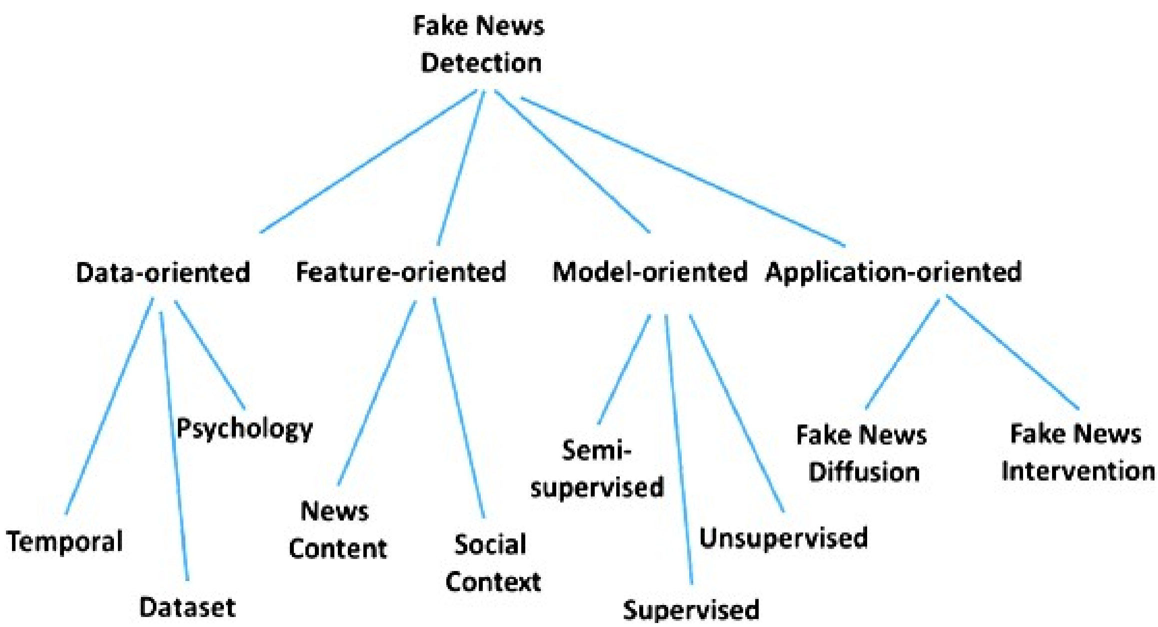

Figure 2: Future directions and open issues for fake news detection on social media.

Various techniques have been proposed in past to identify fake reviews based on types of data like labeled data (for example, supervised learning), unlabeled data (for example, unsupervised learning), and partially labeled data (for example, semi-supervised learning) that is described below.

A. Supervised Learning techniques

Before applying the classification method, different preprocessing steps are performed; these steps include stemming, removal of punctuation marks and stop word removal. They use linguistic feature to identify fake reviews. Linguistic feature contains POS and bag-of-words. Bag-of-words features consist of individual word or group of words that are found in given text. Then different classification algorithms are applied like decision tree, random forest, support vector machine, Naivebayes and gradient boosted trees. Here Naivebayes and support vector machine give better result.

B. Unsupervised Learning techniques

Main advantage of unsupervised learning approach is that, without any labeled dataset, we can classify fake and genuine reviews. This concept uses different features based on review data, reviewer data and product information based on difference in behavioral pattern of reviews. Here author uses Amazon cell phone reviews dataset to identity fake and genuine reviews.

C. Semi-Supervised Learning techniques

Positive Unlabeled (PU) learning technique is combination of some positive label and unlabelled dataset. PU-learning technique is semi supervised technique, which only uses two class classifiers positive as deceptive and unlabeled without having negative as truthful training example. In this algorithm, firs unlabeled data are considered as negative class. In next step, classifiers are trained based on initial set of positive instances. Then classifiers are applied only on unlabeled instances and generate labeled instances. After, classified positive and negative instances, the positive instances as deceptive reviews are eliminated from unlabeled instances and rest of them are considered as negative instances. Again classifiers are applied into negative instances. This process is repeated until the stop criteria, which classify fake and genuine reviews. Here two classifiers are applied in PU learning, support vector machine and Naivebayes.

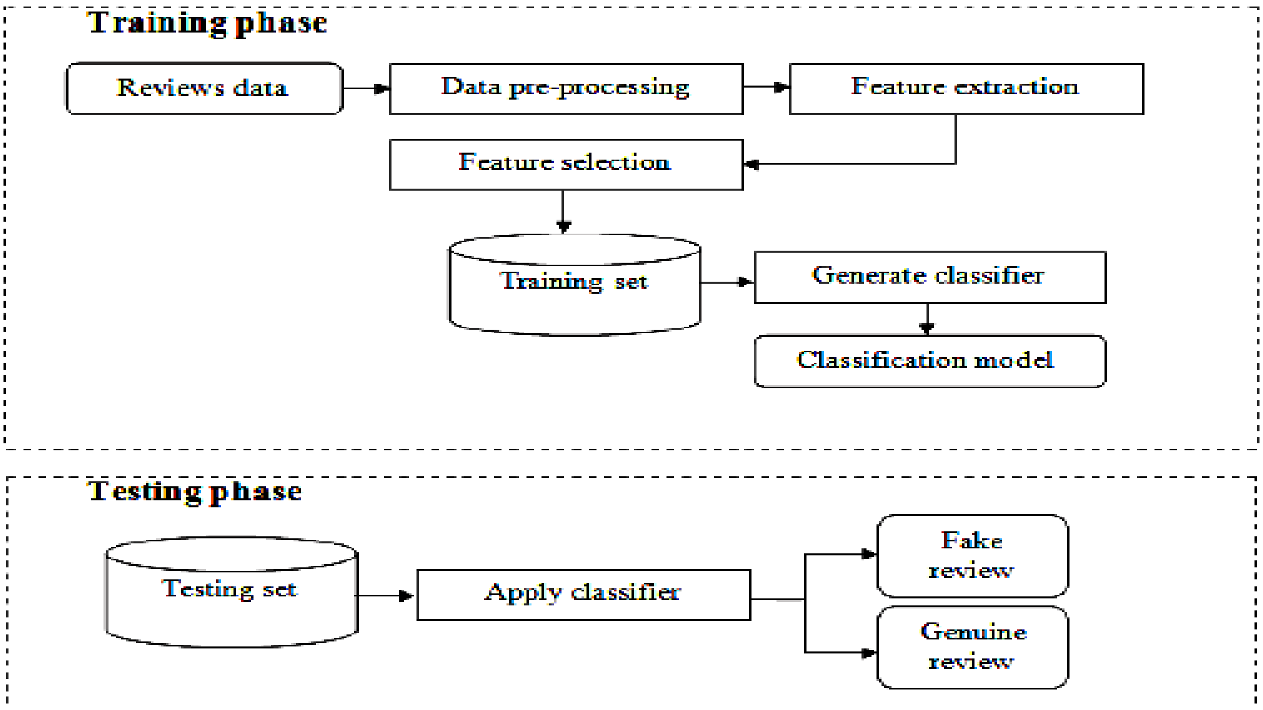

Figure 3: Machine Learning based Fake Review Detection.

There are different approaches to distinguish counterfeit surveys. AI strategy is one of the approaches to distinguish counterfeit audits [10]. AI model learns and make forecast [2]. The fundamental advances associated with AI are information handling, highlight extraction, include determination, and order model age. This cycle is appeared in Fig. 3: AI approach for counterfeit survey location functions as follows:

- **Data collection:** In this stage, survey information will be assembled from different stages like Amazon. Theseaudits could be for item or administration like inn surveys.
- **Data pre-processing:** In subsequent stage, information preprocessing is applied like accentuation marks expulsion, stemming, stop word evacuation and so forth. In accentuation marks evacuation, the entire content is isolated into sentences, expressions or sections. In the stemming cycle, stem will be made from each word in dataset. In stop word evacuation stage, as often as possible utilized gathering of words like determiners, articles and relational word will be distinguished and taken out. Subsequent to eliminating these words, just significant words will be held for the followingstage.

Classifier model construction and testing: For preparing reason, little arrangement of named information is utilized. In this stage, grouping model is created by utilizing the preparation audit dataset. The surveys utilized for this design are as of now marked as phony or veritable audit. When the classifier is prepared, it will be tried utilizing test dataset. The distinctive AI calculations which can be utilized for model development are guileless bayes order, choice tree calculation, uphold vector machine, k-closest neighbor, strategic relapse, and so forth. The exhibition of phony survey identification strategy relies upon named information utilized for preparing reason, right determination of highlights and information digging procedures utilized for recognition.

IV. Conclusion

Machine Learning uses a statistical technique to give the computer the ability to learn with data hence it is widely used in the detection of fake news. Methods used for taking parameters and for categorizing the type of news are also discussed. With the increasing popularity of social media, more and more people consume news from social media instead of traditional news media. However, social media has also been used to spread fake news, which has strong negative impacts on individual users and broader society. From the literature review it has been observed that the accuracy for predicting fake news in social media is much higher than any other online news media hence we have targeted online news media fake news detection along with website verification. In this article, we review the fake news detection based on the data mining and the machine learning approach. In future work, our proposed model will be tested for fake news detection by using standard dataset and apply machine learning based algorithm for detection and validation. Our approach will improve the performance parameters and make an efficient algorithm.

References:

[1]. C. Castillo, M. Mendoza, and B. Poblete. Information Credibility on Twitter. In Proceedings of the 20th International Conference on World Wide Web, WWW ’11, pages 675–684, 2011.

[2]. H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi. Truth of varying shades: Analyzing Language in Fake News and Political Fact- Checking. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP ’17, pages 2931–2937, 2017

[3]. R. T. Yla and W. P. James. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. Journal of Language and Social Psychology, 29(1):24–54, 2010.

[4]. W. Y. Wang. Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), ACL ’17, pages 422–426, 2017.

[5]. S. Vosoughi, D. Roy, and S. Aral. The Spread of True and False News Online. Science, 359(6380):1146–1151, 2018.

[6]. A. Giachanou, P. Rosso, and F. Crestani. Leveraging Emotional Signals for Credibility Detection. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR ’19, pages 877–880, 2019.

[7]. B. Ghanem, P. Rosso, and F. Rangel. An Emotional Analysis of False Information in Social Media and News Articles. ACM Transactions on Internet Technology (TOIT)., 20(2):1–18, 2020.

[8]. K. Shu, S. Wang, and H. Liu. Understanding User Profiles on Social Media for Fake News Detection. In Proceedings of the 2018 IEEE Conference on Multimedia Information Processing and Retrieval, MIPR ’18, pages 430–435, 2018.

[9]. N. Vo and K. Lee. Learning from Fact-checkers: Analysis and Generation of Fact-checking Language. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR ’19, pages 335–344, 2019.

[10]. Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao. EANN: Event Adversarial Neural

Networks for Multi-modal Fake News Detection. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD'18, pages 849–857, 2018.
[11]. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A Large-scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '09, pages 248–255, 2009
[12]. D. Khattar, J. S. Goud, M. Gupta, and V. Varma. MVAE: Multimodal Variational Autoencoder for Fake News Detection. In Proceedings of the 2019 World Wide Web Conference, WWW '19, pages 2915–2921, 2019.
[13]. [S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh. SpotFake: A Multi-modal Framework for Fake News Detection. In Proceedings of the IEEE 5th International Conference on Multimedia Big Data, BigMM, pages 39–47, 2019.
[14]. D. Zlatkova, P. Nakov, and I Koychev. Fact-checking meets Fauxtog- raphy: Verifying Claims about Images. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP '19, pages 2099–2108, 2019.
[15]. X. Zhou, J. Wu, and R. Zafarani. SAFE: Similarity-aware Multi-modal Fake News Detection. arXiv preprint arXiv:2003.04981, 2020.