Investigating the Efficacy of Artificial Intelligence and Natural Language Processing for Fake News Detection

Emad E. Abdallah[1], Kinda A. Magableh, Feras Hanandeh[2], Alaa E. Abdallah[3]

[1] Department of Information Technology, Faculty of Prince Al-Hussein Bin Abdullah II for Information Technology, The Hashemite University, P.O. Box 330127, Zarqa 13133, Jordan
[2] President of Ajloun National University Box Office 43 - Ajlun - 26810 Jordan
[3] Department of Computer Science, Faculty of Prince Al-Hussein Bin Abdullah II for Information Technology, The Hashemite University, P.O. Box 330127, Zarqa 13133, Jordan

***Abstract*— The fast growth of internet platforms and social media has led to in an incredible flood of information, making it increasingly difficult to distinguish between authentic news and fake content. Fake news poses serious threats to public debate, and societal stability. To solve this issue, researchers used machine learning and natural language processing (NLP) approaches to create automated systems for detecting fake news.**

This paper provides an overview of cutting-edge strategies for detecting false news using machine learning and NLP. We investigate text preparation techniques such as tokenization, stemming, and stop-word removal to prepare data for analysis. We investigated our findings using a variety of machine learning algorithms, including logistic regression, decision trees, support vector machines, random forests, gradient boosting, and neural networks.

We demonstrate that the combination of machine learning and NLP techniques provides fascinating possibilities for countering the spread of fake news. We were able to improve our ability to detect deceptive content. Random Forest (RF), CatBoost, and XG Boost are among the best-performing algorithms. When applied to the Fake News Net dataset, the Random Forest method demonstrated exceptional performance inside experiments using a partitioning ratio of 70-30 for training and testing, respectively, with a stunning accuracy of 99.02%.

***Index Terms*— fake news, machine learning, natural language processing (NLP), real news, feature extraction.**

I. Introduction

As an increasing portion of our time is spent engaging online via social media platforms, a growing number of individuals are leaning to look for news through social media instead of relying on traditional news organizations. The factors contributing to this shift in consumption patterns are rooted in the characteristics of these social media platforms. Firstly, news consumption on social media is often prompter and more cost-effective when compared to traditional news media like newspapers or television. Secondly, sharing, commenting on, and discussing news articles with friends or other readers on social media is simpler.

Despite the advantages offered by social media, the quality of news on these platforms is generally lower compared to traditional news organizations. This is primarily due to the affordability of online news provision and the speed and ease with which news can be disseminated through social media. Consequently, a significant volume of fake news circulates widely. The widespread dissemination of fake news can have severe adverse effects on individuals and society. Firstly, it disrupts the authenticity and balance of the news system. Secondly, fake news deliberately influences consumers to adopt biased or false beliefs. Thirdly, fake news alters the way people interpret and respond to genuine news. The conventional solution to this task is to ask professionals, such as journalists, to check claims against evidence based on previously spoken or written facts. However, it is time-consuming and expensive [1].

Detecting fake news has been a significant challenge in the era of digital media, were information spreads rapidly and often without proper verification. Researchers have conducted numerous studies and developed various approaches to identify and combat the dissemination of fake news. Certainly, detecting fake news has been a topic of interest for researchers for quite some time. Here are a few earlier studies and approaches related to detecting fake news: The authors of [2]. In this paper, the authors defined fact-checking as a computational task and generated a dataset to support it. They focused on allegations made in political discussions and news stories and looked into strategies for determining their reliability.

The authors of [3] suggested a system for detecting fake news based on multiple textual indicators, including linguistic patterns, transmission patterns, and user interaction. They employed machine learning algorithms to determine if news articles were fake or true. The authors of [4] provide a study that specifies the types of fake ads that are made online, whether through spam or prominent advertising sites such as Craigslist

DOI:http://doi.org/10.24086/cocos2024/paper.1536

or online dating, and we show how to prevent ad fraud.

The authors of [5] focused on stylometric criteria to distinguish between hyperpartisan and false news parts. Researchers identified language characteristics and patterns of false news items. This study [6] investigated the automatic categorization of rumors and false material on social media using machine learning techniques. They employed language features, dissemination features, and user interactions to detect misleading content.

In [7], researchers investigated the dissemination of true and deceptive data on Twitter and discovered that false news spreads far faster and wider than true news. They identified numerous characteristics that can assist in identifying between true and false news, including language style, sentiment, and content type. In [8], the authors suggested a system for automatically detecting fake news based on language and contextual variables. They trained machine learning models for detecting fake news using variables such as sentiment, readability, and content.

In [9] This survey article gives an overview of different computational fact-checking methodologies. It discusses approaches such as content analysis, network analysis, and user behavior analysis to identify deception and fake news. [10] introduces Fake News Net, a large dataset for researching fake news on social media. The dataset consists of news items, user engagements, and contextual information. The authors highlight the difficulty of gathering and interpreting false news data, in addition to providing baseline studies for detecting fake news.

The work [11] examines the stylometrics of hyper partisan and fraudulent news items. The authors investigate language traits, writing styles, and trends to discern between trustworthy and untrustworthy news sources. They describe studies on a large dataset of news stories to show that stylometric techniques are excellent in detecting fake news. This study of [12] focused on detecting fake news on social media platforms. The authors proposed a framework that incorporates linguistic, social, and network-based features for classification. Their experiments demonstrated the effectiveness of their approach in identifying fake news.

II. PROPOSED METHODOLOGY

The proposed research examines and discusses key areas of false news detection. The dataset used in this study was carefully organized. Several necessary operations were carried out during the data preprocessing phase, including the removal of URLs, punctuation marks, and empty columns, to ensure data cleanliness and consistency. Following that, text preparation was completed, which included tokenizing longer texts to split them into individual words or shorter lines. This technique converted raw text into a format that could then be examined and modeled.

The following step was to extract textual features from the preprocessed data. This feature extraction stage is important in gathering the information and patterns needed to efficiently train the model. The model was then trained using several Machine Learning (ML) techniques. These strategies were combined with retrieved features to improve the model's performance. A variety of evaluation measures were used to assess the efficacy and accuracy of the ML algorithms used.

These metrics provided a quantitative measure of the model's performance, allowing for a complete analysis of its features and potential areas for improvement. In conclusion, this study takes a methodical approach, beginning with data collection and preprocessing, feature extraction, and model training. To evaluate the model's performance, specific assessment metrics are used, giving a complete analysis and useful insights. Figure 1 depicts the recommended methods for distinguishing between fake news and real news.

A. Data Collection

A recent dataset was used in this research to train and evaluate fake news detection models. The dataset Fake News Net has total records: 23481. This dataset includes both fictional and actual news stories sourced from fact-checking websites PolitiFact and GossipCop. It includes news item content, tweets about news articles, and social activities like replies, retweets, and favorites. Overall, the dataset contains 13749 genuine news articles and 9732 fake news articles [13]. This dataset has been used previously to train machine learning models and extract features for identifying fake news (Github, 2020).

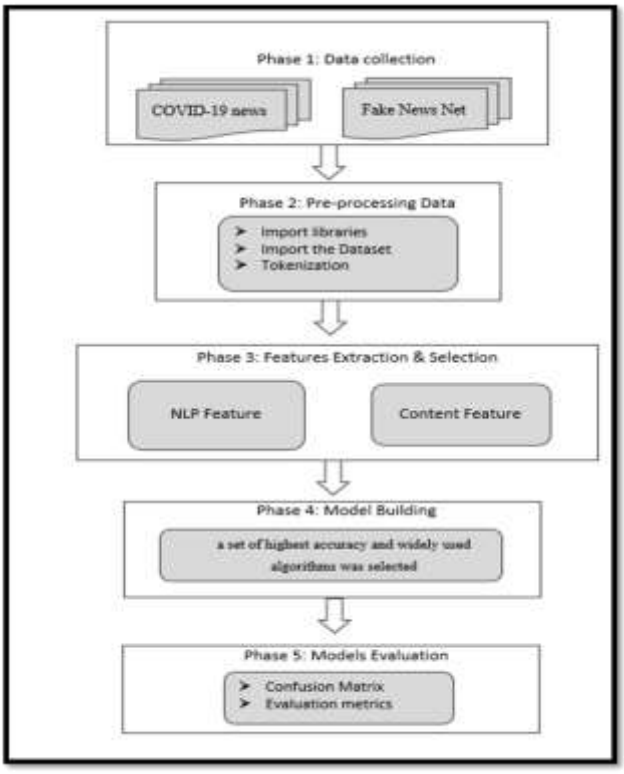

Figure 1. The Proposed Methodology for detecting fake news

DOI:http://doi.org/10.24086/cocos2024/paper.1536

B. Pre-processing Data

This work takes a clean dataset into account. But the dataset needs to be cleaned up of some unused symbols that have an impact on how the news is ultimately categorized. As part of the preprocessing step, URLs are eliminated from the dataset to eliminate any extra symbols, such as punctuation marks, and empty columns from the text of the news articles. The models were then made by taking features out of the text that had already been processed. Finally, the output of these steps is stored in a CSV file.

- Import libraries

Before beginning to build the model, import libraries, which are functions that perform actions without requiring the user to write code.

- Import the Dataset

The dataset used in this thesis is composed in CSV format.

- Remove Missing Values

Missing data in datasets must be removed or replaced in order for machine learning algorithms to predict accurately.

- Tokenization

The data is subjected to a crucial word tokenization procedure in which the text of each record is divided into distinct words using the delimiter (,). By using this tokenization technique, the text is converted into a random "bag of words" (BOW) model, which is represented by vectors of set and predetermined lengths. The frequency of specific words appearing in the text is encoded by these vectors. This tokenization procedure follows the "directed" methodology frequently used with BOW models.

The BOW model is based on the identification of a vocabulary, which is a set of words taken from the record. The delimiter (,) is used to differentiate and isolate each word, resulting in the successful construction of the vocabulary. The number of times each word appears in the record is methodically counted and logged. As a result, a numerical representation of the text's content develops, accurately capturing the word frequencies required for further analysis and modeling tasks. The BOW model's robustness originates from its capacity to generate a compact and structured representation of textual data, which allows for more efficient and effective data processing and the extraction of important insights from the dataset.

C. Feature extraction

This work's main contribution is the extraction of significant features from the fake news dataset. As it narrows the feature space's dimension by taking only the most crucial features into account. Feature extraction is crucial to text processing. In our work, the named-entity recognition (NER) method is employed to extract the features. Unstructured text can be categorized using the NER, features are produced from the relevant fake news dataset for this study. Table 4.2 displays the features that were extracted.

The output of this process is stored in 21 files, content features and NLP features all in CSV format. In an additional step when studying the data, there are non-numeric values for the Encoding feature, which must be converted into numeric values to be used as inputs for Machine learning algorithms. This is called encoding categorical variables. The values of the Encoding feature do not require the order, so the Label encoding method was applied, which is the method that assigns a numerical value to each category without regard to the order.

III. EXPERIMENTAL RESULTS

The investigations on the false news dataset are carried out using Google Colab, an open-source cloud-based graphical processing unit (GPU) platform provided by Google Inc. Python 3.7 is the programming language utilized to implement the machine learning techniques.

The ML algorithms are assessed and validated using a variety of training and testing methodologies on the fake news dataset. The effectiveness of machine learning algorithms is evaluated using a variety of metrics, including accuracy, precision, recall, and the F1 measure. Furthermore, we discuss the machine learning algorithms' performance on the bogus news dataset before and after feature extraction.

Two tests were carried out on the six chosen ML algorithms, which included training on two acquired datasets. The primary difference between the experiments is the algorithm training method. In all experiments we employed a 70-30 split training testing strategy.

We used the Fake News Network dataset. The dataset was utilized to train the model in the experiment, which included six machine learning techniques. The experiment tests and validates the ML algorithms using 30% of the false news dataset, while the remaining 70% is used to train the ML algorithms. The first experiment was separated into three segments, each with a distinct type of feature added. The features and outcomes will be provided in the sections below:

A. Fake News Experiment Results after content feature extraction only

In this part of the experiment, we use content features only. Table 1 compares the accuracy, precision, recall, F measure for the six used ML which clearly shows that the proposed approach achieves high performance accuracy. Evaluating the accuracy metrics, using Fake News Net dataset the "Random Forest algorithms" exhibit exceptional performance, achieving an impressive accuracy of 99.2%. Moreover, this algorithm demonstrates high precision with a score of 1, indicating a remarkable ability to correctly classify true positives. The recall score of 0.99 and the F1 score of 0.99 reflects the balance between precision and recall. Furthermore, the "CATBOOST algorithms and "XGBOOST" algorithms emerge as the second-

DOI:http://doi.org/10.24086/cocos2024/paper.1536

best performer, attaining an accuracy rate of 98%. These algorithms also precision with a score of 0.97, the recall score of 0.99 and the F1 score of 0.98 confirms its overall robustness in achieving a balance between precision and recall.

Table 1. Fake News Experiment Results After Content Feature Extraction Only

	algorithms	precision	recall	f1-score	accuracy
Fake News Net	Naive Bayes algorithms	0.93	0.94	0.93	92.95%
	Random Forest algorithms	1.00	0.99	0.99	99.2%
	XGBOOST algorithms	0.97	0.99	0.98	98%
	CATBOOST algorithms	0.97	0.99	0.98	98%
	Gaussian NB algorithms	0.93	0.93	0.93	93%
	K-Neighbors Classifier	0.94	0.86	0.90	90.4%

These findings provide clear evidence of the strengths and capabilities of the algorithms tested on the Fake News Net dataset. The "Random Forest algorithms" got good accuracy and precision ratings, while the "CATBOOST" and "XGBOOST" algorithms performed outstandingly, demonstrating their potential utility in detecting and classifying fake news situations. Figure 2 displays an algorithmic accuracy comparison.

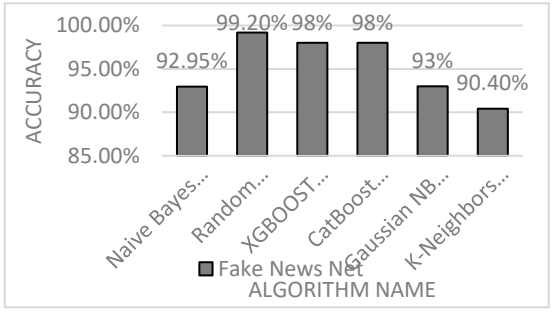

Figure2. Accuracy Achieved in first Experiment After Content Feature Extraction Only

B. *Fake News Experiment Results after extract both type of features content and NLP*

In this part of the experiment, we use both NLP and content features. According to the results in Table 2, when using Fake News Net, the "Random Forest" algorithm regularly outperforms the others, with an astonishing 99.5% accuracy rate. This model achieves a precision score of 1, showing few false positives, as well as a respectable recall score of 0.99 and an F1 score of 0.99, demonstrating its resilience in identifying between fake and genuine news articles.

The "CATBOOST" and "XGBOOST" algorithms come in second with an accuracy percentage of 98.6%. Despite not outperforming the Random Forest, both algorithms are highly effective in accurately identifying news stories, demonstrating their usefulness in countering misinformation.

Table 2. Fake News Experiment Results After Extract Both Type of Features Content and NLP

	algorithms	precision	recall	f1-score	accuracy
Fake News Net	Naive Bayes algorithms	0.93	0.94	0.94	94%
	Random Forest algorithms	1.00	0.99	0.99	99.5%
	XGBOOST algorithms	0.97	0.99	0.98	98.6%
	CATBOOST algorithms	0.97	0.99	0.98	98.6%
	Gaussian NB algorithms	0.93	0.93	0.93	93.89%
	K-Neighbors Classifier	0.94	0.86	0.90	92.4%

As shown in table 2, the used algorithms are capable of detecting fake news the with high accuracy, Next figure 3 showing the best algorithms accuracy achieved using the dataset Fake News Net. However, the superiority of the Fake News Net dataset appears evident.

First experiment summary, a random forest classifier is a adaptable machine learning algorithm known for its robustness, accuracy, and ability to handle complex datasets. Its key features include ensemble learning, random feature selection, and the use of decision trees, making it a valuable tool in various classification and regression tasks, including those involving high-dimensional data or noisy datasets. Therefore, it is expected to obtain the highest results in this experiment. Several experiments with various training ratios and training

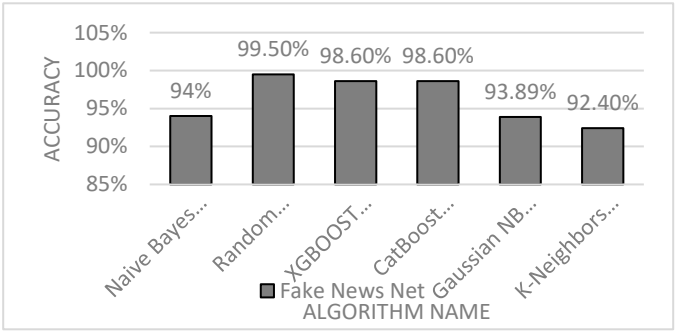

Figure 3. Best Accuracy Achieved in First Experiment After Extract Both Type of Features Content And NLP

methodologies yielded to that the Random Forest approach exceeded other algorithms in terms of accuracy, reaching 99.5%.

IV. CONCLUSION

In this research, we suggested a machine-learning system for detecting fake news based on natural language processing approaches. We evaluated the suggested method's performance using the Fake News Net dataset. The random forest and XGBOOST algorithms outperformed the others, achieving a high accuracy of 99.5% for the random forest and 99.65% for the XGBOOST algorithms. We reported two experiments in this paper. The experiments aimed to determine the most effective feature computations for detecting fake news. In every investigation, the ML algorithms are tested and confirmed on 30% of the false news dataset, while the remaining 70% is used to train the ML algorithm. For future work, we intend to investigate strategies such as deep learning, adversarial

DOI:http://doi.org/10.24086/cocos2024/paper.1536

training, and ensemble methods to improve the robustness of the proposed false news detection models. This may lead to higher accuracy rates.

REFERENCES

[1] Derakhshan, & Wardle. (2017). Information disorder. Toward an interdisciplinary framework for research and policymaking. Council of Europe.

[2] De Beer, D., & Matthee, M. (2021). Approaches to identify fake news: a systematic literature review. Integrated Science in Digital Age, 13-22.

[3] hu, K., Mahudeswaran, D., Wang, S., Dongwon Lee, & Liu., H. (2020). Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. . Big data, 8(3):171–188.

[4] Alzghoul, J. R., Abdallah, E. E., & Al-khawaldeh, A. H. S. (2024). Fraud in Online Classified Ads: Strategies, Risks, and Detection Methods: A Survey. Journal of Applied Security Research, 19(1), 45-69.

[5] Potthast, M., K. J., Reinartz, K., Bevendorff, J., & Stein, B. . 2017)). A stylometric inquiry into hyperpartisan and fake news. arXiv preprint arXiv:1702.05638

[6] Ruchansky, N., Seo, S., & Liu., a. Y. (2017). Csi: A hybrid deep model for fake news detection. In Proceedings of the ACM on Conference on Information and Knowledge Management,, pages 797–806.

[7] Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. science, 359(6380):1146–1151.

[8] Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017). Automatic detection of fake news. . arXiv preprint , arXiv:1708.07104.

[9] D'Ulizia, A., Caschera, M. C., Ferri, F., & Grifoni, P. (2021). Fake news detection: a survey of evaluation datasets. PeerJ Computer Science, 7, e518.

[10] Shu, K., Mahudeswaran, D., Wang, S., Dongwon Lee, & Liu., H. (2020). Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. . Big data, 8(3):171–188.

[11] Potthast, M., Kiesel, J., Reinartz, K., Janek Bevendorff, & Stein., B. (2017). A stylometric inquiry into hyperpartisan and fake news. . arXiv preprint arXiv:1702.05638

[12] Lazer, David MJ, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger et al. "The science of fake news." Science 359, no. 6380 (2018): 1094-1096.

[13] Github. (2018). Retrieved from Github: https://github.com/KaiDMML/FakeNewsNet

DOI:http://doi.org/10.24086/cocos2024/paper.1536