International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.41, September 2023 – www.ijais.org

Enhancing Fake News Identification in Social Media through Ensemble Learning Methods

Timothy Moses
Senior Lecturer,
Department of Computer Science,
Faculty of Computing, Federal University Lafia
P.M.B 146, Lafia, Nasarawa State, Nigeria

Henry Egaga Obi
M.Sc. Student,
Department of Computer Science,
Faculty of Computing, Federal University Lafia
P.M.B 146, Lafia, Nasarawa State, Nigeria

Christopher Ifeanyi Eke
Senior Lecturer,
Department of Computer Science,
Faculty of Computing, Federal University Lafia
P.M.B 146, Lafia, Nasarawa State, Nigeria

Jeffrey Agushaka
Senior Lecturer,
Department of Computer Science,
Faculty of Computing, Federal University Lafia
P.M.B 146, Lafia, Nasarawa State, Nigeria

ABSTRACT
The proliferation of deliberately misleading information, commonly known as fake news, poses a significant challenge in shaping public opinions. This paper presents a cutting-edge methodology for effectively identifying and combating fake news by harnessing the power of ensemble learning techniques. Recognizing the widespread influence of fake news and its detrimental societal effects, there is an urgent need for robust and adaptable identification models. Existing approaches often suffer from biases and lack adaptability due to their reliance on single algorithms or limited datasets. To address these limitations, the study introduces an ensemble learning model that incorporates a diverse range of algorithms, enhancing accuracy and adaptability across various fake news contexts. Leveraging a benchmark dataset, the established model attained an exceptional accuracy rate of 97.86% using the test dataset, outperforming existing architectures. Through this research, the researchers aim to mitigate the adverse impact of fake news on social media platforms and provide a more reliable means of content verification.

General Terms
News Classification, Text Recognition.

Keywords
Machine Learning – Fake News – Ensemble Learning – Identification Models – Social Media

1. INTRODUCTION
The prominence and dissemination of fake news have endured and expanded throughout history, predating the emergence of the internet [1]. This assertion, supported by research, demonstrates a longstanding pattern of influential figures throughout history leveraging media outlets to shape and control narratives that align with their interests [2]. Subsequent studies have made attempts to elucidate the characteristics associated with attributing information to fake news. Various terms have been used, including lies, falsehood, selective falsification, and intentional deception, all aiming to describe the act of deliberately misleading a specific audience [3]. But another research like the one carried out by Kaliyar [4], ascribed the prevalence of fake news in society to stem from readers' tendency to swiftly embrace information that aligns with their beliefs, without thoroughly verifying its authenticity before sharing or spreading it.

The impact of fake news has been widespread, affecting numerous individuals, including prominent figures in society. This is evident as even the former President of the United States of America expressed the sentiment that "Everything is true and nothing is true" [5]. The propagation of misinformation has played a role in deepening societal divisions, as different groups gravitate towards information that aligns with their individual lives or interests. A notable illustration of this phenomenon is a study revealing that 75% of conservatives in the United States perceive news from the Cable News Network (CNN) as fake, while 59% of liberals hold a similar view towards reports from Fox News [6]. This is no surprise, as observed by Mark Thompson, the Chief Executive Officer (CEO) of the New York Times, that "falsehood flies and the truth comes limping at it" [5]. This indicates that the impact of fake news extends beyond individuals and also affects corporate entities and government-owned institutions. As a result, it is crucial to implement measures for identifying and combating fake news to mitigate the challenges it poses to society as a whole.

Various approaches have been adopted, with organizations opting to educate information seekers. However, the automated approach, which leverages machine learning (ML), stands out as it minimizes human involvement and offers promising solutions [7]. There are other efforts to combat fake news that go beyond just the automated methods. Some employ additional strategies such as analyzing writing patterns, distribution styles, and source legitimacy to identify and address false information [8]. In their research, other scholars focused on the identification and classification of online content that falls under the umbrella of fake news.

They successfully outlined seven distinct categories, which include journalism, false news, misreporting, persuasive information, commentary, and satire [9].

Although efforts have been made to tackle fake news through the utilization of the Natural Language Processing (NLP) method [10], others evaluated three different datasets on several ML algorithms [11]. However, this study aims to enhance and make a valuable impact on the field of fake news identification by employing a synergistic combination of multiple algorithms. The objective is to improve the accuracy and effectiveness of identifying fake news by utilizing various textual properties extracted from a publicly accessible domain. The dataset utilized in this study comprises 20,800 news articles, specifically labeled as 1 for fake news and 0 for true news. To enhance the performance accuracy of the model, an ensemble learning technique is employed, harnessing the strengths of several ML algorithms, including Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), and Gradient Boost (GB) classifier. Through the amalgamation of these algorithms trained using the best-performed feature extraction technique, an advanced fake news identification model is developed, aiming to achieve significantly improved accuracy in the identification process.

Therefore, this study contributes to knowledge through the following:

i. The research significantly enhances the understanding of both the concept of fake news and the models employed for its identification.

ii. The model unveils the ensemble learning method, utilizing it to achieve a superior performance in comparison to the existing model.

iii. The study employs the textual properties of both fake and true news articles to add clarity to the classification of fake and true news.

iv. The research explored several feature extraction techniques to train the model using the best dataset features extracted with the most suitable feature extraction technique.

v. Uniquely, multiple ML algorithms are applied to evaluate a publicly available dataset on Kaggle. The performance of the entire model is then measured using a comprehensive confusion matrix that encompasses accuracy, F1 score, recall, and precision.

The next section exposes various definitions, effects, and ML studies carried out by research to help reduce the impact of the spread of fake news.

2. LITERATURE REVIEW

This section takes a look at several research works carried out concerning the identification of fake news. The inquisitive nature of humans has made many people sort information for various reasons and from various sources without checking if the source is legitimate or not. The phrase fake news has drawn a lot of attention but it is still important to understand the scholarly acceptable definition of the term

2.1 Definition of Fake News

The term fake news has evolved and people tend to misunderstand the extent to which it covers. Even as it has blown and played a part in some crises in democracy, justice, politics, journalism, economics, and public trust [8]. That's why it's important to be acquainted with different views about the term fake news. In an attempt to define fake news, Gelfert [12] opined that fake news should be used for instances where the act is typically aimed at falsehood, misinformation, or misrepresenting assertions and disseminating the same as news. In the same vein, a study shows that the term has expanded beyond just meaning false information to including scamming followers with defamation against those who hold opposing political views [9]. Also, satirical stories along with the intent to create such, are considered when defining the term fake news. This is because sometimes a satirical story may appear to be too real and can end up confusing readers [9]. In another study, the author Tandoc et al. [13] believed that fake news is incomplete without an audience. That is, the in-authenticity of any news is dependent on how it is perceived by the targeted audience. Another research referred to fake news as "Frankfurtian bullshit asserted in the form of news publication" [14]. A separate opinion by Wang [15], considers fake news as a report that is mostly focused on new stories and happenings that contain wrong or false information but do not report the incorrectness of the information. Also, fake news has been described as information that is fabricated and structured in the form of legitimate news content [16]. It was further defined as a meddler of factual information so that it ends up being untruthful, misleading, and mostly deceives those who have seen, heard, or read the bullshit [12]. Research carried out by Baptista and Gradim[17] explains the term as a kind of online false information that contains misleading reports that are mostly related to happenings that are real, designed mainly to manipulate a particular set of persons using unscrupulous means to attain success. This explains why the government, companies, and legislative bodies have put in measures to limit the contact between citizens and false information due to the belief that it is not healthy in a democratic system [18]. In summary, the term is used to describe information that is not true, misleading, or fabricated, and its source cannot be traced, found, or verified.

2.2 Effects of Fake News

Different countries, states, and provinces have felt the harm caused as a result of fake news dissemination. If left to discuss the effect that fake news has caused on human life, it will be discussed at length. This section seeks to unmask various reports as carried out by different researchers regarding the effects caused by fake news.

A study was carried out to scrutinize various fabrications weeks before the Irish abortion referendum which was held for the amendment of the constitutional right of an unborn child to life. Murphy et al.[19] conducted an online survey that constituted both true and false events and the result showed that nearly part of the respondent recounted the memory of a fabricated event. The researcher proposed that electorates in actual election canvass are most likely to grow fake news that will support their opinion, mostly in cases where they have little intellectual power [19]. In India, an award was presented to Banaji [20] for conducting very good research on some of the atrocities caused by fabricated information disseminated on WhatsApp. The author [20] cited instances like lynching, killings, and injury carried out on the victimized set of persons in Indian societies, such as Christians, Muslims, Adivasis, and Dalits. This set of persons are accused of slaughtering cows, trafficking cows, childsnatching, or stealing human organs. Those who hold bad intent against this group of persons, fabricated lies that were mostly shared or disseminated by peer-to-peer communication through WhatsApp [20].

Another study carried out by Gaozhao[18] showed that fake news has affected democratic systems around the world by reducing the level of the genuineness of citizens in decision-making, as they are badly influenced by the different disseminated information they find on social media. This problem has prompted many governments, and other decision-making bodies to put in the effort to see how fake news contact with citizens can be limited, thereby increasing trust between the government and its citizens [18]. This goes a long way to justify the opinion of the researcher Linden et al. [21] who opined that the more fake news spreads and the top in the society discusses it, the further it kills democracy.

Talking about killing democracy, researchers have opined that the 2016 general election in the United States of America was concentrated on fake news and is considered the reason for the popularity of fake news in modern days [6]. Fake news has become one major tool used by politicians against their opponents. It is believed so because research by the scholar shows that 1% of Twitter users are responsible for 80% of political fake news [6]. Another researcher delved into the aspects of fake news affecting Italians, with a 2018 study comparing how fake news affected electorates in 2013 with how fake news affected them in 2018 [22]. The research shows that fake news affected the Italian electorates more in 2018 and favored the populist political parties more. Talking about the effects of fake news on elections, a report from the findings of an author [23] shows that fake news gains more ground in a country that runs a democratic system of government like America, hampering the free and fair selections of political office holders. When citizens of a country, state, or province are constantly exposed to fake news, it tends to affect their psychological well-being, thereby remaining uninformed or provoking them to make wrong decisions [24].

In the health sector, a study was carried out to create awareness of how the dissemination of fake news affects the health sector because of politicking. Using autism policies, the research shows different ways of spreading fake news about the methodological discipline of Applied Behaviour Analysis, which contributed to the support from North Americans whereas the Europeans were against it [25]. It doesn't just end there, the fight to understand how far fake news has eaten into and affected the society has been on. Such effort was undertaken by Jang [26] to verify the third-person effect of fake news. The study shows that most persons believe that fake news is likely to affect others rather than the fabricators themselves, with partisan identity, social undesirability subjects, and external political efficacy being the indicators of this fact [26].

2.3 Review of Machine-Learning-Related Works

Researchers did not only try to know the meaning of the term or understand the length at which fake news has ravaged society at large, but efforts have been put in place to find and isolate such information before it gets hold of information seekers who are mostly the catalyst for the dissemination of fake news they come across on social media platforms, which contributes as the most medium through which the fabricators of fake news use to get to their targets. An extensive review of some recent and notable ML models and systems developed to help the identification of fake news are highlighted below:

Firstly, the research looks at the study undertaken by Agudelo et al.[27], based on Naïve Bayes (NB) and some NLP libraries like the Python Data Analysis, Natural Language Toolkit, and Scitkit-learn which were used for classification, regression, and clustering purposes. The author used a dataset obtained from the GitHub repository consisting of 10,550 news articles that cut across 2015 to 2017 news information that was labeled as either fake or true. After performing cleaning, training, and testing their model, the result shows that Count-Vectorizer outperformed others with an accuracy of 89.30% [27].

Also, a system that uses keywords from users to search across different sites like Google News, Feedly, News360, etc. was developed based on ML algorithms such as NB, SVM, and Semantic Investigation methods to check the authenticity of information. The system works by accepting news from users as input and verifying them by searching through popular sites for the authenticity of the news [28]. According to the researcher, the system can accurately provide 93.50% results [28].

The third review of ML-related research for the identification of fake news pointed out here is the research carried out by Waikhom [29], which was based on an ensemble method (AdaBoost). The model was trained with the LIAR dataset extracted from the PolitiFact site, containing URLs from as far back as 2007 to 2016, having categories such as true, mostly true, barely true, pants-on-fire, false, and half-true. The model produced an accuracy of 75.00%.

In another effort, Veda et al. [30] developed a model based on SVM, NB, LR, and RF combined with web scrapping technology to identify fake news. The obtained dataset was split into three distinct parts namely the labeled dataset which consists of data labeled as fake or real, the testing dataset, and the training dataset. The web scrapping keeps updating the model dataset to make it learn more. The result from the model shows that the SVM had the best result for the model with an accuracy of approximately 90%.

Another set of authors developed a model using the LR algorithm trained using the LIARs dataset. The dataset consists of 12,000 short statements [31]. Bag-of-words, N-grams, and Term Frequency-Inverse Document Frequency (TF-IDF) were techniques that were used to capture the frequency of the dataset. The effectiveness of the model was increased using the K-fold cross-validation technique and Grid Search Parameter Optimization to increase the performance of the LR algorithm used in the model which produced an accuracy of 75%-93% [31].

In another research, Yerlekar [32] used the NB classifier to design a system for identifying fake news. According to the author, the system is designed to provide users with insight into the identification of fake news. A dataset of 6,335 rows and 4 columns comprising 3,171 real and 3,164 fake information was cleaned of anomalies remaining 6,060 for the model design. A combination of Count-Vector, TF-IDF, etc. was used by the author to extract the necessary features required for the model implementation. Though the results as indicated by the author show a time-saving and straightforward option for identifying fake news with an accuracy of 80%, the amount of dataset used for the implementation of the model is limited as hoaxing has continued to grow in terms.

Another researcher used NLP techniques to identify the authenticity of the news spreading on social media [33]. The model was divided into three stages: identification of stances, author credibility verification, and ML-based classification. A fake news dataset, which contains five features and 20,718 entries, was used in the study. Finally, four different algorithms were used to identify fake news using ML-based classification: DT, RF, LR, and SVM algorithms. The experimental results show that the SVM algorithm has an accuracy of 93.15%, a precision of 92.65%, a recall of 95.71%, and an F1-score of 94.15%. It is worth noting that, the developed fake news identification system uses this model as its existing system [33].

Hossain et al. [34] focused their research on Bangla, combining several ML and Deep Learning (DL) algorithms. Trained using a dataset of 57.000 Bangladesh news articles obtained from the Bangladesh news portal, filtered using Bangladesh Language Took-Kit (BLTK) library, and a combination of TF-IDF, Word2Vec, Fast-Text as well as GloVe model for feature extraction purposes. The research yielded a 96% accuracy with the Bidirectional Long-Short-Term Memory (BLSTM) model.

Finally, Ali et al. [35] developed a fake news identification model based on a sequential ensemble of DL techniques. The author used LIAR and ISOT datasets but trained the model separately using the two datasets. The LIAR dataset comprises 12,825 news content with 8 subjects while the ISOT dataset comprises 44,898 news articles. Although the model performed well with an accuracy of 100%, the model was trained with a dataset that contains only short sentences.

The next section discusses the methods, procedures, and algorithms used for the development of the fake news identification model.

3. METHODOLOGY

3.1 Proposed Framework

The research presents a unique approach by employing Ensemble learning, which harnesses the strengths of multiple algorithms to mitigate bias arising from their weaknesses. The model framework encompasses various algorithms such as SVM, LR, RF, DT, and GB. These algorithms were chosen based on their relevant properties and their previous utilization in fake news identification tasks. Additionally, the algorithms were trained using a benchmark dataset. The established model aims to evaluate the textual content of a dataset, comprising both lengthy and concise sentences that include author names, information content, and titles. The objective is to achieve a robust and high-performance model for identifying fake news.

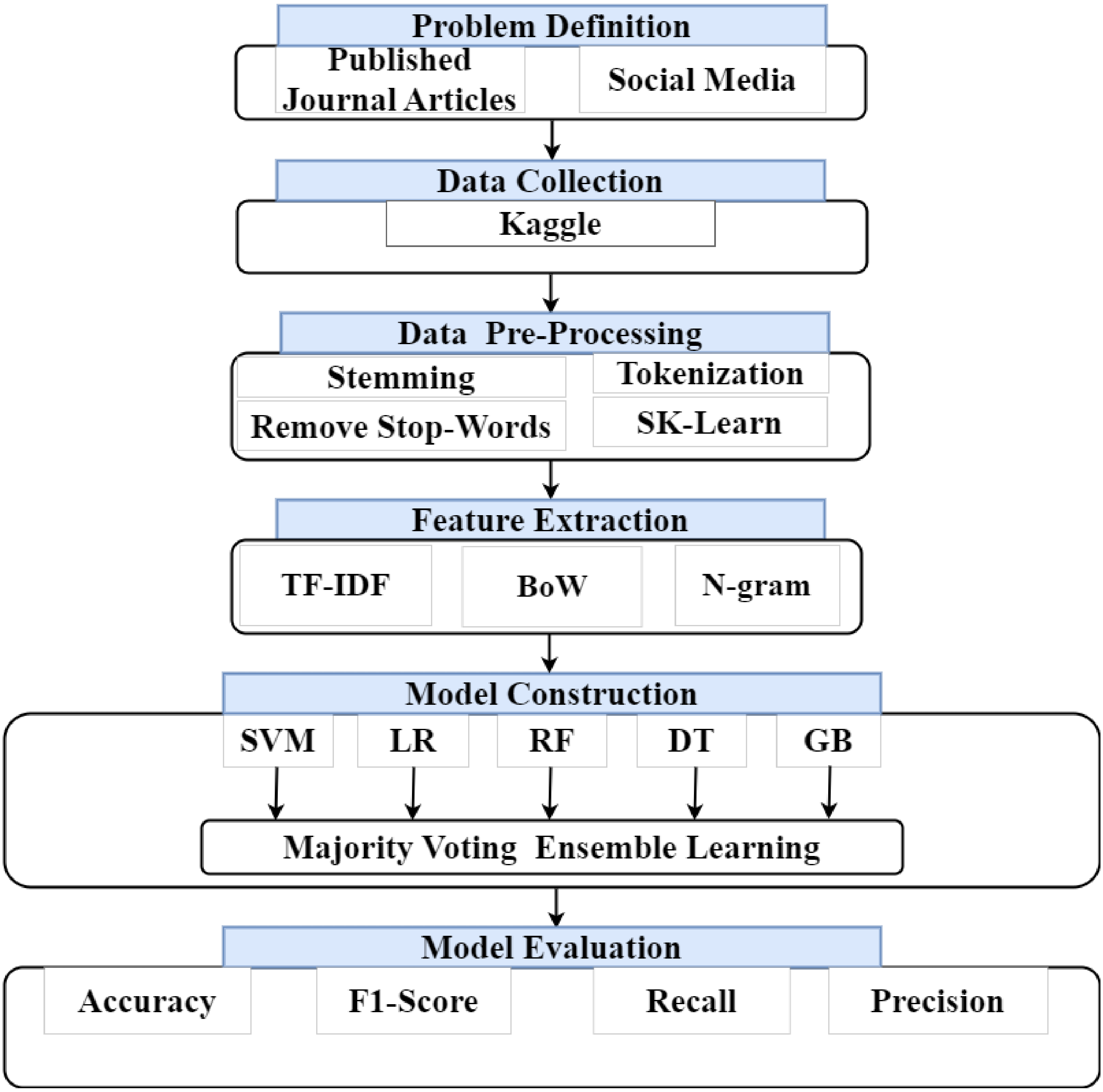

Fig. 1 Shows the Ensemble Learning Framework for Fake News Identification

3.2 Dataset Used

The dataset used in this model was sourced from Kaggle, a publicly accessible repository for datasets. This dataset comprises 20,800 instances of English language text and includes five features: id, authors, title, text, and labels. The labels are represented as 1 for false and 0 for true, indicating whether the instances are classified as false or true, respectively. A pictorial view of the dataset and a short explanation of each feature of the dataset is provided below:

	id	title	author	text	label
0	0	House Dem Aide: We Didn't Even See Comey's Let...	Darrell Lucus	House Dem Aide: We Didn't Even See Comey's Let...	1
1	1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	Daniel J. Flynn	Ever get the feeling your life circles the rou...	0
2	2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired October 29, ...	1
3	3	15 Civilians Killed In Single US Airstrike Hav...	Jessica Purkiss	Videos 15 Civilians Killed In Single US Airstr...	1
4	4	Iranian woman jailed for fictional unpublished...	Howard Portnoy	Print \nAn Iranian woman has been sentenced to...	1

Fig. 2 Shows the Pictorial View of the Dataset

Table 1. **Shows the description of the features of the dataset**

Column	Description
Id	This is a unique identifier assigned to each news article.
Title	The title of the news article.
Text	The news text.
Author	The news originator.
Label	The tag of each news article.

3.3 Data Preprocessing

To achieve optimal results, it is crucial to preprocess and clean the obtained dataset. Anomalies within the dataset can adversely affect the performance of the algorithms during training. With this understanding, preprocessing was performed on the dataset obtained from the Kaggle domain using Python libraries such as Numpy, Pandas, and Scikit-learn. The preprocessing phase involved several steps.

First of all, all the empty cells that are within the dataset were removed, to make sure that all the instances in the dataset had complete information. The next step was to remove all the English Language punctuation marks in other to enhance the consistency of the text. Subsequently, the dataset was also divided into the training and testing datasets, using a ratio of 80:20. This specifically implies that 80% of the dataset was utilized for the training of the models, while 20% which is the remaining portion of the dataset was used for the testing of the models.

Furthermore, this phase of the model construction stage also includes the elimination of stop words and white spaces from the dataset, thereby reducing unnecessary noise from the dataset. In addition to the above explanation, other preprocessing techniques were also carried out, tokenization was also carried out to further enhance the quality of the dataset, to ensure that any other anomalies remaining within the dataset were properly addressed.

3.4 Feature Extraction

Feature Extraction is another important phase in the stages of developing a fake news identification model. This stage is aimed at reducing the presence of instances that were repeated or are not important within the dataset. This stage is essential to enhance the dataset quality, and the accuracy of the established model, reduce the computational time of the model, and finally foster a deeper understanding of the underlying patterns.

In other to accomplish the task of extracting features from the dataset, a Python stemming technique was utilized to stem words, reducing the words to their root or base form. The stemming process helps to normalize the text and remove the various variations of words that are the same, thereby reducing redundancy and improving the dataset coherence.

Additionally, TF-IDF, Bag-of-Words (BoW), and N-gram methods were assessed before selecting the best (TF-IDF) to extract relevant features that are contained in the dataset. The TF-IDF method works by assigning weights to all the terms in the dataset based on the frequency in a particular document and the inverse frequency across the whole dataset. The advantage of the feature extraction phase is the identification of relevant terms within the dataset and downplay the relevance of common words, thereby refining the content of the dataset and capturing the important information needed for the model. A quick look at the three ML feature extraction techniques assessed during the feature extraction phase of this research:

Term Frequency-Inverse Document Frequency: This approach focuses on evaluating the significance of terms within the dataset. It assigns higher importance to rare words and lower importance to common words. The method involves tokenizing the data, calculating the inverse and present frequency of each term, and subsequently encoding new records based on this analysis. By giving priority to rare words, the method aims to capture unique and meaningful information while downplaying the impact of commonly occurring terms. This process helps in extracting valuable features from the dataset and encoding new records with the derived insights.

Bag of Words: Written in short as BoW is also a technique used in ML to extract features from datasets in other to enhance the performance of a given model. This research seeks to enhance the performance of the model for fake news detection. This technique works by recording the number of occurrences of each bag that is created for each instance type or word disregarding the orders of the clusters [36]. This is a computational and conceptual method that makes the feature extraction process simple.

N-gram: N-gram is one of the feature extraction methods used in this research to extract relevant features from the dataset for training the model for the fake news detection model. It is an easy method that is also reliable for text classification[37]. The main idea behind how this model works is by sequencing the letters in a dataset sequentially, which is formed by determining the value of n. It is also a good method because it works regardless of the language of the dataset.

3.5 Model Construction

During this phase of model construction, the selected algorithms are trained using the training dataset. Choosing the appropriate algorithms can be challenging, but for this research, the selection process was informed by an extensive literature review conducted during the problem definition stage. Selecting the classifiers used for the development of this model was influenced by the prevalence of those algorithms in the already existing fake news identification models and the effectiveness of the learning approach of the classifiers.

The algorithms selected for the development of this model include DT, RF, LR, SVM, and GB. Each of the algorithms was independently trained using the same dataset. However, after training them separately they were then combined using the ensemble learning technique. Ensemble learning works by allowing the entire model to leverage the strength of the individual algorithm, enhancing the general performance and accuracy by aggregating the predictions or leveraging their balancing abilities. These algorithms and the ensemble learning technique are discussed below:

Decision Tree Classifier: The described algorithm is a widely employed supervised learning technique that adopts a tree-like structure. It is particularly popular in organizing tasks due to its ability to make decisions and identify patterns based on the given input. The algorithm operates by iteratively splitting the input data using different features, eventually constructing a structured flowchart-like representation that captures these divisions.

This algorithm has found extensive application in various research domains, including ML, statistics, pattern recognition, disease analysis, fake news identification, and more. Its versatility and adaptability make it a valuable tool for diverse fields seeking to leverage data analysis and decision-making processes [38].

Random Forest Classifier: This is used for both classification and regression ML problems [39]. This classifier stands out by employing a unique approach to classification and training. It operates by harnessing the power of randomized data subsets to deliver accurate prediction results. With its versatility, this classifier can effectively tackle a wide range of classification problems while mitigating the risk of overfitting.

What sets this classifier apart is its utilization of multiple DTs. By leveraging the knowledge and insights from these trees, it achieves exceptional prediction performance. The ensemble nature of this approach allows for the aggregation of individual DT predictions, resulting in a robust and reliable overall prediction outcome. In doing so, this classifier offers a distinctive and effective solution to classification challenges.

Logistic Regression Classifier: This classifier specializes in binary classification tasks, focusing on studying the connection between input features and binary outcomes. Its primary objective is to determine the probability of a positive result. In the context of fake news identification, this classifier carefully analyzes the dataset and adjusts its parameters to generate accurate predictions and probabilistic outcomes.

Due to its ability to handle statistical problems effectively, this classifier is widely recognized as a classification ML technique. By leveraging its classification capabilities, it provides valuable insights into the likelihood of a given outcome, making it a valuable tool for addressing various statistical challenges [40].

Support Vector Machine: The described classifier is a versatile supervised ML algorithm that is widely utilized. Its functionality revolves around mapping different points in the dataset to create a high-dimensional feature space. By doing so, the algorithm aims to enable the clear separation of various classes within the dataset using a hyperplane.

The primary goal of this classifier is to establish a decision boundary that effectively discriminates between different classes in the dataset. Mapping the data points into a higher-dimensional space enhances the ability to find hyperplanes that can accurately separate the classes. This approach allows for the classification of new data points based on their relative position with respect to the established hyperplanes [41]. Due to its effectiveness in handling complex classification tasks, this classifier is a popular choice in ML applications. Its ability to create a high-dimensional feature space and identify discriminative hyperplanes makes it a valuable tool for a wide range of classification problems.

Gradient Boost Classifier: This is a powerful ML method commonly employed to tackle both regression and classification problems. It is an ensemble approach that combines the performance of multiple weak DTs to create a more robust and accurate predictive model.

Similar to other ensemble learning techniques, this method leverages the collective strength of the individual weak learners. By combining their predictions, it boosts the overall predictive power of the model. However, what sets this technique apart is its ability to optimize an arbitrary differentiable loss function, which allows for greater flexibility and adaptability to various problem domains.

This ensemble method excels at generalizing the performance of the weak learners and producing a stronger predictive model. By optimizing the chosen loss function, it can effectively handle diverse regression and classification tasks. This makes it a valuable tool for building accurate and robust predictive models in ML applications [42].

Ensemble Learning (Majority Voting Approach): The method used for the identification of fake news involved combining classifiers to take advantage of their strengths and weaknesses. This approach aimed to improve the model's reliability and reduce any biases that might occur when relying on one classifier.

To achieve the results the model was trained using a dataset. This decision enabled the developed model to surpass existing models focused on identifying fake news. By leveraging the insights gained from research and utilizing a used dataset the model was trained to effectively identify fake news, with superior accuracy. The approach employed for the news identification model involved incorporating learning techniques and using a benchmark dataset. This ensured enhanced resilience, minimized biases, and outperformed models, in the field.

3.6 Model Evaluation

When evaluating a developed model, many experts in the field of machine learning find the confusion matrix to be a straightforward approach. Model evaluation is a step in creating a reliable model. This stage focuses on determining whether the proposed model accomplishes its goals and improves upon existing models. The confusion matrix provides insights into how the model performs by showing a comprehensive overview of predicted and actual classes. It allows for calculating performance metrics like accuracy, precision, recall, and F1-score, which helps to give an understanding of how effectively the model classifies instances. By using the confusion matrix during evaluation the research quantitatively assesses how well the proposed model performs compared to established benchmarks and existing models. This helps researchers determine if the developed model achieves its goals successfully and represents an improvement, over other models.

Accuracy: The evaluation of classifiers often depends mostly on the utilization of an accuracy assessment method, which works based on the measurement of the percentage of instances that are correctly classified. This is a commonly used approach when assessing the effectiveness of classifiers in accurately assigning instances to the appropriate classes [43]. As stated by Eke et al. [44], accuracy is a metric used to quantify the percentage of correctly predicted outcomes in relation to the total number of predicted classes. It serves as a measure to evaluate the model's ability to accurately predict and allot instances to their respective classes. Mathematically, accuracy is defined as follows:

$$Accuracy = \frac{(TN+TP)}{(TP+FP+TN+\text{FN`})} \quad (1)$$

TN, TP, FP, and FN' are acronyms used in classification evaluation. TN represents instances that are correctly classified as negative, TP represents instances that are correctly classified as positive, FP denotes instances that are incorrectly categorized as positive, and FN' signifies instances that are incorrectly categorized as negative. These expressions represent an important role in analyzing the performance of classification models and calculating evaluation metrics.

Precision metrics point to the model's reliability by calculating the accuracy of positive predictions, reflecting the proportion of true positives amid the predicted positive instances [45]. Mathematically,

$$Precision = \frac{(TP)}{(FP+TP\text{`})} \quad (2)$$

TP stands for True Positive, representing instances correctly classified as positive by the model. On the other hand, FP stands for False Positive, indicating instances falsely classified as positive when they are negative.

Recall assessment metrics gauge the model's accurate prediction of true positive instances, showcasing its ability to identify actual positive cases. It measures the proportion of correctly identified positive instances out of all actual positive instances [45]. Mathematically,

$$Recall = \frac{(TP)}{(TP+\text{FN`})} \quad (3)$$

TP stands for True Positive, denoting instances correctly identified as positive by the model, while FN' represents False Negative, indicating instances wrongly classified as negative when they are positive.

F1-Score: The F1-score provides a comprehensive evaluation metric by combining precision and recall assessments [45]. It indicates that as precision increases, recall decreases, and vice versa, establishing a robust association between these evaluation metrics [45], [46]. Mathematically,

$$F1\ Score = 2 * \frac{(Recall * Precision)}{(Recall + Precision)} \quad (4)$$

Experimental Set-up

Experimental Procedure

This section describes the various stages involved in constructing the fake news identification model, which is designed to identify fake news content in text on social media platforms like Facebook and Twitter. The development of the model involved employing state-of-the-art processes, including problem definition, data collection, pre-processing, feature extraction, model construction, and model evaluation stages. The Google Collaborator, an online resource with the features of a Jupyter Notebook for Python programming language, was used for this research work. The dataset used for training and testing purposes was obtained from the publicly available platform Kaggle in a .csv file format. Features from the dataset were extracted using three distinct ML feature extraction techniques. The dataset was also split into separate parts for training and testing, and both splits were utilized to train and test five ML algorithms as well as the ensemble method. The decision to train separate models and combine them into a single ensemble model was made to address biases and enhance robustness. This ensemble model, comprising different types of classifiers working together, contributes to improved model performance. The experiment was conducted on a 64-bit Windows 10 operating system with an Intel Core i5 CPU, and 12GB RAM, operating at a speed of 2.65GHz.

The following section contains the results obtained after the proposed model was developed, discussing the various results obtained and interpreting how the established model outperformed others and the performance of the individual models combined to form the ensemble learning model.

4. RESULTS AND DISCUSSIONS

The Fake News Identification model, developed using the Ensemble learning method, commenced with the training of various ML algorithms on the same dataset intended for the proposed model. Consequently, both the performance of individual ML algorithms and the collective performance of the ensemble method were assessed using a binary classification confusion matrix. This confusion matrix is termed 'binary classification' due to its ability to classify news articles as either fake or true. The resulting matrices for each algorithm, as well as the ensemble's performance, are presented below, accompanied by a comprehensive analysis of the accuracy obtained using the training and testing datasets. This is followed by an evaluation of the performance of the individual models in comparison to the ensemble model. All of these results are presented based on the different feature extraction techniques applied to the dataset. Finally, the research compares the developed ensemble model with other existing works.

4.1 Term Frequency-Inverse Document Frequency Feature Extraction

As pointed out earlier, during the feature extraction phase of developing the fake news detection model, one of the techniques used for extracting relevant features from the dataset is TF-IDF. The results obtained from various machine learning models after applying the TF-IDF feature extraction method to extract relevant features from the dataset are presented below:

Support Vector Machine: In Fig. 3, the utilization of the SVM algorithm to identify fake news is depicted through a confusion matrix, demonstrating its efficacy. The outcomes reveal that the SVM model proficiently distinguished between fabricated and genuine news articles, achieving a remarkable accuracy of 96.97% with the testing data. Furthermore, the model yielded consistent scores of 97% for Precision, Recall, and F1-score.

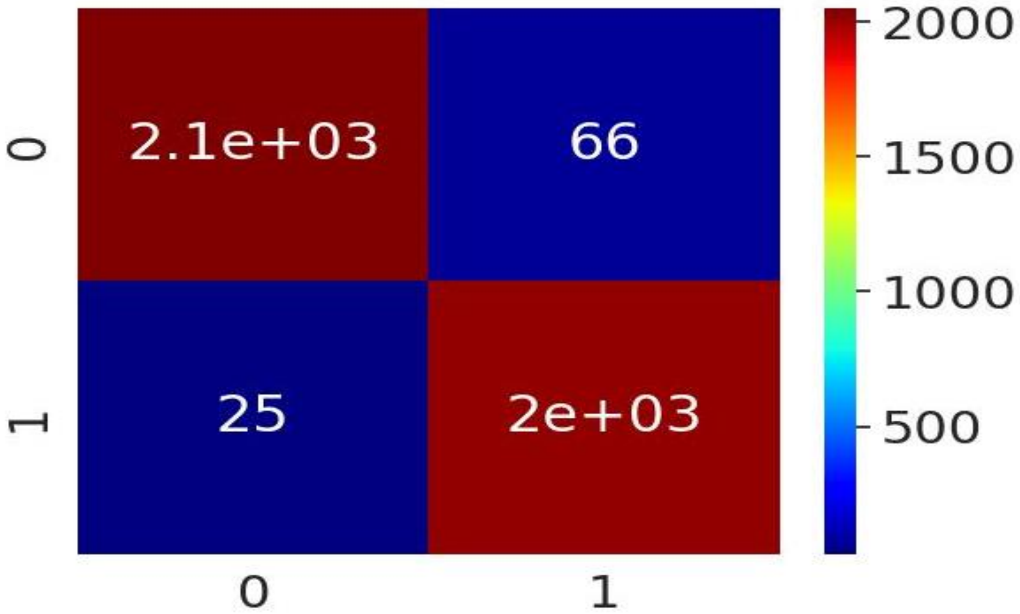

Fig. 3 Shows the Confusion Matrix for the SVM Algorithm Trained on a TF-IDF-extracted dataset

Decision Tree: Similar to the SVM approach, the DT algorithm underwent a comparable process. To identify fake news, the DT algorithm was trained using the training portion of the split dataset. The outcomes demonstrated that the model achieved a remarkable accuracy of 96.41% when assessed with the testing dataset. Moreover, it attained a precision of 96%, recall of 96%, and an F1-score of 96% when evaluated with the same testing dataset. Fig. 4 exhibits the confusion matrix for the DT algorithm.

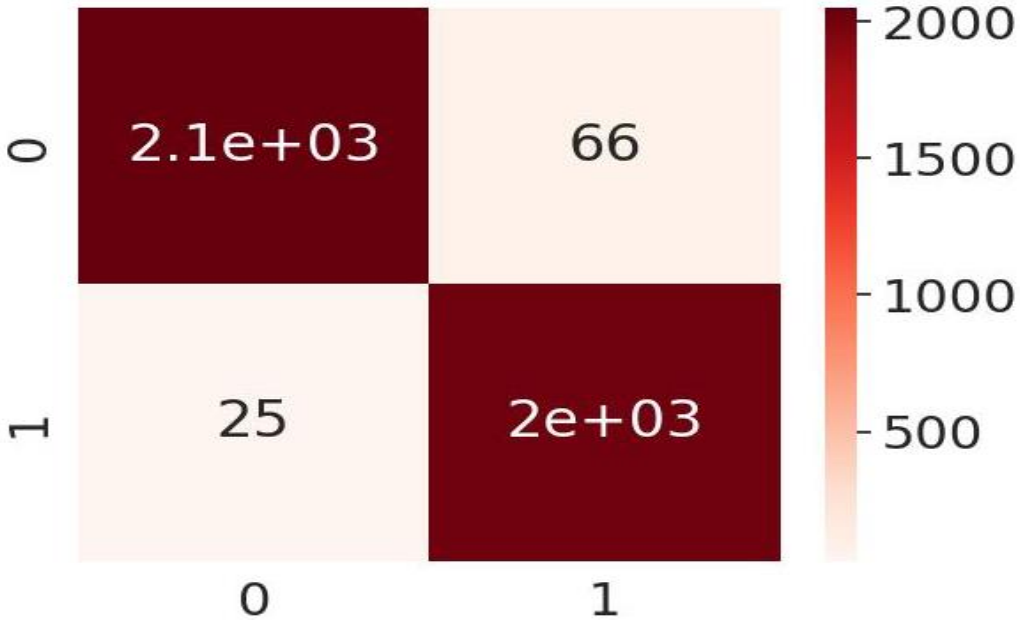

Fig. 4 Shows the Confusion Matrix for the DT Algorithm Trained on a TF-IDF-extracted dataset

Logistic Regression: The confusion matrix for the LR algorithm is depicted in Fig. 5. This matrix is derived from training the LR algorithm using the same fake news dataset utilized in the SVM approach. The model's performance on the testing dataset yielded an accuracy of 95.43% showing it learned well from the training dataset. Similarly, on the testing dataset, the algorithm achieved precision, recall, and F1-Score of 95%. The matrix is presented below:

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.41, September 2023 – www.ijais.org

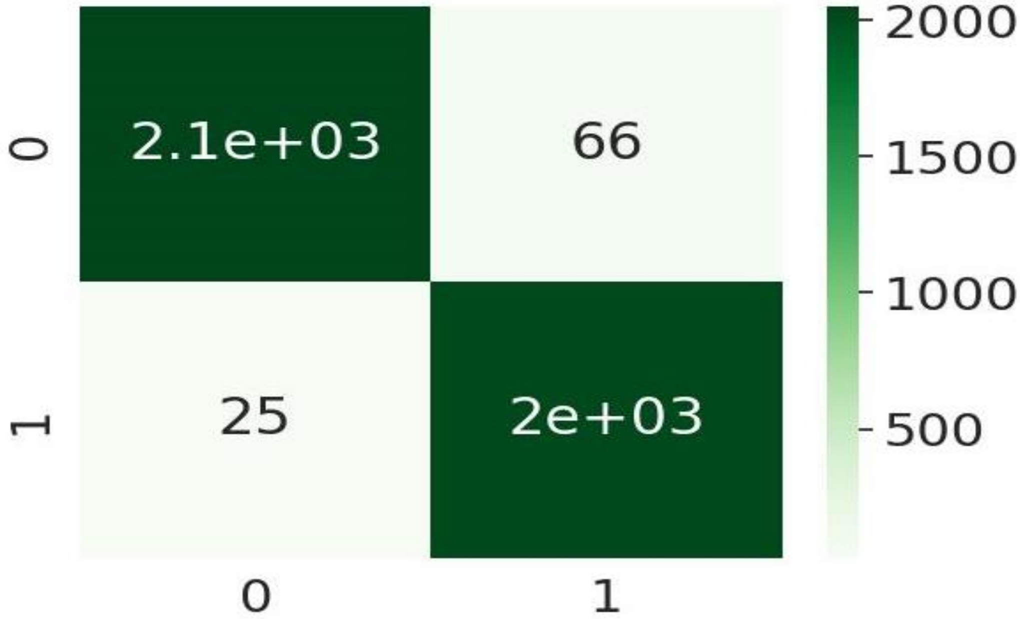

Fig. 5 Shows the Confusion Matrix for the LR Algorithm Trained on a TF-IDF-extracted dataset

Random Forest: Both the LR and RF algorithms employed the same dataset for identifying fabricated articles. The RF model effectively identified fake news using this dataset. When evaluated with a distinct dataset, it consistently achieved a commendable accuracy of 92.50%. Moreover, the model demonstrated impressive precision, recall, and F1-score metrics, all measuring 93%. Fig. 6 visually represents the performance of the confusion matrix for the RF model.

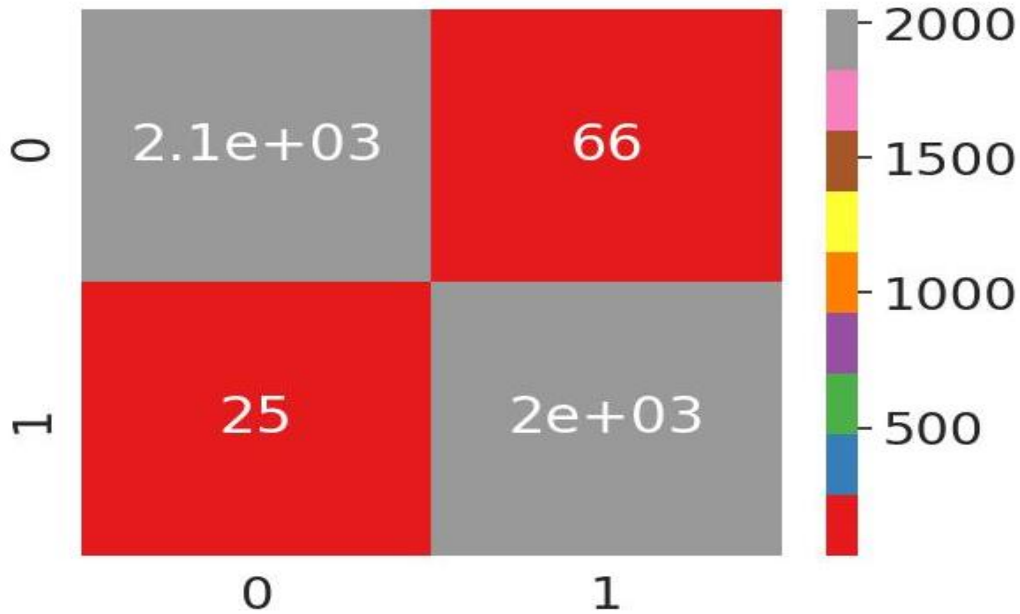

Fig. 6 Shows the Confusion Matrix for the RF Algorithm Trained on a TF-IDF-extracted dataset

Gradient Boost: The successful identification of fabricated news within the training subset of the dataset by the GB algorithm is supported by the depiction of the confusion matrix. However, when evaluated using the test dataset, the GB algorithm demonstrated an accuracy of 96.41%. Additionally, the algorithm exhibited an average precision, F1-score, and recall of 96%. Fig. 7 provides a visual representation of the aforementioned results.

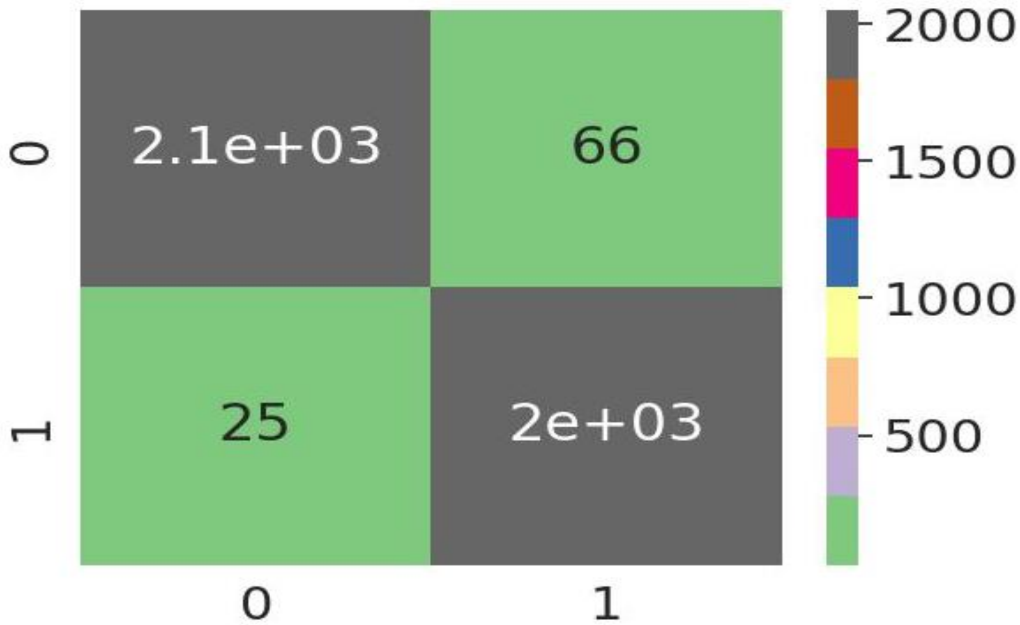

Fig. 7 Shows the Confusion Matrix for the GB Algorithm Trained on a TF-IDF-extracted dataset

Ensemble Model: In conclusion, the ensemble model assumes a pivotal role as the cornerstone of this research, aimed at constructing a resilient model that mitigates bias and ensures consistency. This model was deployed on the identical dataset procured. Fig. 8 showcases the confusion matrix for the ensemble model, emphasizing its proficiency in precisely categorizing fabricated news instances within the training dataset extracted using the TF-IDF feature extraction technique. Upon evaluation using the testing dataset, the ensemble model exhibited a substantial level of 97.86 accuracy, while maintaining an average score of 98% for F1-score, Precision, and Recall.

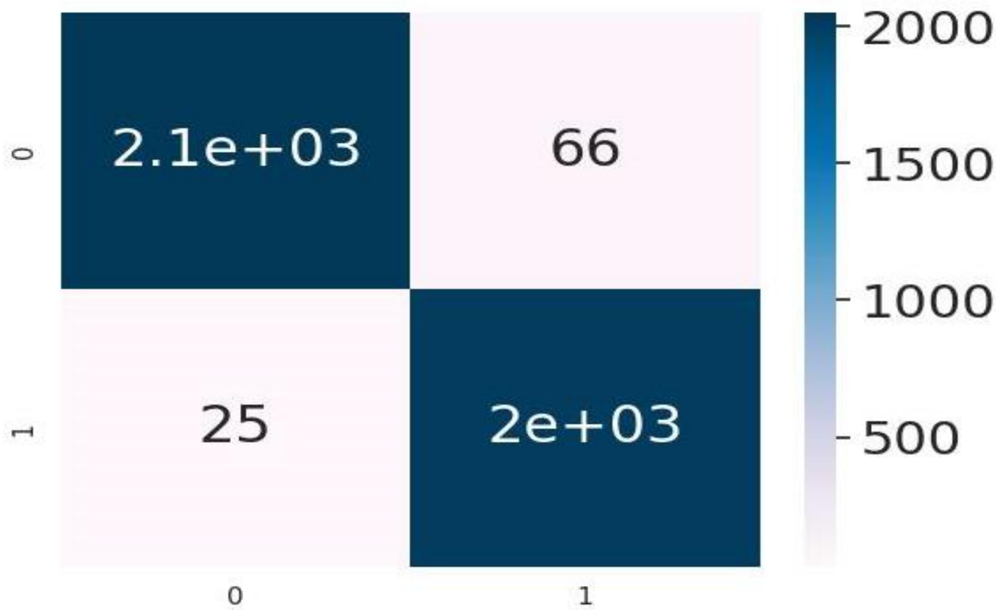

Fig. 8 Shows the Confusion Matrix for the Ensemble Model Trained on a TF-IDF-extracted dataset

Comparative Analysis of the Algorithms Trained on a TF-IDF Extracted Dataset: The task of selecting an optimal ML model with superior performance trained on a specific dataset extracted properly using a good feature extraction technique can pose challenges. This decision relies on the algorithm's ability to effectively learn from the dataset's features. Consequently, prior research has employed multiple ML algorithms and feature extraction techniques, conducting thorough performance evaluations to identify the most effective model. In this study, the choice of ML algorithms was guided by the literature review, emphasizing the importance of combining the strengths of selected algorithms into a unified model called the Ensemble Model. Hence, the employed algorithms for developing the fake news identification model included SVM, LR, DT, RF, and GB. In this case, the TF-IDF feature extraction technique was assessed, and the results are presented below.

To determine the best-performing ML model among these five algorithms and the Ensemble Model using the TF-IDF, a comparative analysis is essential. The performance results of these ML techniques are presented, providing a summary of their predictive performance metrics such as Precision, Recall,

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.41, September 2023 – www.ijais.org

F1-score, and Accuracy. These performance evaluation metrics are elaborated in Table 2, offering a comprehensive overview of the algorithms' effectiveness in identifying fake news.

Table 2. Shows the Comparative examination of the six(6) ML models trained on a TF-IDF extracted dataset

Machine Learning Models	Assessment in Percentages			
	Accuracy	**Precision**	**Recall**	**F1-score**
SVM	96.97%	97.00%	97.00%	97.00%
LR	95.43%	95.00%	95.00%	95.00%
RF	92.50%	93.00%	93.00%	93.00%
DT	96.41%	96.00%	96.00%	96.00%
GB	96.41%	96.00%	96.00%	96.00%
Ensemble Method	**97.86%**	**98.00%**	**98.00%**	**98.00%**

The analysis depicted in Table 2 demonstrates that each of the algorithms effectively captured the features present in the dataset, enabling precise identification of fake news based on the dataset attributes. The results indicate that the Ensemble Model outperformed the other ML models, showcasing superior performance with accuracy rates of 97.86%.

4.2 Bag-of-Words Feature Extraction

Because this research aims to develop an efficient and robust fake news detection model, the BoW feature extraction technique was also employed to extract features from the dataset before training various models. The results obtained from the machine learning models, after applying the BoW feature extraction method to extract relevant features from the dataset, are presented below:

Logistic Regression: Figure 9 displays the confusion matrix for the LR algorithm, which was trained using a dataset extracted using the BoW feature extraction technique. This matrix originates from the LR algorithm's training on the identical fake news dataset prepared using BoW feature extraction. The LR model exhibited outstanding performance on the testing dataset, achieving an impressive accuracy of 96.63%, indicating its strong learning from the training dataset. Furthermore, on the testing dataset, the LR algorithm demonstrated remarkable precision, recall, and F1-Score, all at 97%. The confusion matrix is presented below for reference:

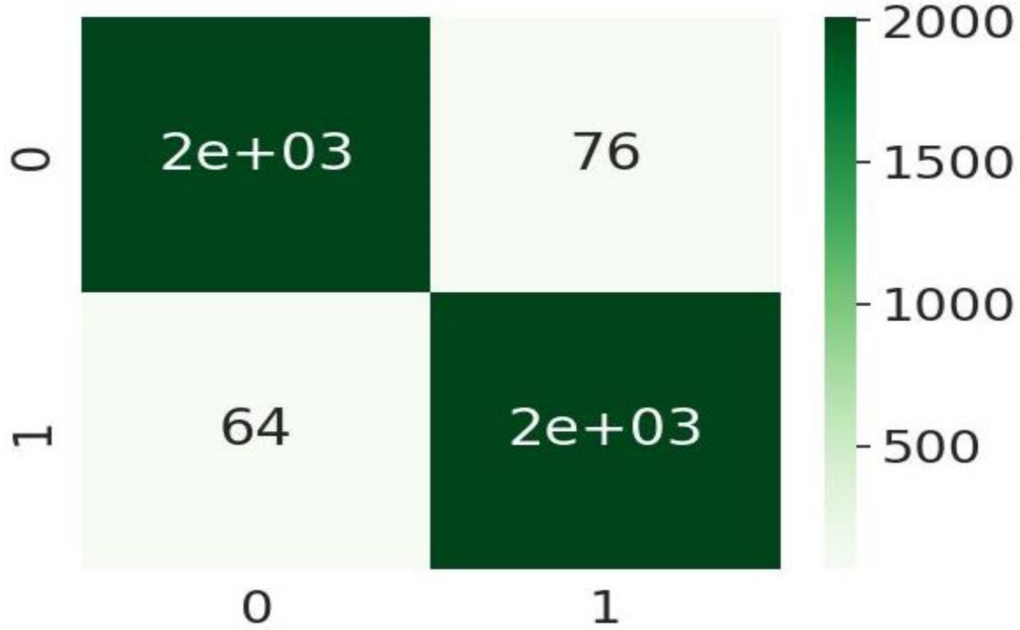

Fig. 9 Shows the Confusion Matrix for LR Algorithm Trained on a BoW-extracted dataset

Decision Tree: DT was the next model trained using the features from the dataset, which were also extracted using the BoW feature extraction technique. The model learned effectively from the training dataset, and when tested with the validation dataset, it achieved an accuracy of 96.37%, while maintaining 96% for the precision, recall, and the F1-score. To further report the performance of the DT model, Fig. 10 below shows the confusion matrix obtained from the model.

International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.41, September 2023 – www.ijais.org

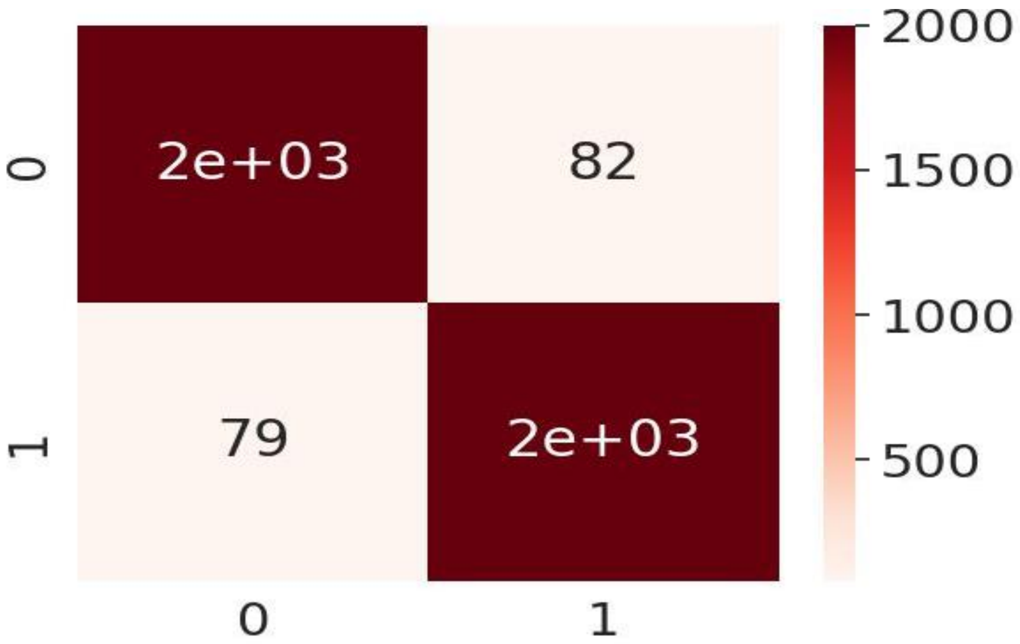

Fig. 10 Shows the Confusion Matrix for a DT Algorithm Trained on a BoW-extracted dataset

Support Vector Machine: Fig. 11 illustrates the application of the SVM algorithm for fake news detection using a confusion matrix, following the model's training with a dataset extracted via the BoW method. The results indicate that the SVM model, trained on the BoW-extracted dataset, effectively differentiated between fabricated and genuine news articles, achieving an accuracy of 96.12%. Moreover, the model consistently achieved precision, recall, and F1-score values of 96%.

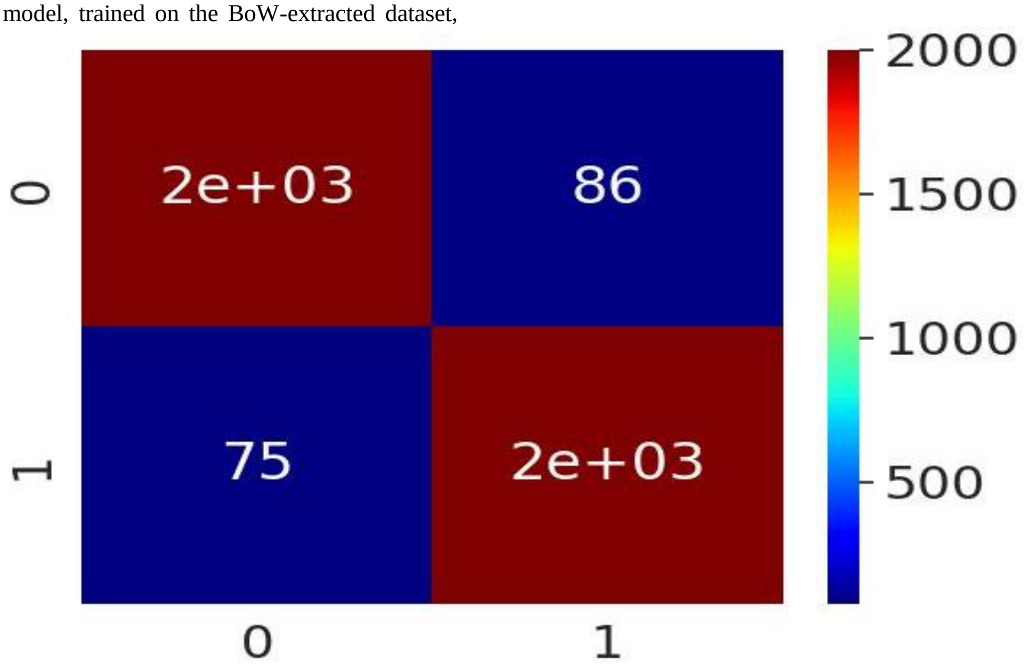

Fig. 11 Shows the Confusion Matrix for an SVM Algorithm Trained on a BoW-extracted dataset

Random Forest: The RF algorithms utilized the same dataset extracted using the BoW feature extraction technique. The RF model achieved an accuracy of 92.78% when evaluated with a separate dataset. Additionally, the model exhibited precision, recall, and F1-score metrics, all at 93%. Fig. 12 provides a visual representation of the confusion matrix for the RF model's performance.

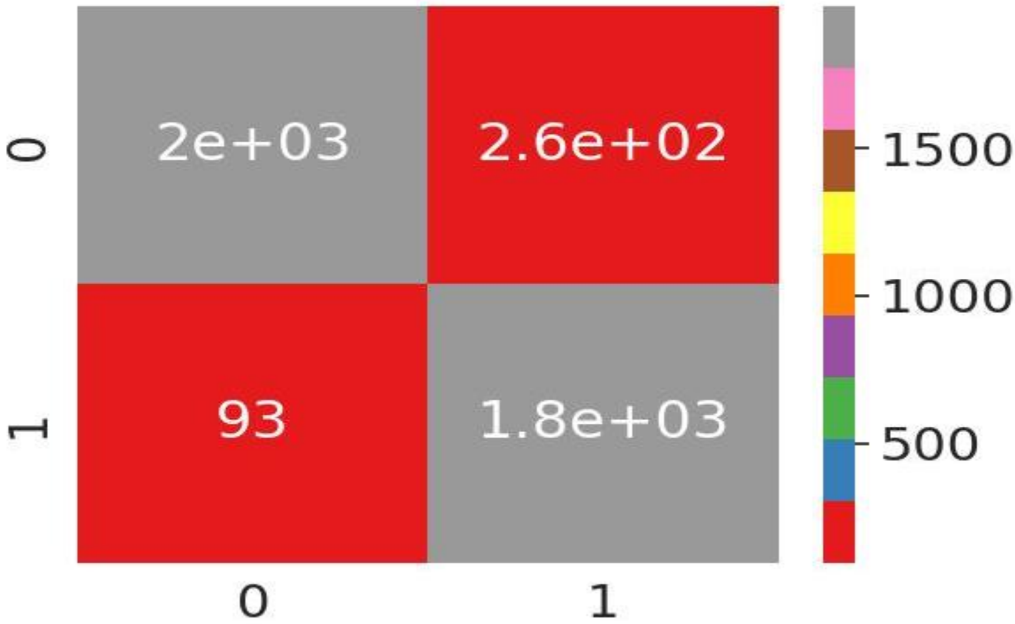

Fig. 12 Shows the Confusion Matrix for an RF Algorithm Trained on a BoW-extracted dataset

Gradient Boost: The GB algorithm was trained with the same dataset obtained using the BoW feature extraction method. When assessed using the test dataset, the GB algorithm achieved a noteworthy accuracy of 96.58%. Furthermore, it consistently maintained an average precision, F1-score, and recall of 97%. To view these results visually, please refer to Fig. 13.

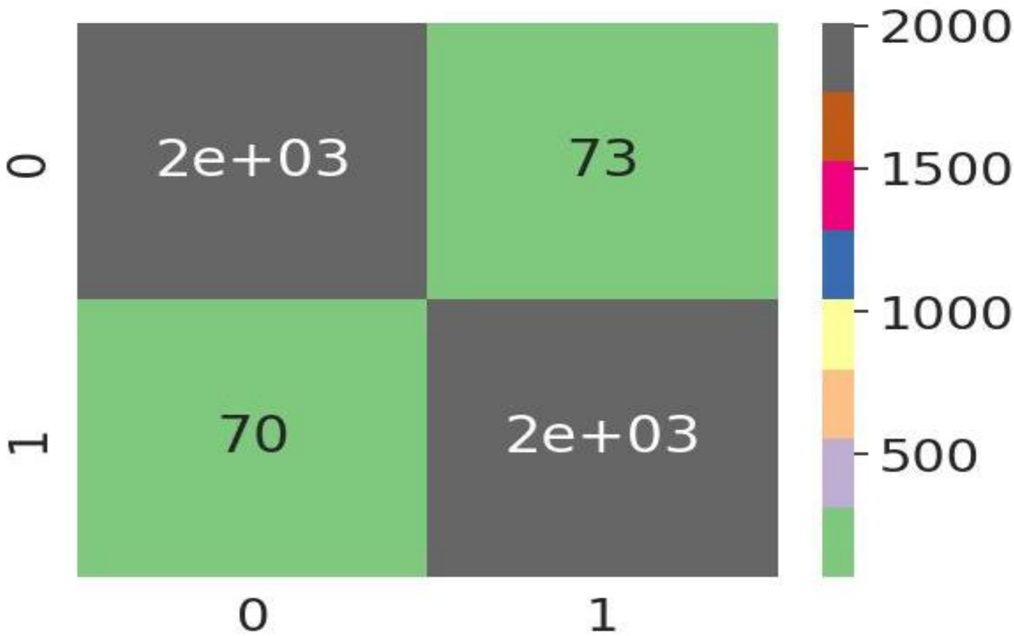

Fig. 13 Shows the Confusion Matrix for a GB Algorithm Trained on a BoW-extracted dataset

Ensemble Model: After training individual models using BoW-extracted features, the study followed a similar procedure to assess the performance of the ensemble model with the same dataset. Here, the study integrated various individual models trained on the BoW-extracted dataset. Fig. 14 shows the confusion matrix for this ensemble model, achieving an impressive accuracy rate of 97.45%. Furthermore, the ensemble model consistently maintained an average score of 97% for F1-score, Precision, and Recall.

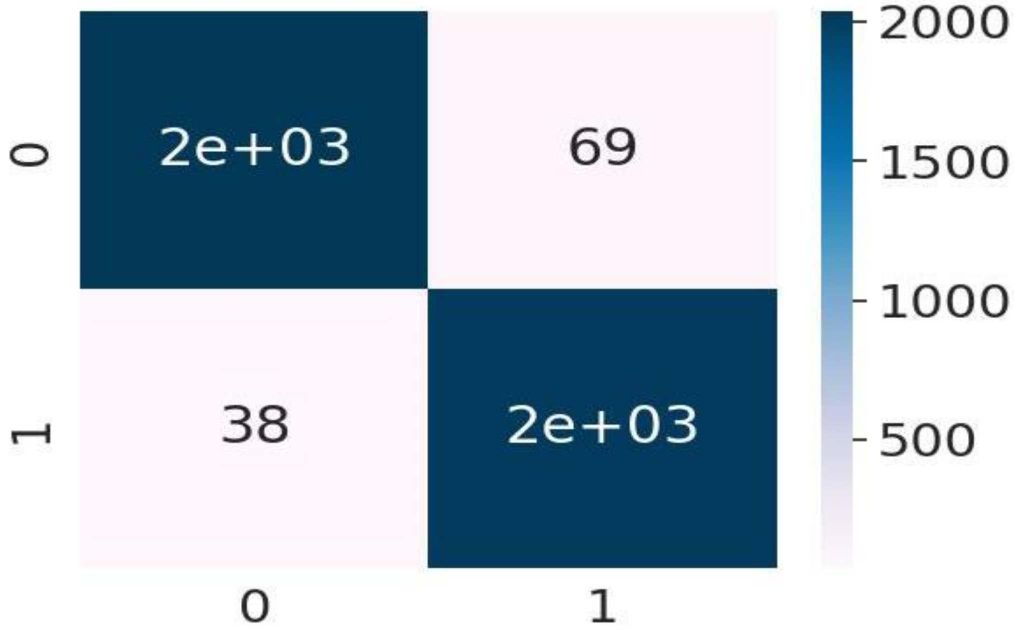

Fig. 14 Shows the Confusion Matrix for an Ensemble Model Algorithm Trained on a BoW-extracted dataset

Comparative Analysis of the Algorithms Trained on the BoW-Extracted Dataset: Similar to the TF-IDF feature extraction section, this is a comparative analysis to identify the best-performing model trained on the BoW-extracted dataset. Several ML algorithms were trained using the BoW technique, and an emphasis was placed on combining the strengths of these selected algorithms into an ensemble model. These algorithms align with those trained using the TF-IDF extracted dataset.

To ascertain the top-performing ML model among these five algorithms and the Ensemble Model, a comparative analysis is essential. The results of these ML techniques' performance are presented, summarizing their predictive metrics, including Precision, Recall, F1-score, and Accuracy. These performance metrics are comprehensively detailed in Table 3, offering an overview of the algorithms' performance.

Table 3. Shows the Comparative examination of the six(6) ML models trained on a BoW-extracted dataset

Machine Learning Models	Assessment in Percentages			
	Accuracy	Precision	Recall	F1-score
LR	96.63%	97.00%	97.00%	97.00%
DT	96.37%	96.00%	96.00%	96.00%
SVM	96.12%	96.00%	96.00%	96.00%
RF	92.78%	93.00%	93.00%	93.00%
GB	96.58%	97.00%	97.00%	97.00%
Ensemble Method	**97.45%**	**97.00%**	**97.00%**	**97.00%**

The analysis depicted in Table 3 demonstrates that each of the algorithms effectively captured the features extracted using the BoW. The results indicate that the Ensemble Model outperformed the other ML models, showcasing superior performance with accuracy rates of 97.45%.

4.3 N-gram Feature Extraction: Bag-of-Words Feature Extraction

As mentioned earlier in the feature extraction phase of model development, one of the techniques assessed was the N-gram feature extraction method. This method was used to extract features from the dataset to obtain the best results when training the fake news detection model. The following section presents the results obtained from the ML models after training them using the dataset that underwent feature extraction using the N-gram method to capture relevant dataset features:

Logistic Regression: Fig. 15 provides the confusion matrix for the LR algorithm, which was trained using a dataset extracted with the N-gram feature extraction method. This matrix results from the LR algorithm's training on the same fake news dataset extracted using N-gram feature extraction. The LR model established a performance on the testing dataset, attaining an accuracy of 96.63%, with precision, recall, and F1-Score all having 97%.

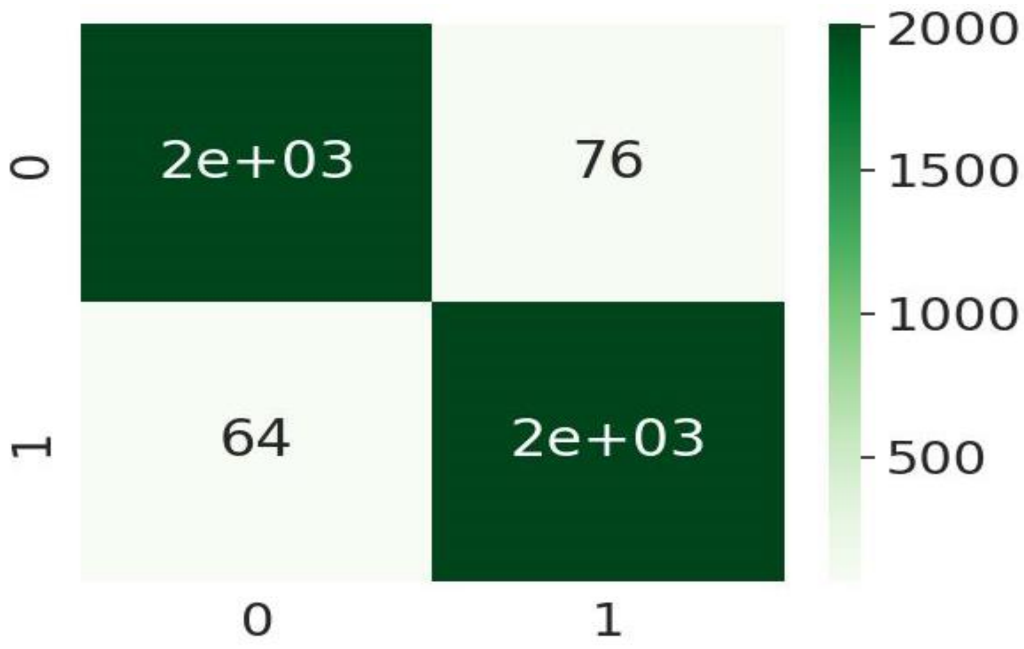

Fig. 15 Shows the Confusion Matrix for LR Algorithm Trained on a N-gram-extracted dataset

Decision Tree: The DT model was also trained with features extracted from the dataset using the N-gram feature extraction method. The model exhibited good capability during training and achieved an accuracy of 96.46% when evaluated with the testing dataset. Notably, both precision and recall maintained 96%, along with the F1-score. Fig. 16 provides the confusion matrix.

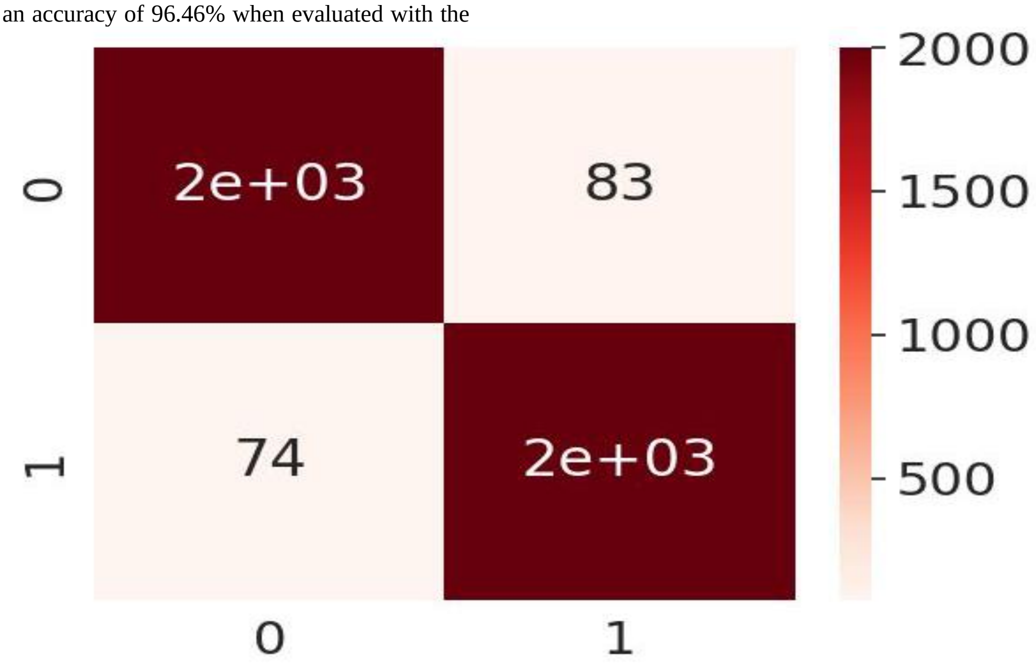

Fig. 16 Shows the Confusion Matrix for a DT Algorithm Trained on a N-gram extracted dataset

Support Vector Machine: Following the training of the SVM model with a dataset extracted using the N-gram method, it achieved an accuracy of 96.12%. Notably, the model consistently maintained precision, recall, and F1-score values at 96%, matching the results when trained with a BoW-extracted dataset. The corresponding confusion matrix for this model is presented below in Fig. 17:

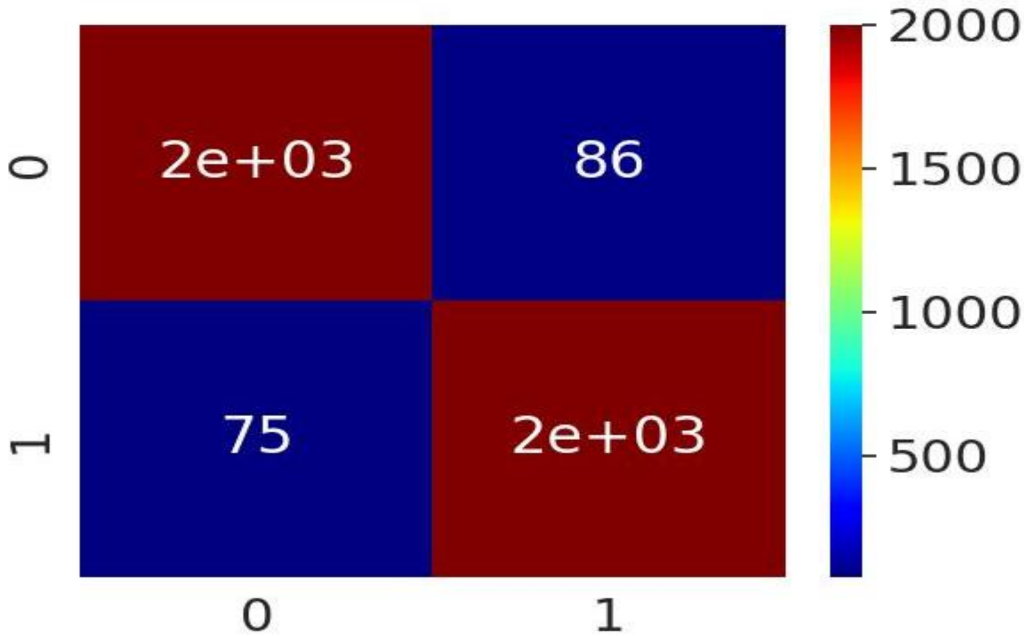

Fig. 17 Shows the Confusion Matrix for an SVM Algorithm Trained on an N-gram extracted dataset

Random Forest: The RF algorithms utilized the same dataset extracted using the N-gram feature extraction technique. The RF model achieved an accuracy of 92.83% when evaluated with a testing dataset. Additionally, the model exhibited precision, recall, and F1-score metrics, all at 93%. Fig. 18 provides a visual representation of the confusion matrix for the RF model's performance.

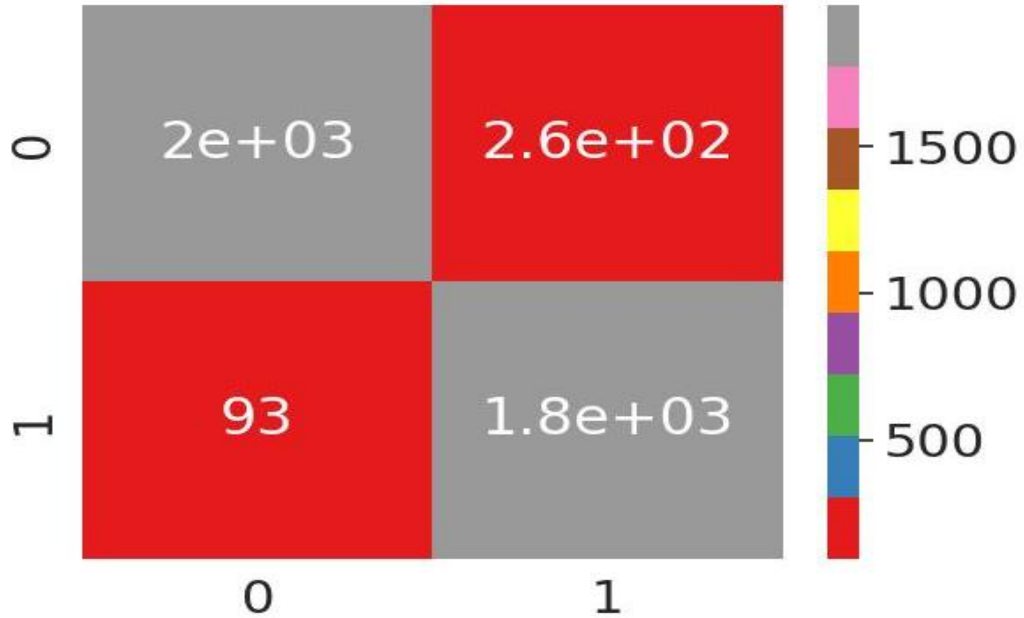

Fig. 18 Shows the Confusion Matrix for an RF Algorithm Trained on an N-gram extracted dataset

Gradient Boost: The GB algorithm was trained using the N-gram extracted dataset. Upon evaluation of the test dataset, the GB algorithm displayed an impressive accuracy of 96.58%. Furthermore, the algorithm consistently achieved an average precision, F1-score, and recall of 97% which is a similar result obtained when trained with a BoW-extracted dataset. To visualize these results, please refer to Fig. 19.

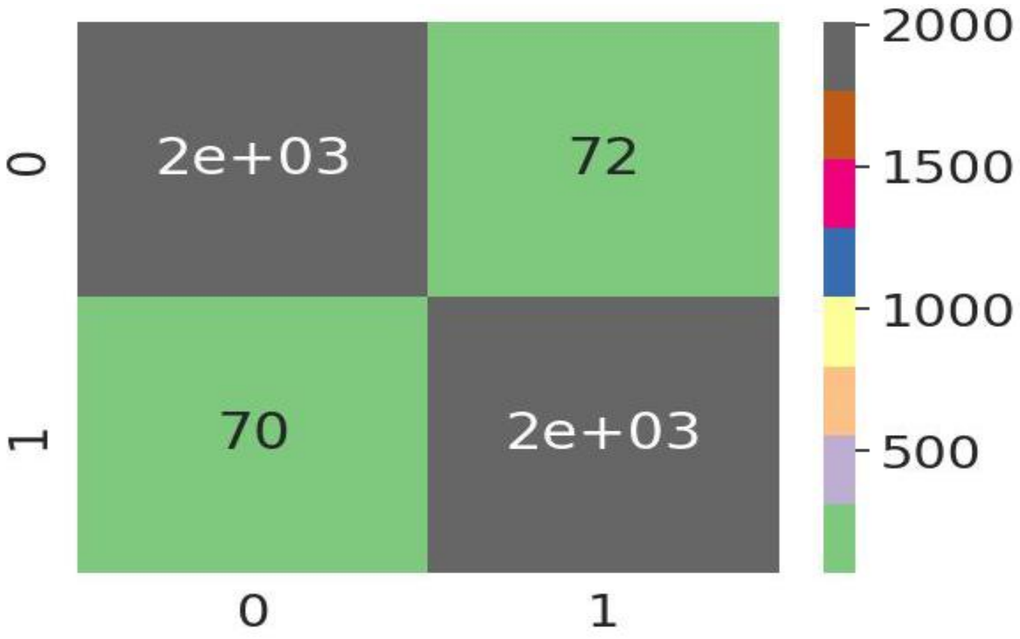

Fig. 19 Shows the Confusion Matrix for a GB Algorithm Trained on a N-gram extracted dataset

Ensemble Model: In conclusion, after reporting on the single models trained using the dataset extracted using the N-gram feature extraction technique, we also trained an ensemble model using the same dataset. The results from this ensemble model show an accuracy of 97.40% while consistently maintaining an average score of 97% for F1-score, Precision, and Recall.

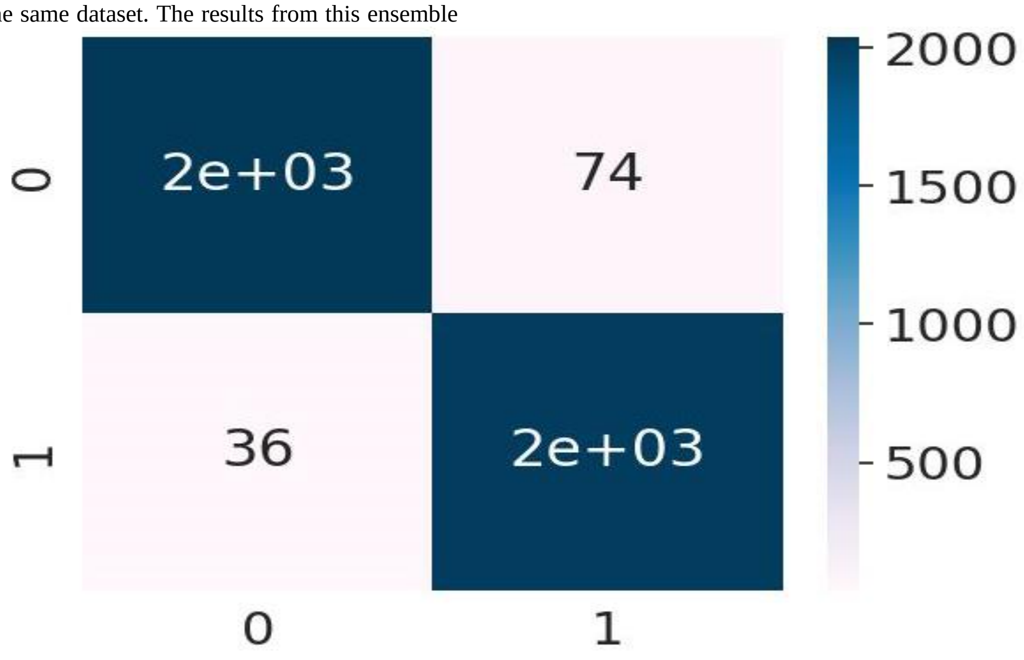

Fig. 20 Shows the Confusion Matrix for an Ensemble Model Algorithm Trained on a N-gram extracted dataset

Comparative Analysis of the Algorithms Trained on N-gram Extracted Dataset: Similar to the previous sections analyzing the results of the BoW and TF-IDF feature extraction techniques, this section aims to present a concise assessment of various models trained using the N-gram extracted dataset. The goal here is to identify the best-performing ML model among these five algorithms and the Ensemble Model trained with the N-gram feature extraction dataset. The section will provide the performance results of these ML techniques, summarizing their predictive performance metrics, including Precision, Recall, F1-score, and Accuracy. A detailed breakdown of these performance evaluation metrics is in Table 4, which offers a comprehensive overview of the algorithms' performance.

Table 4. Shows the Comparative examination of the six(6) ML models trained on a N-gram extracted dataset

Machine Learning Models	Assessment in Percentages			
	Accuracy	Precision	Recall	F1-score
LR	96.63%	97.00%	97.00%	97.00%

DT	96.46%	96.00%	96.00%	96.00%
SVM	96.12%	96.00%	96.00%	96.00%
RF	92.83%	93.00%	93.00%	93.00%
GB	96.58%	97.00%	97.00%	97.00%
Ensemble Method	**97.40%**	**97.00%**	**97.00%**	**97.00%**

The analysis depicted in Table 4 demonstrates that each of the algorithms effectively captured the features extracted using the N-gram. The results indicate that the Ensemble Model outperformed the other ML models, showcasing superior performance with accuracy rates of 97.40%.

4.4 Comprehensive Analysis of the Three (3) Feature Extraction Methods

To provide a comprehensive conclusion and understanding of the research, which applies ML techniques to develop an enhanced fake news detection model on social media, utilizing state-of-the-art ML methodologies and tools, a comparative analysis of the ML feature extraction method is provided. These methods include TF-IDF, Bag of Words (BoW), and N-gram techniques, along with an examination of various trained models. Table 5 provides a detailed comparative analysis of ML models trained using these diverse feature extraction methods.

Table 5 Shows the Comprehensive Analysis of the Three (3) Feature Extraction Method

ML Models	Assessment of Feature Extraction Methods											
	TF-IDF				BoW				N-gram			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
LR	95.43%	95.00%	95.00 %	95.00 %	96.63%	97.00%	97.00 %	97.00 %	96.63%	97.00%	97.00 %	97.00 %
DT	96.41%	96.00%	96.00 %	96.00 %	96.37%	96.00%	96.00 %	96.00 %	96.46%	96.00%	96.00 %	96.00 %
SVM	96.97%	97.00%	97.00 %	97.00 %	96.12%	96.00%	96.00 %	96.00 %	96.12%	96.00%	96.00 %	96.00 %
RF	92.50%	93.00%	93.00 %	93.00 %	92.78%	93.00%	93.00 %	93.00 %	92.83%	93.00%	93.00 %	93.00 %
GB	96.41%	93.00%	93.00 %	93.00	96.58%	97.00%	97.00 %	97.00 %	96.58%	97.00%	97.00 %	97.00 %
E M	**97.86%**	**98.00%**	**98.00 %**	**98.00 %**	**97.45%**	**97.00%**	**97.00 %**	**97.00 %**	**97.40%**	**97.00%**	**97.00 %**	**97.00 %**

As depicted in Table 5 above, the performance of various models based on the specific feature extraction technique employed to extract the dataset used in their training is shown. The overall results point towards the ensemble model emerging as the top performer, boasting the highest accuracy among all models.

The research investigated the most optimal combination, it became evident that the TF-IDF feature extraction technique, in conjunction with the ensemble model, delivers the most impressive results. This combination achieves an accuracy rate of 97.86%, with precision, recall, and F1-score all consistently hitting the 98.00% mark.

It's interesting to note that the BoW and N-gram feature extraction techniques yield results that are quite closely matched. Moreover, the individual models trained using either the BoW or N-gram feature extraction technique exhibit slightly superior performance compared to those trained with data extracted using the TF-IDF feature extraction technique.

4.5 Conclusion

In line with the purpose of this research work which is aimed at developing an effective and robust fake news detection model. After analyzing the results obtained from various models trained with different ML feature extraction techniques, the results have shown that the ensemble model trained with the TF-IDF feature extraction technique outperformed other methods.

Given that, a thorough examination is conducted to compare the effectiveness of this study with five recently established models in identifying false information. The selection of these benchmark studies is based on their relevance to the field of fake news identification and the specific features of the datasets used. Through rigorous experimentation, the newly developed fake news identification model has outperformed the five benchmark studies. A comprehensive overview of the improved model's performance concerning the existing cutting-edge models for identifying fake news is provided in Table 6.

Table 6. Shows the Comparative Analysis of the Improved Model with Existing Model(s)

Authors	Learning Model	Percentage Accuracy
[27]	Naïve Bayes	89.30%
[29]	AdaBoost	75.00%
[47]	Logistic Regression	90.00%
[32]	Naïve Bayes	80.00%
[48]	Legitimacy	96.90%
Developed Model	**Ensemble Model**	**97.86%**

The analysis of Table 6 reveals the remarkable superiority of the developed Fake News Identification model across all evaluation metrics, outperforming all existing models. In the initial benchmark study, the developed model achieved an accuracy of 97.86%, surpassing the existing model based on the NB algorithm, which achieved an accuracy of 89.30% for fake news identification. In contrast, the second baseline model, an ensemble AdaBoost model, exhibited a significantly lower accuracy of 75.00% compared to the developed model. Similarly, the third existing study achieved an accuracy of 90.00% in identifying inauthentic information, yet still fell short of the developed model's performance. Despite both utilizing the same ML algorithm, the fourth existing model achieved a lower accuracy of 80.00% compared to the first model. Lastly, the fifth baseline study, employing a Legitimacy model that combines two DTs and two Natural Language Process models in an ensemble, yielded an accuracy of 96.90%. This comparative analysis unequivocally demonstrates the indisputable superiority of the developed model in terms of Fake News Identification when compared to the five existing models.

A unique notable improvement over the existing studies lies in the handling of dataset features, involving meticulous extraction and effective combination of text features. Additionally, the integration of multiple algorithm performances into a unified model through ensembling constitutes a significant enhancement. Consequently, the exceptional performance of the developed model can be attributed to its ability to address the weaknesses and biases observed in the existing model. Moreover, the developed model emphasizes comprehensive data cleaning and feature extraction, aspects that are sometimes overlooked in existing studies.

5. SECTIONS

In this groundbreaking study, a cutting-edge model for identifying fake news is being introduced. The primary aim of this innovative tool is to counteract the rampant spread of unverified information in the modern era. As the purveyors of inauthentic content persistently refine their methods to create and circulate fabricated news, the model harnesses the power of a benchmark dataset, meticulously for comprehensive training and testing.

To give information seekers and social media users the tools to make informed decisions before sharing content, the Scikit-learn Python library was brought into pre-processing the dataset. By leveraging the TF-IDF ML technique, essential features crucial for accurate classification were extracted.

To evaluate the model's performance, established ML evaluation metrics like Accuracy, F1-Score, Precision, and Recall were employed. These metrics provided dependable benchmarks to measure the approach's effectiveness. A comparison of the outcomes with existing fake news identification models revealed that the improved model surpasses its forerunners.

This study represents a crucial milestone in the automated identification of fake news, actively contributing to the ongoing battle against the dissemination of fabricated information, particularly across social media platforms. By combining advanced techniques, rigorous evaluation, and a commitment to empowering users, the model represents a significant step toward fostering a more informed and trustworthy information ecosystem.

DECLARATIONS

Funding:

This research received no specific grant from any funding agency. The authors independently conducted this study without any financial support.

Conflicts of Interests:

The authors declare that they have no competing interests related to this work. There are no financial or non-financial interests that could influence the research, results, or conclusions presented in this paper.

Availability of data and material:

The dataset used for this research work is the Kaggle dataset, which is available at https://www.kaggle.com/competitions/fake-news/data. The dataset contains both true and false news labeled as 1 for fake and 0 for true, along with the authors of the news, the title, and the textual sentence associated with the news.

Code availability:

The code used for the experimental side of the research wasn't made public because of the issue of copyright restrictions. However, interested researchers can request access to the code by sending an email to the corresponding author via (obihenryegaga@gmail.com). Striving to provide measured responses to inquiries, all the while ensuring compliance with relevant licensing terms and restrictions.

Authors contributions:

Timothy Moses provided the overall supervision and mentorship, at the same time facilitating research resources, offering critical insights during the study's progression, and reviewing the final version of the manuscript.

Obi, Henry Egaga conceived and designed the study, collected and curated the dataset, analyzed the dataset, and drafted the manuscript.

Christopher Ifeanyi Eke provided guidance and oversight throughout the time of this research work, including refining the research design, assisting with data interpretation, and critically reviewing and editing the manuscript.

Jeffrey Agushaka offered expertise in the research area, provided feedback on the research design, and contributed to revising the manuscript.

6. REFERENCES

[1] M. Aldwairi and A. Alwahedi, "Detecting fake news in social media networks," *Procedia Comput. Sci.*, vol. 141, pp. 215–222, 2018, doi: 10.1016/j.procs.2018.10.171.

[2] Y. Liu and Y. F. B. Wu, "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks," in *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, 2018, pp. 354–361.

[3] T. Finneman and R. J. Thomas, "A family of falsehoods: Deception, media hoaxes and fake news," *Newsp. Res. J.*, vol. 39, no. 3, pp. 350–361, 2018, doi: 10.1177/0739532918796228.

[4] R. K. Kaliyar, "Fake news detection using a deep neural network," *2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018*, pp. 1–7, 2018, doi: 10.1109/CCAA.2018.8777343.

[5] J. M. Burkhardt, "History of fake news," *Libr. Technol. Rep.*, vol. 53, no. 8, pp. 1–33, 2017.

[6] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer, "Political science: Fake news on Twitter during the 2016 U.S. presidential election," *Science (80-.).*, vol. 363, no. 6425, pp. 374–378, 2019, doi: 10.1126/science.aau2706.

[7] D. De Beer and M. Matthee, *Approaches to Identify Fake News : A Systematic Literature Review*, no. Macaulay 2018. Springer International Publishing, 2021. doi:

10.1007/978-3-030-49264-9.

[8] X. Zhou and R. Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," *ACM Comput. Surv.*, vol. 53, no. 5, 2020, doi: 10.1145/3395046.

[9] M. D. Molina, S. S. Sundar, T. Le, and D. Lee, "'Fake News' Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content," *Am. Behav. Sci.*, vol. 65, no. 2, pp. 180–212, 2021, doi: 10.1177/0002764219878224.

[10] R. Oshikawa, J. Qian, and W. Y. Wang, "A survey on natural language processing for fake news detection," *Lr. 2020 - 12th Int. Conf. Lang. Resour. Eval. Conf. Proc.*, pp. 6086–6093, 2020.

[11] J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," *Mach. Learn. with Appl.*, vol. 4, p. 100032, 2021, doi: 10.1016/j.mlwa.2021.100032.

[12] A. Gelfert, "Fake news: A definition," *Informal Log.*, vol. 38, no. 1, pp. 84–117, 2018, doi: 10.22329/il.v38i1.5068.

[13] E. C. Tandoc, Z. W. Lim, and R. Ling, "Defining 'Fake News': A typology of scholarly definitions," *Digit. Journal.*, vol. 6, no. 2, pp. 137–153, 2018, doi: 10.1080/21670811.2017.1360143.

[14] N. Mukerji, "What is Fake Make-Up?," *Fiji Times*, no. 679, p. 3300, 2016, [Online]. Available: http://www.consumersfiji.org/media/feature-articles/feature-articles-2016/what-is-fake-make-up

[15] C. C. Wang, "Fake news and related concepts: Definitions and recent research development," *Contemp. Manag. Res.*, vol. 16, no. 3, pp. 145–174, 2020, doi: 10.7903/CMR.20677.

[16] D. M. J. Lazer *et al.*, "The science of fake news," *Science (80-.).*, vol. 359, no. 6380, pp. 1094–1096, 2018, doi: 10.1126/science.aao2998.

[17] J. P. Baptista and A. Gradim, "A Working Definition of Fake News," *Encyclopedia*, vol. 2, no. 1, pp. 632–645, 2022, doi: 10.3390/encyclopedia2010043.

[18] D. Gaozhao, "Flagging fake news on social media: An experimental study of media consumers' identification of fake news," *Gov. Inf. Q.*, vol. 38, no. 3, p. 101591, 2021, doi: 10.1016/j.giq.2021.101591.

[19] G. Murphy, E. F. Loftus, R. H. Grady, L. J. Levine, and C. M. Greene, "False Memories for Fake News During Ireland's Abortion Referendum," *Psychol. Sci.*, vol. 30, no. 10, pp. 1449–1459, 2019, doi: 10.1177/0956797619864887.

[20] S. Banaji and R. Bhat, "WhatsApp Vigilantes: An exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India," *LSE Media Commun.*, vol. 2, pp. 1–14, 2020, [Online]. Available: https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/WhatsApp-Misinformation-Report.pdf

[21] S. van der Linden, C. Panagopoulos, and J. Roozenbeek, "You are fake news: political bias in perceptions of fake news," *Media, Cult. Soc.*, vol. 42, no. 3, pp. 460–470, 2020, doi: 10.1177/0163443720906992.

[22] M. Cantarella, N. Fraccaroli, and R. G. Volpe, "Does Fake News Affect Voting Behaviour?," *SSRN Electron. J.*, vol. 18, no. 6, 2020, doi: 10.2139/ssrn.3629666.

[23] T. Lee, "The global rise of 'fake news' and the threat to democratic elections in the USA," *Public Adm. Policy*, vol. 22, no. 1, pp. 15–24, 2019, doi: 10.1108/pap-04-2019-0008.

[24] V. Bakir and A. McStay, "Fake News and The Economy of Emotions: Problems, causes, solutions," *Digit. Journal.*, vol. 6, no. 2, pp. 154–175, 2018, doi: 10.1080/21670811.2017.1345645.

[25] M. Keenan and K. Dillenburger, "How 'fake news' affects autism policy," *Societies*, vol. 8, no. 2. 2018. doi: 10.3390/soc8020029.

[26] S. M. Jang and J. K. Kim, "Third person effects of fake news: Fake news regulation and media literacy interventions," *Comput. Human Behav.*, vol. 80, no. March, pp. 295–302, 2018, doi: 10.1016/j.chb.2017.11.034.

[27] G. E. R. Agudelo, O. J. S. Parra, and J. B. Velandia, "Raising a model for fake news detection using machine learning in Python," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11195 LNCS, pp. 596–604, 2018, doi: 10.1007/978-3-030-02131-3_52.

[28] A. Jain, A. Shakya, H. Khatter, and A. K. Gupta, "a Smart System for Fake News," pp. 2–5, 2019.

[29] L. Waikhom and R. S. Goswami, "Fake News Detection Using Machine Learning," *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3462938.

[30] K. S. Veda, K. Sudarshana, and Amulya, "A Novel Technique for Fake News Detection using Machine Learning Algorithms and Web Scrapping," 2020, [Online]. Available: https://semanticscholar.org/paper/a1abd8cf5cb9630946f59207d40cdec3dab5b033

[31] U. Sharma, S. Saran, and S. M. Patil, "Fake News Detection Using Machine Learning Algorithms," 2020, [Online]. Available: https://semanticscholar.org/paper/d8e7a9bc3e5dbebc4c5f8711e7cc5422a00631c9

[32] P. A. Yerlekar, "Fake News Detection using Machine Learning Approach Multinomial Naive Bayes Classifier," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 5, pp. 1304–1308, 2021, doi: 10.22214/ijraset.2021.34509.

[33] N. Islam *et al.*, "Ternion: An autonomous model for fake news detection," *Appl. Sci.*, vol. 11, no. 19, pp. 1–15, 2021, doi: 10.3390/app11199292.

[34] E. Hossain, N. Kaysar, A. Zahid, J. Uddin, M. Rahman, and W. Rahman, "(2) A Study towards Bangla Fake News Detection Using Machine Learning and Deep Learning | Request PDF," *Springer*, no. December, pp. 79–95, 2022, doi: 10.1007/978-981-16-5157-1.

[35] A. M. Ali, F. A. Ghaleb, B. A. S. Al-Rimy, F. J. Alsolami, and A. I. Khan, "Deep Ensemble Fake News Detection Model Using Sequential Deep Learning Technique," *Sensors*, vol. 22, no. 18, 2022, doi: 10.3390/s22186970.

[36] W. A. Qader, "An Overview of Bag of Words ; Importance , Implementation , Applications , and Challenges," *2019 Int. Eng. Conf.*, no. June 2019, pp. 200–204, 2020, doi: 10.1109/IEC47844.2019.8950616.

[37] Ş. B. Hayta, H. Takçi, and M. EmİnlÍ, "Language identification based on n-gram feature extraction method by using classifiers N-GRAM FEATURE EXTRACTION METHOD BY USING CLASSIFIERS," no. July 2020, 2013.

[38] T. J. Bahzad and A. M. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," no. January, 2021, doi: 10.38094/jastt20165.

[39] N. Mohapatra, K. Shreya, and A. Chinmay, "Optimization of the Random Forest Algorithm," in *International Journal of Geoinformatics*, 2020, pp. 201–208. doi: 10.1007/978-981-15-0978-0_19.

[40] S. Sperandei, "Understanding logistic regression analysis," *Biochem. Medica*, pp. 12–18, 2014, doi: 10.11613/BM.2014.003.

[41] T. P. Bagchi, "S upport Vector Machines," no. January, 2022.

[42] S. Fafalios, P. Charonyktakis, and I. Tsamardinos, "Gradient Boosting Trees," no. April, pp. 1–13, 2020, [Online]. Available: https://www.gnosisda.gr/wp-content/uploads/2020/07/Gradient_Boosting_Implementation.pdf

[43] Y. Liu, Y. Zhou, S. Wen, and C. Tang, "A Strategy on Selecting Performance Metrics for Classifier Evaluation," *Int. J. Mob. Comput. Multimed. Commun.*, vol. 6, no. 4, pp. 20–35, 2014, doi: 10.4018/IJMCMC.2014100102.

[44] C. I. Eke, A. A. Norman, and L. Shuib, "Context-Based Feature Technique for Sarcasm Identification in Benchmark Datasets Using Deep Learning and BERT Model," vol. 9, pp. 48501–48518, 2021, doi: 10.1109/ACCESS.2021.3068323.

[45] A. Tasnim, M. A. Rahman, and J. Akhter, "Performance Evaluation of Multiple Classifiers for Predicting Fake News," pp. 1–21, 2022, doi: 10.4236/jcc.2022.109001.

[46] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, "Sarcasm identification in textual data: systematic review, research challenges and open directions," *Artif. Intell. Rev.*, vol. 53, pp. 4215–4258, 2020.