

Fake News Filtering Hybrid Algorithm - A Systematic Review

Ekene Frank Ozioko¹, Johnpaul Nwaigwe¹, Izuchukwu Akwudi²

¹Department of Computer Science, Enugu State University of Science and Technology, ESUT, Enugu, Nigeria

²Information and Communication Technology Centre, Enugu State University of Science and Technology, ESUT, Enugu, Nigeria.

Date of Submission: 18-05-2024

Date of Acceptance: 28-05-2024

ABSTRACT—Spreading of “rumors and mistrust information” has been identified as one of the significant concerns of abuse that has plagued social media platforms. The attractiveness of fake news across social media users is one of the key drivers for its fast dissemination. Online hoax news spreads rapidly because most social media platforms have no technique for analysing and detecting fake news. There exist challenges in developing a robust algorithm for detecting and evaluating what fake news contains before online publication and broadcasting. This paper looked at specific assessment parameters for news before adjoining it as fake or real.

a) Keywords::Fake news, filtering, detection, news categorisation, news dissemination platform

I. INTRODUCTION

Fake news can be defined as false information that is disseminated or published as a piece of news for fraudulent or politically motivated purposes. It can also be defined as news or stories with intentionally false information [1]. Fake news is different from news where the source is unsure or has not performed a thorough search on the subject, which is called misinformation because fake news is purposely released to deceive people [2]. Fake news grow quickly and should be reserved for cases of deliberate presentation of false or misleading claims as news, where these are misleading by design. This means that fake news is purposely manufactured by malicious persons to deceive people or misled the community [3]. Some authors have researched into this topic to find out how some of these evils are perpetuated. The author in [4] tried to find how some online users use fake profile in social media and proposed a model using machine learning and Natural Language Processing (NLP) techniques that will help in detecting these wrong profile. This research work was carried out using Oroma language which is a language spoken in Ethiopia, Kenya, Somalia, and

Egypt. [5] analysed, 14 million messages spreading 400 thousand claims on Twitter during and following the 2016 U.S. presidential campaign and election. They used Hoaxy which was an open search engine to collect fake news, the result showed that social bots were the main source of biased and fake claims. Network analysis can also help in identifying and mitigating fake news. In trying to use networks in tracking the challenges imposed by fake news, the writers in reviewing the network introduced popular network types and proposed how they can be used in identifying and mitigating fake news.

II. LITERATURE REVIEW

The researchers in [6] examined if filter bubbles and fake news have effect in the outcome of U.S presidential election in 2016. Through the review of literature, the author was able to ascertain that filter bubble and fake news affect the result of that election and that Facebook and Google are trying to explore ways of reducing it. [7] applied quantitative and qualitative content analysis of online comment to find out if fake news is supported by filter bubbles and also whether echo chambers are man made. Their result showed that they found hint on filter bubbles. The researchers equally discovered it was not possible to describe all information patterns following fake news with the theory of selective exposure but rather with individual cognitive structures like non-argumentative or off-topic behaviour, denial and so on. The authors in [8] used Naive Bayes Classifier to detect fake news. The research used a questionnaire and interview as methodologies and gave a classification accuracy of (74%). In [9] the writers used supervised Artificial Intelligence Algorithms to detect fake news within online social media. They proposed twostep techniques for identifying fake news on social media. Study [10], discussed the various techniques to fight fake news. These techniques include determining text features using linguistic natural language processing

methods, detect spam bots in social networks to identify those using machine learning methods and confirm the facts in online documents by approaches used in search engines. These techniques can show a high level of accuracy in filtering fake news. In [11], the authors proposed an algorithm that was able to detect and filter what constitute online fake news. From the pieces of literature reviewed in this chapter, fake news has been defined as “news articles that are intentionally and verifiably false, and could mislead readers”. Fake news can be characterized as news that comes from an unfamiliar website, the headline is outrageous or doesn’t match the article, or news that fails to provide proof of claim (Ireton, 2018). In the review, some techniques and tools was considered for fake news detection. Arvin, Timothy, and Jeanette (2021) used Blending Ensemble Learning for detection of online fake news. Monti et al. (2019) used geometric deep learning (GDL) to identify and extract fake news on social networks and cyberspace. Using natural language processing (NLP) and deep convolution neural network (CNN), the researchers identified general and counterfeit deception news. The accuracy of the algorithm was 92.7% presents a summary of the related literature and the technique applied with the results achieved.

A. Classification Matrix of The Existing Methodologies on Fake News Detection

The detailed classification matrix for the existing fake news filtering methodologies is represented in table I

Although there exist research works that use only exploitative research models, linguistic features (NLP), and text representation analysis which are individually not enough for fake news detection. However, none has combined the exploitative research model with the linguistic features analysis (NLP) for automatic detection of fake news on social media. The addition of the Network confirmation approach should be considered in this existing model for a better result. This should verify the IP address of the source of the

news, and crosscheck that this IP address has not sent fake news before. Therefore, a data set combining the network source data-set and linguistic feature data-set will be used in this research. Most of the pieces of literature reviewed used Sentiment Analysis for fake news detection. The system only detects fake news but doesn’t block the news from public view. So this research introduces the use of a decision tree algorithm in fake news filtering and it can also block fake news articles from public view

III. PROPOSED METHOD

The proposed method is a hybrid model comprising Linguistic Analysis to extract meaning from the news posts on social media and the exploratory research model used in [12] to study how cognitive, visual, effective, and behavioural cues of the news post and the associated comment can be used for automatic filtering of fake news. Then the model will be trained with a supervised machine learning approach. The labeled data set comprises of features such as comments, views, the IP address of the source, author of the post, date of the post, URLs, and spam word.

A. Fake News Detection Methods

The wide usage of social media platforms worldwide has provided a fertile ground for the widespread dissemination of online fake news in an unprecedented way. The social network is flooded with massive, diverse, and heterogeneous information (both real and fake), and spreads rapidly on these platforms causing severe impact to the whole society. Therefore, many researchers and technical giants are working together to detect fake news in online media. The traditional automatic rumor detection methods were based on handcrafted features but with the advent of big data and a huge base of user-generated data we have seen a shift to deep-level features. In this section, we discuss various state-of-the-art

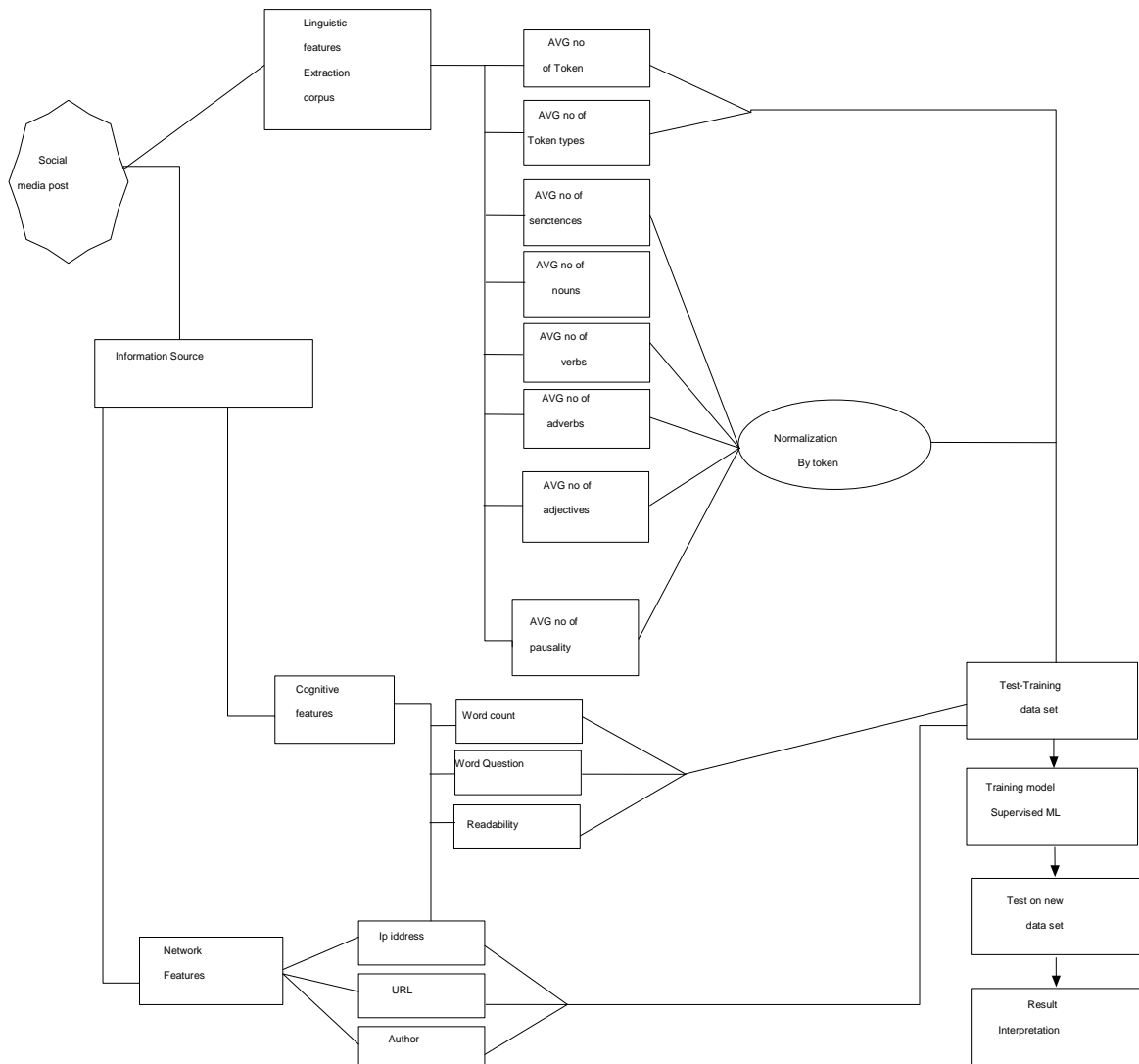


Fig. 1: System Block Diagram

studies on fake news detection (figure 2.1) under the broader umbrella of content and social context of the news article.



Fig. 2.1: Fake News Detection Methods (Zhou, 2019)

B. Content Based

The content-based fake news detection method aims to detect fake news by analyzing the content of the article (Zhou, 2019), i.e., either the text or image or both within the news article. For automatically detecting the fake news, the researchers

often relying on either latent or hand-crafted features of the content.

C. Knowledge Based

Knowledge-based approaches utilize fact checking method in which the given claim is compared with the external sources to verify the

authenticity of the given claim. The existing factchecking methods can be categorized as manual (using experts or by crowd-sourcing) and automatic fact-checking. The manual fact-checking can be broadly divided into expertbased and crowd-sourced fact-checking. The expert-based methods use an expert-oriented approach and rely on human experts working in specific domains for decision-making. Factchecking websites like Snopes, PolitiFact, GossipCop use this approach. These methods are reliable but are time-consuming and do not scale well with the huge volume of content available on social media. Many researchers use these websites for creating their own datasets among these are the benchmark datasets LIAR (Wang, 2017) and FakeNewsNet (Shu, 2018). For crowdsourced approaches, “wisdom of crowd” helps to check the accuracy of the news articles. A similar approach is used by Fiskkit that provides a platform for people to discuss important news articles and finds out their accuracy. Crowdsourced fact-checking is even though relatively difficult to manage, biased, has conflicting annotations, is less credible but has better scalability as compared to expert based factchecking (Zhou, 2020). CREDBANK (Mitra, 2015) is a publicly available large-scale benchmark fake news dataset that is annotated by fact-checkers. The datasets that are created using this approach needs to be filtered for non-credible users and the conflicting annotations need to be resolved beforehand. Some similar datasets can also be created and then be annotated using crowd-sourcing marketplaces such as AMT (Amazon Mechanical Turk).

The Manual fact-checking approaches do not scale well with the huge volume of data especially generated with the use of social media as a result, to address this issue automatic factchecking techniques have been deployed. Instead of relying on human intelligence these methods heavily rely on Natural Language Processing (NLP), Data Mining, Machine Learning (ML) techniques and network/graph theory (Zhou, 2020). The automatic fact-checking process can be divided into two stages:

- Fact extraction which is related to collection of facts and construction of a Knowledge Base and
- Fact-checking which is used to assess the authenticity of news articles by comparing that with the facts in the knowledge base.

It uses open web source and knowledge base/graph to check whether the given claim is true/false. The real-world datasets for fake news detection are usually incomplete, unstructured, unlabeled and noisy which make automatic detection a bit complex (Shu, et. al. 2017).

D. Style Based

Style-based fake news detection follows the same approach like knowledge-based fake news detection by analyzing the news content. However, instead of evaluating the authenticity of news content this method assess the intention of writer to mislead the public (Shu, 2018). Fake news publishers usually have an intent to influence large communities while spreading distorted and misleading information. To make the titles catchy fake titles use mostly all capitalized words, significantly more proper nouns, and fewer stop words (Shu, 2018). Style-based approaches capture the distinguishing characteristics of writing styles between legitimate users and anomalous accounts to detect fake news.

E. Linguistic Based

Twenty-six linguistic-based textual features were proposed in (Hakak, 2021). In Gravanis, (2019) authors proposed an enhanced set of linguistic features to discriminate between fake and real news. Despite the success of this method in various scenarios, it poses a limitation in case of detecting misinformation on popular social media platforms where the messages are short, and thus, the linguistic features extracted from them are often inadequate for machine learning algorithms to make accurate predictions. Additionally, these approaches cannot be used to detect fake news that contains no text content but only a photo or a video.

F. Visual Based

Visual content is often viewed as evidence that can increase the credibility of the news article and hence fake news publishers tend to utilize provocative visual content to attract and mislead readers. In Jin, (2017) various visual and statistical image features are extracted for new authentication. Verifying Multimedia Use task under the MediaEval-16 benchmark addresses the problem of detecting digitally manipulated (tampered) images (Boididou, 2016).

G. Social Context Based

There are three major aspects of the social context i.e.: user profiles, user posts and responses, and network structures (Shu, et. al. 2018). It represents how the news proliferates over time and provides useful information to determine the veracity and stance of news articles. Recent studies have explored various context-based approaches for fake news detection.

H. Network Based

Network-based fake news detection studies different social networks like friendship, tweet-re-tweet, and post-repost networks to detect fake news. It detects who spreads the fake news, relationships among the spreaders, and how fake news propagates on social networks. Users tend to create various networks on online platforms media in terms of their common interests and similarities, these networks serve as paths for information diffusion. Wu, et. al. (2015) model the pattern of message propagation as a tree, which along with the relation among posts gives additional information about the temporal behavior and the sentiment of the posts.

I. Temporal Based

Studies have shown that news stories on the Internet are not static but are constantly evolving over time by adding new information or twisting the actual claim. This is very much evident in cases where the rumors resurge multiple times after the original news article is posted. The lifecycle analysis of rumor helps in understanding this phenomenon and Shin, et. al. (2018) examine the recurring rumors at the message level across different time periods. Kwon, et. al. (2017) provide a deep understanding of the diffusion patterns of rumors over time.

J. Credibility Based

The credibility of the claim, publisher, and spreader is often assessed by its news quality and trustworthiness/ credibility. Rath, et. al. (2018) identify the users spreading rumor by leveraging the concept of believability. Alrubaian, et. al. (2018) proposed a credibility analysis system for evaluating credibility of a Tweet and prevents the proliferation of fake or malicious information.

IV. RESULTS

The main goal of the proposed system is to apply a set of filtering algorithms to obtain a classification model in order to be used as a scanner for fake news by details of news detection and embed the model in python application to be used as a discovery for the fake news data. For the fake news filtering, the actual news data (body of the news article) is being considered as features. Fake news filtering in social media aims to extract useful features and build effective models from existing social media datasets for detecting fake news in the future. Thus, 22a comprehensive and large-scale dataset with multidimensional information in the online fake news ecosystem is important. The multidimensional information not only provides more signals for detecting fake news but can also be used for re- search such as understanding fake news propagation and fake

news intervention. In this project, the system developed uses a decision tree algorithm to classify the news. The project proposes a methodology to create a model that will detect if an article is authentic or fake based on its words, phrases, sources, and titles, by applying a decision tree algorithm on an annotated (labeled) dataset that are manually classified and guaranteed. Decision Tree was used for data classification. A decision tree is an important tool that works based on flowchart-like structure that is mainly used for classification problems. Each internal node of the decision tree specifies a condition or a "test" on an attribute and the branching is done on the basis of the test conditions and result. The advantage of the decision tree algorithm is that it can work with category and dependent variable. They are good at identifying the most important variables and they also depict the relation between the variables quite suitably. They are significant in creating new variables and features which is useful for data exploration and predicts the target variable quite efficiently. Tree-based learning algorithms are widely used with predictive models using supervised learning methods to establish high accuracy. They are good at mapping nonlinear relationships. They also solve classification or regression problems quite well.

This study tends to develop an algorithm that filters fake news. The methodologies of the already existing algorithms were reviewed and the pros and cons of each of the methodologies were noted. Figure 3.2 shows the user requirements functions that are performed by the users on the system. The users of the proposed system are categorized into users and admin. The activities of these users are described in Figure 3.2 using use-case diagrams.

V. SUMMARY, CONCLUSION AND DIRECTION FOR FUTURE WORK

6.1 Summary

The fake news filtering system was developed utilizing a suitable online dataset. In the software developed, users can sign up to the platform, read news, upload news, verify news, and also be able to comment on news items. The software was developed using Django, a Python web development platform. The user can check the news article or keywords online; also check the authenticity of the news.

So having achieved this, the system developed will assist in a

big way to checkmate the spread of fake news thereby giving more credibility to online news platforms.

6.2 Conclusion

The growing problem of fake news only makes things more complicated and tries to change or

hamper the opinion and attitude of people towards use of digital technology. When a person is deceived by the real news two possible things happen- People start believing that their perceptions about a particular topic are true as assumed. Thus, in order to curb the phenomenon, we have developed a Fake news filtering system that takes input from the user and classify it to be true or fake.

In this project, the task of automatic identification of fake news was addressed. The project developed classification models that rely on a combination of lexical, syntactic, and semantic information, as well as features representing text readability properties. The datasets are static datasets that can be updated from time to time. Through these, we can only test the data which is present in the predefined datasets.

6.3 Recommendation

It is recommended that government should adopt this fake news filtering system as developed in this project so as to detect fake news on social media and stop it from circulating.

REFERENCES

- [1]. K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.
- [2]. D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild et al., "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
- [3]. K. Shu, H. R. Bernard, and H. Liu, "Studying fake news via network analysis: detection and mitigation," in *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*. Springer, 2019, pp. 43–65.
- [4]. K. L. Arega, "Social media fake account detection for afanoromo language using machine learning," 2020.
- [5]. C. Shao, G. L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, and F. Menczer, "The spread of low-credibility content by social bots," *Nature communications*, vol. 9, no. 1, pp. 1–9, 2018.
- [6]. D. DiFranzo and K. Gloria-Garcia, "Filter bubbles and fake news," *XRDS: Crossroads, The ACM Magazine for Students*, vol. 23, no. 3, pp. 32–35, 2017.
- [7]. F. Zimmer, K. Scheibe, M. Stock, and W. G. Stock, "Fake news in social media: Bad algorithms or biased users?" *Journal of Information Science Theory and Practice*, vol. 7, no. 2, pp. 40–53, 2019.
- [8]. M. Granik and V. Mesyura, "Fake news detection using naive bayes classifier," in *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*. IEEE, 2017, pp. 900–903.
- [9]. F. A. Ozbay and B. Alatas, "Fake news detection within online social media using supervised artificial intelligence algorithms," *Physica A: Statistical Mechanics and its Applications*, vol. 540, p. 123174, 2020.
- [10]. V. Klyuev, "Fake news filtering: Semantic approaches," in *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2018, pp. 9–15.
- [11]. S. M. Sirajudeen, N. F. A. Azmi, and A. I. Abubakar, "Online fake news detection algorithm," *Journal of Theoretical & Applied Information Technology*, vol. 95, no. 17, 2017.
- [12]. C. Janze and M. Risius, "Automatic detection of fake news on social media platforms," *Pacis*, vol. 261, 2017.
- [13]. S. Gilda, "Notice of violation of ieeepublication principles: Evaluating machine learning algorithms for fake news detection," in *2017 IEEE 15th student conference on research and development (SCOREd)*. IEEE, 2017, pp. 110–115.
- [14]. R. M. Silva, R. L. Santos, T. A. Almeida, and T. A. Pardo, "Towards automatically filtering fake news in portuguese," *Expert Systems with Applications*, vol. 146, p. 113199, 2020.
- [15]. V. Singh, R. Dasgupta, D. Sonagra, K. Raman, and I. Ghosh, "Automated fake news detection using linguistic analysis and machine learning," in *International conference on social computing, behavioral-cultural modeling, & prediction and behavior representation in modeling and simulation (SBP-BRIMS)*, 2017, pp. 1–3.
- [16]. D. Varshney and D. K. Vishwakarma, "Hoax news-inspector: a realtime prediction of fake news using content resemblance over web search results for authenticating the credibility of news articles," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–14, 2020.

TABLE I: CLASSIFICATION MATRIX OF EXISTING LITERATURE/ METHODOLOGIES ON FAKE NEWS DETECTION

S/N0	Title	Main Idea/ Description/Methodology	PRO'S
1	Gilda, S. (2017, December). Evaluating machine learning algorithms for fake news detection. In 2017 IEEE 15th Student Conference on Research and Development (SCORED) (pp. 110-115). IEEE. [13]	<ul style="list-style-type: none"> • Detecting fake news through the application of NLP • TF-IDF of bi grams and probabilistic context-free grammar (PCFG) was applied to a corpus of about 11,000 articles. • The datasets were trained and tested on multiple machine-learning classification algorithms. 	<ul style="list-style-type: none"> • TF-IDF is potentially predictive of fake news. • The stochastic gradient Descent Model is the best-performing model when trained on the TF-IDF feature set only.
2	Silva, R. M., Santos, R. L., Almeida, T. A., and Pardo, T. A. (2020). Towards automatically filtering fake news in Portuguese. Expert Systems with Applications, 146, 113199. [14]	<ul style="list-style-type: none"> • They proposed and manually built a public and real labeled (true or false) dataset of fake news in Portuguese. • Linguistic base features and features generated through text representation techniques (BoW, wordsvect and fasttext) were used to train the model. • They tried to determine the best features or combination of features and the best machine learning methods to be used for the automatic detection of fake news. 	<ul style="list-style-type: none"> • Bow constitutes the best set of features for fake news filtering. • Combination of the results in BoW with the results obtained with linguistic-based feature using ensemble and stacking of classifiers outperformed results obtained by individual classifiers.
	Granik, M., and Mesyura,	<ul style="list-style-type: none"> • A software system was 	
	Conference	<ul style="list-style-type: none"> • Linguistic 	yield high