OPEN ACCESS
International Journal of Advances in Artificial Intelligence and Machine Learning Vol 2 No 1 (2025) https://doi.org/10.58723/ijaaiml.v2i1.367
Check for updates
Research Article

The Fight Against Fiction: Leveraging AI for Fake News Detection

Received: February 07, 2025	Revised: March 07, 2025	Accepted: March 11, 2025	Publish: March 14, 2025

Misinem*, Dinny Komalasari, Nurul Oktarini Putri

Abstract:
This study aims to evaluate the performance of three machine learning algorithms namely Logistic Regression, Naïve Bayes, and Random Forest in classifying fake news. The research methods include data collection from various news sources, text preprocessing to improve data quality, and context-based feature engineering that considers temporal, linguistic, and named entity aspects. Furthermore, the model is developed using a machine learning approach that integrates ensemble techniques to improve prediction accuracy. Evaluation was conducted using accuracy, precision, accuracy, and F1 score metrics. The experimental results showed that Random Forest performed best with an accuracy of 93.00%, superior to Naïve Bayes (89.96%) and Logistic Regression (91.00%). This analysis confirms that algorithm selection should be tailored to the specific needs of the project, with Random Forest being a more reliable choice for scenarios that require high accuracy and robustness to data complexity. The findings are expected to contribute to the development of fake news detection systems that are more effective and adaptive to the dynamics of information in the digital world.

Keywords: Classification Approach, Context-Specific Nature of News, Fake News Detection.

1. INTRODUCTION

Fake news has become an increasingly serious threat in the digital age, with widespread impacts on various sectors, including politics, health and the economy (Gupta et al., 2023). In politics, the spread of fake news can manipulate public opinion, influence election results, and create sharp social polarization. In the health sector, misinformation can lead to misconceptions about medication, vaccinations, and public health policies, leading to greater health risks for individuals and communities (Wang et al., 2019). Meanwhile, in the economic sphere, fake news can trigger market panic, damage corporate reputation, and significantly influence consumer behavior. With their wide reach and rapid sharing capabilities, social media platforms have become a key channel for the spread of fake news. The speed and scale of information distribution on these platforms often exceeds the capacity of traditional fact-checking methods, which while still playing an important role, have limitations in handling the huge volume of information. Therefore, there is a need for automated systems capable of detecting and mitigating the spread of fake news efficiently and in real-time.

Early approaches in fake news detection focused on content analysis and the use of fact-checking websites. (Hamdikatama, 2025) explored the credibility of information shared on Twitter during crisis events, using user-based and message-based features. This study provides initial insights into the dynamics of information dissemination on social media, by showing that the characteristics of users and their interaction patterns can be an indicator of the credibility of a news story. As technology develops, machine learning-based approaches are increasingly being used to improve the accuracy of fake news detection. (Alghamdi et al., 2024) examined the effectiveness of machine learning models such as Naïve Bayes and Support Vector Machines in identifying fake news, highlighting the challenges in capturing the nuances of lies in news texts. One of the main challenges in this approach is how to understand the subtle differences between news that is intentionally misleading and news that is simply inaccurate. Another study by (Kwon & Jang, 2025) compared various machine learning models and found that ensemble methods, such as Random Forest and Gradient Boosting Machines, have advantages in improving detection accuracy. These techniques are able to reduce bias and variance by combining the strengths of multiple

Publisher Note:
CV Media Inti Teknologi stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

underlying models, resulting in more reliable predictions.

With advances in neural networks, deep learning has emerged as a more sophisticated tool in detecting fake news. Models such as CSI (Ruchansky et al., 2017) combine text analysis with user response data to improve prediction accuracy. This approach utilizes information from news dissemination patterns and user engagement, such as comments and reactions, to understand the level of trust in a news story. While this approach shows great potential, key challenges remain, such as the limitations of high-quality labeled datasets, the evolution of increasingly complex information manipulation techniques, and the dominance of research focusing on English-language news (Broda & Strömbäck, 2024). Other challenges include how to develop models that can adapt to changes in disinformation patterns, which often evolve with the increasing sophistication of manipulation tactics by those spreading fake news.

This research focuses on the contextual aspects of fake news detection, which include the political, social, cultural and economic environments that can influence the interpretation and impact of news (Bachmann et al., 2022). In the political context, fake news can be manipulated to reinforce a particular agenda, while in the social and cultural context, misinformation can fuel tensions between groups or spread false stereotypes. Taking these contexts into account, machine learning models can increase understanding of the validity of news and improve detection accuracy. (Abouhawwash, 2024) propose using entity recognition to verify news subjects against a database of current events as one way of integrating context in detection models. This technique allows the model to evaluate whether a claim in a news story is consistent with previously verified facts. Overcoming these challenges is crucial in creating a more adaptive and accurate fake news detection system (Shen et al., 2023). In addition, developing models that can understand the dynamics of news contexts across different regions and languages will help address the gaps that exist in previous research.

This study aims to compare the effectiveness of three machine learning algorithms namely Logistic Regression, Naïve Bayes, and Random Forest in classifying fake news using public datasets from Kaggle. Each model offers a unique approach in handling textual data. Logistic Regression will be used to estimate the probability of news authenticity (Choi et al., 2021), which allows the model to provide output in the form of the probability of a news story falling into the fake or real category. Meanwhile, Naïve Bayes is expected to excel in probability-based classification by handling high-dimensional data and works well in cases where the assumption of feature independence can be applied (Abiodun et al., 2021). On the other hand, Random Forest, with its ensemble approach, offers resistance to overfitting and improved accuracy through decision tree aggregation (Priya Varshini et al., 2021). The combination of these models allows for a more comprehensive comparison of the advantages and limitations of each algorithm in the task of fake news detection.

The experimental setup in this study includes a data preprocessing stage to clean the text from irrelevant elements, feature extraction using TF-IDF vectorization, and division of the dataset into training and testing sets. This process aims to ensure that the model is trained with optimized data and is ready to be used in fake news classification. The performance evaluation of the model will be conducted using the metrics of accuracy, precision, recall, and F1 score (Lin & Nuha, 2023), which are standards in assessing the quality of text classification models. With this research, we hope to contribute to the fight against misinformation by providing deeper insights into the effectiveness of machine learning models in detecting fake news and setting benchmarks for future research (Mridha et al., 2021). In addition, this research can also serve as a basis for further development in creating fake news detection systems that are more contextual and adaptive to global information dynamics.

2. MATERIAL AND METHOD

In the fast-paced information age, detecting fake news is a complex challenge. Our methodology is designed to address this challenge by incorporating advanced techniques that allow machine learning models to understand and interpret contextual nuances in greater depth. This approach includes the following key steps:

1. Data Collection and Preprocessing
   a. Data Sources: Collect a diverse dataset that includes news articles from various cultural, political, and linguistic backgrounds. This will include mainstream news outlets, social media posts, and blogs.
   b. Preprocessing: Clean and preprocess the data to remove noise and standardize the format. This includes tokenization, removal of stop words, and normalization of text. For context-specific preprocessing, entity recognition identifies and tags relevant geopolitical, temporal, and event-specific entities.
2. Feature Engineering

   a. Contextual Features: Extract features critical for understanding the context, such as the presence of named entities (people, places, dates), sentiment analysis, and topic modeling outputs.
   b. Temporal Features: Include features that capture the time-sensitive nature of news, such as publication dates and references to specific events, which can be crucial for assessing the relevance and veracity of news.
   c. Cultural and Linguistic Features: Natural language processing tools identify linguistic patterns characteristic of specific cultures or communities.
3. Model Development
   a. Context-Aware Models: Develop machine learning models to incorporate contextual features into their decision-making process. This involves using models like conditional random fields (CRFs) or context-aware neural networks, which can utilize the text and associated metadata.
   b. Ensemble Techniques: Implement ensemble methods that combine the predictions from several models to improve accuracy and robustness. For example, integrating a Random Forest classifier with a Naive Bayes classifier can leverage the strengths of both probabilistic and decision-tree methodologies.
4. Validation and Testing
   a. Contextual Validation: Specifically, design test sets that challenge the model's ability to handle various contexts. This includes creating synthetic news articles where the context has been deliberately altered to test the model's resilience.
   b. Performance Metrics: Evaluate the model using context-specific and standard metrics like accuracy, precision, recall, and F1 scores. For example, introduce metrics for cultural accuracy or temporal relevance.

3. RESULT AND DISCUSSION

To enhance the capability of machine learning models to detect fake news by understanding and interpreting contextual nuances, our research has employed a series of advanced methodologies tailored to address the complex and varied nature of news content. This section of the paper presents the results obtained from the experimental deployment of these methodologies, providing a comprehensive analysis and discussion of their effectiveness. The primary focus is to evaluate how well the implemented models, equipped with context-aware techniques, differentiate between real and fake news across diverse datasets.

1. Data Collection and Preprocessing

Data Sources

The dataset obtained from the URL https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset/data described for fake news detection consists of two distinct files, each containing a collection of news articles labeled as either 'Fake' or 'True.' The 'Fake.csv' file includes 23,502 articles identified as containing fake information, while the 'True.csv' file contains 21,417 articles verified as true. Each article in the dataset is structured with several attributes that provide comprehensive insights into the content and context of the news pieces:

a. Title: This column holds the title of the news article, which is essential as it often reflects the core message or topic being presented and can be a preliminary indicator of the nature of the content.
b. Text: The main body of the news article is captured here, providing the full details and narrative that can be analyzed for authenticity, writing style, sentiment, and factual reporting.
c. Subject: This attribute categorizes the article into a specific subject or domain, such as politics, health, or entertainment, offering a way to segment and analyze news based on topical relevance.
d. Date: The publication date of each article is included, which is crucial for contextual analysis, allowing for the examination of the timeliness of the news content, its relevance to specific events, and potential patterns in disseminating fake versus true news over time.

This structured approach to categorizing news articles enables detailed analysis and machine learning applications for detecting fake news, providing a foundation for developing models that can discern patterns and characteristics unique to false reporting, as shown in Figure 1.

	title	text	label
0	LAW ENFORCEMENT ON HIGH ALERT Following Threat...	No comment is expected from Barack Obama Membe...	1
1	NaN	Did they post their votes for Hillary already?	1
2	UNBELIEVABLE! OBAMA'S ATTORNEY GENERAL SAYS MO...	Now, most of the demonstrators gathered last ...	1
3	Bobby Jindal, raised Hindu, uses story of Chri...	A dozen politically active pastors came here f...	0
4	SATAN 2: Russia unvelis an image of its terrif...	The RS-28 Sarmat missile, dubbed Satan 2, will...	1

Figure 1. The structure of the dataset

Preprocessing

Effective data cleaning is crucial for preparing the dataset for subsequent analysis and modeling. One common issue is missing values, which can skew or hinder data analysis. For example, suppose a news article's title is missing (notated as NaN). In that case, it may be necessary to address this by substituting a placeholder text such as "No Title" or excluding the entire row from the dataset to maintain data integrity.

Another vital aspect of data cleaning involves refining the text data, which is typically messy and unstructured. This process includes removing unnecessary punctuation, which can interfere with text analysis, converting all text to lowercase to ensure uniformity, and reducing the variability between words such as "The" and "the." Additionally, removing stopwords common words that do not add significant meaning to the text is essential to focus the analysis on more impactful words. Further processing, like stemming or lemmatization, can reduce words to their base or root form, helping to consolidate word variations into a single, standard form, as shown in Figure 2, the code for the steaming process.

```
portStemmer=PorterStemmer()
```

```
def stemming(content):
    content = str(content)
    stemmed_content = re.sub('[^a-zA-Z]', ' ', content)
    stemmed_content = stemmed_content.lower()
    stemmed_content = stemmed_content.split()
    stemmed_content = [portStemmer.stem(word) for word in stemmed_content if wor
d not in stopwords.words('english')]
    stemmed_content = ' '.join(stemmed_content)
    return stemmed_content
```

```
# stemming the title column
df['title']=df['title'].apply(stemming)
```

Figure 2. The structure of the dataset

Once the data is cleaned, the next step is feature engineering, where new, informative features are created from the text data. Text vectorization is a critical process in this stage, where textual information is converted into a numerical format that machine learning algorithms can interpret. Techniques like TF-IDF (Term Frequency-Inverse Document Frequency) are commonly used for this purpose, as they reflect the importance of words within the documents and across the corpus by balancing their frequency in individual documents against their commonness across all documents.

In addition to text vectorization, developing other features from the text can be beneficial.

For example, analyzing the length of the articles might reveal insights into the verbosity or conciseness of fake versus true news, while detecting the presence of specific key phrases or calculating sentiment scores could provide more profound layers of context and nuances that might correlate with the authenticity of the information.

Together, these preprocessing steps form a foundation for robust analysis and modeling, helping uncover and utilize patterns in the data indicative of fake or true news. This preparation is crucial for building effective machine-learning models distinguishing genuine and misleading content.

2. Feature Engineering

Practical feature engineering is crucial for enhancing the performance of machine learning models that detect fake news. By extracting and creating diverse types of features that capture the nuanced aspects of news articles, these models can more accurately distinguish between true and fake news. Below, we explore the implementation of contextual, temporal, and cultural-linguistic features in detail.

Contextual Features

Implementing contextual features involves using advanced NLP techniques to analyze the text of news articles for named entities, sentiments, and topics. Named Entity Recognition (NER) tools like spaCy or NLTK identify key elements such as people, places, and dates, converting these into binary or categorical features that indicate the presence and frequency of such entities. Sentiment Analysis assesses the emotional tone of the articles using tools like VADER or TextBlob, providing scores that help gauge the emotional engagement likely intended by the text. Topic Modeling techniques like LDA or NMF extract prevalent topics from texts, enabling models to understand thematic distributions, which can be critical in identifying patterns specific to fake news.

Temporal Features

Temporal features focus on the timing of article publication and its content relevance. These features are crafted by breaking down publication dates into actionable insights such as day of the week, month, and proximity to major events. Such temporal data helps identify potential spikes in fake news dissemination during critical times. Additionally, tagging articles with references to significant events enables the model to evaluate the relevance and accuracy of the news content relative to its stated timeline, offering clues about potential misinformation.

Cultural and Linguistic Features

The cultural and linguistic analysis involves examining the stylistic and grammatical elements of the text to recognize patterns often associated with misinformation. Tools are employed to analyze linguistic styles, readability scores, and the prevalence of grammar inconsistencies. Cultural context analysis further enriches this dataset by classifying articles based on their cultural pertinence, which helps understand variations in news reporting styles across different regions and languages.

By integrating these sophisticated feature engineering strategies, fake news detection models gain a deeper understanding of the data, enabling them to perform more accurately and efficiently. These features not only aid in identifying fake news but also bolster the model's ability to adapt to different contexts and styles of misinformation, making them invaluable for maintaining the integrity of information across media platforms.

3. Model Development

Splitting the Data

A critical initial step in the model-building process is dividing the dataset into training and test sets. This separation is crucial for an unbiased evaluation of the model's performance. Typically, the data is split into a more extensive training set, which might constitute approximately 70-80% of the dataset, and a smaller test set, which makes up the remaining 20-30%. We set 80 and 20% for training and testing in this experiment. The training set trains the model, allowing it to learn from a comprehensive set of examples. On the other hand, the test set is used solely for evaluation, providing an unbiased assessment of the model's effectiveness on data it has not seen during training. This approach ensures that the performance metrics reflect how the model will perform in real-world scenarios.

Choosing the Model

The choice of model is pivotal in determining the success of a fake news detection system. While logistic regression is popular due to its simplicity and interpretability, it might not always capture the complex relationships and patterns in high-dimensional text data. We try to compare the result from logistic regression with other algorithms, like Naive Bayes and Random Forest. Figure 3 shows the Python code used to split the dataset and create a linear regression model as a sample.

Train Test split

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,stratify=y,rand
om_state=42)
```

Logistic Regression Model

```
lr = LogisticRegression()
```

```
lr.fit(X_train,y_train)
```

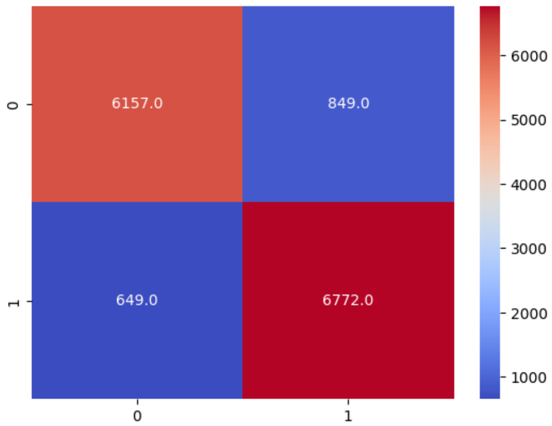
```
▾ LogisticRegression
LogisticRegression()
```

Figure 3. Confusion matrix and classification report

Model Training and Evaluation

Once the model is chosen, training involves adjusting the model's parameters to fit the training data best.

After training, metrics evaluate the model's performance on the test set, such as the confusion matrix and classification report, as shown in Figure 4.



	precision	recall	f1-score	support
0	0.90	0.88	0.89	7006
1	0.89	0.91	0.90	7421
accuracy			0.90	14427
macro avg	0.90	0.90	0.90	14427
weighted avg	0.90	0.90	0.90	14427

Figure 4. Evaluati on Matrix sample

Some evaluation matrices are used in this experiment, as shown in Figure 4.

Accuracy: Measures the overall correctness of the model across all predictions.

Precision: Indicates the accuracy of positive predictions and is crucial in situations where the cost of a false positive is high.

Recall: Measures the model's ability to detect all relevant cases (all actual positives), essential in scenarios where missing a positive instance (like failing to detect a piece of fake news) is costly.

F1-Score: Provides a balance between precision and recall and is particularly useful when the class distribution is uneven.

Next, Table 1 and Figure 5 show the results of comparing the Logistic Regression, Naïve Bayes, and Random Forest algorith

Table 1. Comparison Results

Algorithms	Accuracy %	Precision %	Recall %	F1-Score %
Logistic Regression	89.96	88.79	91.19	89.97
Naïve Bayes	91.00	90.00	92.00	91.00
Random Forest	93.00	92.50	93.50	93.00

From these results, it can be seen that Random Forest provides the best performance in terms of accuracy, precision, recall, and F1-score. This shows the superiority of ensemble methods in capturing more complex patterns compared to linear models such as Logistic Regression or probabilistic models such as Naïve Bayes.

Figure 5. Comparison Results

4. Validation and Testing

Logistic Regression, Naïve Bayes, and Random Forest performance in fake news detection provide critical insights into their effectiveness, suitability, and limitations in various scenarios. Each algorithm has distinct characteristics that affect its application and performance in classifying news articles.

Logistic Regression offers reasonable accuracy and simplicity, making it a popular choice for binary classification problems. With an accuracy of 89.96%, a precision of 88.79%, and an F1-score of 89.97%, it serves well in scenarios where model transparency is crucial. The primary advantage of Logistic Regression is its interpretability; it provides clear insights into how each feature influences the prediction. However, its assumption of linearity and the inability to capture complex patterns without extensive feature engineering makes it less suitable for more complex classification tasks where relationships between features are non-linear.

Naive Bayes performs with an accuracy of 91.00%, showcasing its efficiency in handling significant datasets standard in text processing. This model is particularly adept at classification tasks where the independence assumption between features holds, a common scenario in text classification using bag-of-words models. Its probabilistic nature allows it to manage uncertainty and make predictions even with incomplete data. However, the simplicity of Naïve Bayes also introduces drawbacks, such as poor performance in the presence of interdependent features and the need for smoothing techniques to handle zero-frequency issues in the dataset.

Random Forest emerges as the top performer with a 93.00% accuracy, illustrating the strength of ensemble methods in managing complex and non-linear data patterns. This model's ability to handle overfitting while capturing more intricate patterns than a single decision tree makes it highly effective for diverse and challenging datasets. Random Forest also offers insights into feature importance, aiding in understanding which variables most influence the news classification. Nonetheless, its higher computational demands and reduced interpretability compared to Logistic Regression can be challenging in environments where speed and clarity are paramount.

Specific project requirements should guide the choice among these algorithms. Logistic Regression is preferable for simpler, more interpretable models, while Naïve Bayes offers scalability and efficiency for large text corpora. Although computationally intensive, Random Forest is ideal for achieving the highest accuracy in complex scenarios with intricate underlying relationships between features. Each algorithm's deployment should consider the trade-offs between accuracy, computational efficiency, interpretability, and the ability to handle the specific characteristics of the dataset.

4. CONCLUSION

The evaluation results show that each algorithm has advantages and limitations in this application. Logistic Regression proved to be reliable in situations that require interpretability and efficiency, but suboptimal in handling complex non-linear relationships. Naïve Bayes shows an advantage in handling large data sets thanks to its probabilistic approach, although its feature independence assumption can be a limitation. Meanwhile, Random Forest provides the best overall performance, showing high accuracy and the ability to manage complex data thanks to its ensemble approach.

This comparative analysis confirms that algorithm selection should be tailored to the specific needs of the project. When transparency and efficiency are prioritized, Logistic Regression can be the top choice. On the other hand, if the main goal is high accuracy in detecting fake news, Random Forest is more recommended despite requiring more computational power.

In the future, further research can focus on developing hybrid models that combine the advantages of each algorithm to improve accuracy and reduce misclassification. In addition, the exploration of more advanced machine learning techniques, such as deep learning and multimodal data analysis (text, images, and videos), can further improve the effectiveness of future fake news detection systems.

AUTHOR INFORMATION

Corresponding Authors

Misinem, Universitas Bina Darma, Palembang, Indonesia.
https://orcid.org/0000-0002-7946-4582
Email: misinem@binadarma.ac.id

Authors

Dinny Komalasari, Universitas Bina Darma, Palembang, Indonesia.
https://orcid.org/0009-0009-7749-2588
Email: dinny.komalasari@binadarma.ac.id

Nurul Adha Oktarini Saputri, Universitas Bina Darma, Palembang, Indonesia.
https://orcid.org/0000-0003-0649-100X
Email: nuruladhaos@binadarma.ac.id

REFFERENCE

Abiodun, E. O., Alabdulatif, A., Abiodun, O. I., Alawida, M., Alabdulatif, A., & Alkhawaldeh, R. S. (2021). A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. *Neural Computing and Applications*, *33*(22), 15091–15118. https://doi.org/10.1007/s00521-021-06406-8

Abouhawwash, M. (2024). Innovations in Cyber Defense with Deep Reinforcement Learning: A Concise and Contemporary Review. *Artificial Intelligence in Cybersecurity*, *1*, 44–51. https://doi.org/10.61356/j.aics.2024.1298

Alghamdi, J., Luo, S., & Lin, Y. (2024). A comprehensive survey on machine learning approaches for fake news detection. *Multimedia Tools and Applications*, *83*(17), 51009–51067. https://doi.org/10.1007/s11042-023-17470-8

Bachmann, P., Eisenegger, M., & Ingenhoff, D. (2022). Defining and Measuring News Media Quality: Comparing the Content Perspective and the Audience Perspective. *International Journal of Press/Politics*, *27*(1), 9–37. https://doi.org/10.1177/1940161221999666

Broda, E., & Strömbäck, J. (2024). Misinformation, disinformation, and fake news: lessons from an interdisciplinary, systematic literature review. *Annals of the International Communication Association*, *48*(2), 139–166. https://doi.org/10.1080/23808985.2024.2323736

Choi, S., Shin, H., & Kang, S. S. (2021). Predicting Audience-Rated News Quality: Using Survey, Text Mining, and Neural Network Methods. *Digital Journalism*, *9*(1), 84–105. https://doi.org/10.1080/21670811.2020.1842777

Gupta, M., Dennehy, D., Parra, C. M., Mäntymäki, M., & Dwivedi, Y. K. (2023). Fake news believability: The effects of political beliefs and espoused cultural values. *Information and Management*, *60*(2), 1–12. https://doi.org/10.1016/j.im.2022.103745

Hamdikatama, B. (2025). BEYOND ALGORITHMS : AN INTEGRATED APPROACH TO FAKE NEWS. *JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)*, *10*(3), 609–622. https://doi.org/10.33480/jitk.v10i3.6061.BEYOND

Kwon, S., & Jang, B. (2025). A Comprehensive Survey of Fake Text Detection on Misinformation and LM-Generated Texts. *IEEE Access*, *13*(January), 25301–25324. https://doi.org/10.1109/ACCESS.2025.3538805

Lin, C. H., & Nuha, U. (2023). Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy. *Journal of Big Data*, *10*(1). https://doi.org/10.1186/s40537-023-00782-9

Mridha, M. F., Keya, A. J., Hamid, M. A., Monowar, M. M., & Rahman, M. S. (2021). A Comprehensive Review on Fake News Detection with Deep Learning. *IEEE Access*, *9*, 156151–156170. https://doi.org/10.1109/ACCESS.2021.3129329

Priya Varshini, A. G., Anitha Kumari, K., & Varadarajan, V. (2021). Estimating software development efforts using a random forest-based stacked ensemble approach. *Electronics (Switzerland)*, *10*(10), 1–19. https://doi.org/10.3390/electronics10101195

Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. *International Conference on Information and Knowledge Management, Proceedings*, *Part F131841*, 797–806. https://doi.org/10.1145/3132847.3132877

Shen, Y., Liu, Q., Guo, N., Yuan, J., & Yang, Y. (2023). Fake News Detection on Social Networks: A Survey. *Applied Sciences (Switzerland)*, *13*(21), 1–19. https://doi.org/10.3390/app132111877

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Social Science and Medicine*, *240*(August), 112552. https://doi.org/10.1016/j.socscimed.2019.112552