Detecting Fake News Using Machine Learning: A Comparative Study of Techniques

Howaida Abdul Hadi Al ibraheemi
Department of Computer Science,College Of Education for Pure Sciences
University of Kufa,
Najaf, Iraq.
Howaidaa.alibraheemi@student.uokufa.edu.iq
Orcid.org/0009-0007-5972-1100

Mohammed Jabardi
Department of Computer Science, College Of Education for Pure Sciences
University of Kufa,
Najaf, Iraq.
mohammed.naji@uokufa.edu.iq

DOI: http://dx.doi.org/10.31642/JoKMC/2018/1 I0213
Received May. 20, 2024. Accepted for publication Aug. 10, 2024

Abstract**—***Many people get their news via the Internet and social media platforms, and given the rapid growth of these platforms, fake news may now spread easily and quickly. False information that aims to mislead and harm society and the individual is known as fake news. By deliberately spreading false information, the media distort public opinion and threaten the social order by leading people to believe things that are not true. With the massive expansion of social media networks, the spread of fake news has increased dramatically. Although interesting, it poses some difficulties due to limited resources (such as datasets and published research). This paper presents diverse machine-learning techniques to identify fabricated news by analyzing the textual content. Several techniques were used, including SVM, RF, logistic regression, Naive Bayes, Gradient Boosting, AdaBoost, KNN, DT, and XGBoost. Based on comparing the results, who got the best result with an accuracy rate of 0.9967 and the lowest loss of 0.003 The study includes a variety of methodologies, such as natural language processing (NLP), machine learning, and data mining, which have been found to improve the efficiency of text processing to increase accuracy and can save time and effort by automatically identifying fake news, especially in light of the massive amount from materials available on the Internet.*

Keywords— Fake news detection, natural language processing (NLP), Machine learning, TF_IDF, classification text.

I. INTRODUCTION

Misleading information is now widely disseminated due to the rise in popularity and speed of Internet use brought about by the advancement of information technology. News is any information intended to notify the public about local happenings that might affect their social or personal lives [1]. In recent years, online social media has become a well-liked medium for distributing news for political, economic, and entertainment reasons [2]. This component demonstrated both beneficial and detrimental effects [3]. People disseminate false information under the guise of legitimate news to amuse themselves and make money. The fake news circulating on Facebook in recent years has greatly influenced the 2016 US presidential election campaign. [1]. In light of this tragedy, many businesses and research institutions are concentrating on comprehending the phenomenon and reducing fake news. Many scholars who examine false news and social influence have used terms like "rumors," "misinformation," and "fake news." Politics, the economy, and public opinion may all suffer from fake news. One well-known example of false news is the assertion that Barack Obama was hurt in the explosion that destroyed equities worth $130 billion [4]. Several false reports about the COVID-19 epidemic have sowed mistrust and concern, which has caused social media systems to crash. The most unsettling thing on the Internet is people's lack of confidence in one another and bogus news. Classifying news, particularly fake news, is a complex task that has engaged the attention of many scholars. While some have approached it as a binary classification (i.e., real or fake) [5], others have found this approach problematic and have suggested multiclass classification [6], regression, or clustering as alternative methods. In their research, these scholars have employed various techniques to identify and classify false information, each with its criteria.

1) Propagation-based [7] approaches: use people's responses and shares to track the trend of news dissemination.

2) User profile-based [8] approaches: monitor users' actions by utilizing the news posted, forwarded, or commented on.

This information includes other analytical data such as location, sexual orientation, followers, or friends.

3) Content-based [9] methods: There are two types of news:

(*a) Syntactic-based* approaches categorize the news using linguistic and writing patterns such as verbs, memorable characters, and nouns.

(b) using semantic-based approaches. High-level representation and structure of the text in a given document are carried out.

II. LITERATURE REVIEW

Research into fake news detection has been ongoing in recent years. However, most of this research focuses on analyzing and detecting fake information recognition from Internet networks and articles. So, the authors proposed numerous strategies. Harita Reddy et al. [10] collected techniques on stylometric highlights and word vector highlights of the news items' content used to discover fake news with an accuracy of up to 95.49%. Another study like [11] used a theory-driven approach using machine learning methods, including SVM, Random Forest, XGBoost, Naive Bayes, and Logistic Regression, along with misinformation and clickbait features taken from news article text to identify bogus news. Another researcher, such as [12], developed the FakeNewsTracker system, which automatically gathers news data and distinguishes between authentic and false news. Employ word n-grams and character n-grams analysis to test two machine learning methods. They employed word n-grams and character n-grams analysis to test two machine learning methods. A gradient boosting classifier and term-frequency-inverse document frequency (TF-IDF) were also used to get better results with character n-grams in [13], with a 96% success rate. Superior results using character n-grams using a Gradient Boosting Classifier and Term-Frequency-Inverse Document Frequency (TF-IDF) were attained in [14], Using word embedding to improve linguistic features for detecting false news; we applied AdaBoost and bagging ensemble machine learning techniques to 57 features, to compare the effectiveness of manual versus machine false news identification. Moreover, [15] employed linguistic characteristics such as n-grams, punctuation, readability, and syntax-based features. Also, [16] presented a novel semantic approach for identifying fake news that improved accuracy by focusing on social features such as opinions, components, or realities that could be adequately extracted from information. In [17], semantic detection of fake news using machine-learning techniques is presented.

III. PROPOSED SYSTEM

This study used machine learning techniques on fake news through text analysis using natural language (NLP) and extraction techniques. Texts were converted to digital representations using TF-IDF technology. Then, several machine learning models were trained, including SVM, RF, Logistic Regression, Naive Bayes, Gradient Boosting, AdaBoost, KNN, DT, and XGBoost. Finally, the model's performance was evaluated. It was applied to improve fake news detection operations.

A-Data Description

The database is from the Kaggle website by (BHAVIK JIKADARA[1]); it is a data collection of 43 MB of news items with the following features: title, text, subject, and data. Of the articles, 21417 are accurate (the label is 1), and 23481 are false (the label is 0). Figures 1 and 2 show the data before and After preprocessing; we used only the text of the article and the text.

(https://github.com/Bhavik−Jikadara/Fake−News−Detection[1])

	title	text	subject	date	class
0	Donald Tru...	Donald Trum...	News	December 31...	0
1	Drunk Brag...	House Intel...	News	December 31...	0
2	Sheriff Da...	On Friday, ...	News	December 30...	0
3	Trump Is S...	On Christma...	News	December 29...	0
4	Pope Franc...	Pope Franci...	News	December 25...	0

Fig. 1. **A** part of the dataset is uncleaning

	text	class
4893	You know how we often jest, largely because it...	0
17864	When did our Washington DC Mall become a stagi...	0
6464	Now that the side-show is essentially over and...	0
18366	(Reuters) - An opposition leader who said she ...	1
11029	William Stanford Nye or Bill Nye is an America...	0

Fig. 2. A part of the dataset cleaning

The distribution of classes in our dataset is shown in Figure (3); the sample is reasonably balanced, with 23,468 instances of false information and 21,202 instances of real news. This balance is essential to our research because it avoids model bias toward any specific class during training. A balanced guarantee that the performance acquired during training is accurate.

```
1: true
0: fake
Distribution of class:
class
0    23478
1    21211
Name: count, dtype: int64
```

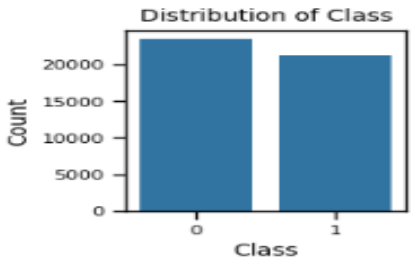

Fig. 3. Dataset Labels of Ratio

B- Data Preprocessing

Preprocessing can be compared to text mining since text expression is challenging because it is noisy and inconsistent.

Several strategies convert inconsistent data into more consistent data that the computer can understand [18]. In this investigation, we used the following techniques: Remove HTML tags: Web scraping often adds HTML components to the text when collecting datasets. These items should be removed, even if they are frequently ignored. Eliminating stop words, including "a," "an," "the," "of," and "is," can significantly speed up processing and draw attention to the main idea of the phrase.

- Remove special characters**:** Since they are not letters or numbers, special characters like &, %, and $ are unreadable.
- Lemmatization and stemming**:** This process reduces words to their primary form.
- Normalization: NLP libraries are case-sensitive, so all text is converted to lowercase.

More extensive input paragraphs are encoded or reduced to a single word.

C- Feature Extraction

Text can be converted to vectors using various methods, including word2vec, TF-IDF, and N-gram. Our study used the TF-IDF technique, which performs well on non-large datasets. TF-IDF vector: One of the most popular feature extraction techniques is inverse document frequency (TF-IDF) [19]. This method has two stages: the term frequency (TF) is calculated in the first stage, and the inverse document frequency (IDF) is calculated in the second. A statistical analytic technique called TF-IDF analyses a word's significance within a corpus or group of documents. It is based on the number of times a term appears in the article and the term frequency-inverse document frequency (TF-IDF). TF divides the number of times a word appears in the article by the total number of words to avoid bias against lengthy papers. With fewer documents containing a word indicating a strong word's capacity to discriminate, IDF refers to the frequency of reverse documents. Regarding longer texts, TF-IDF excels in RNN and other neural network models in text feature extraction [20].

D- models classifier

Our model was developed using nine well-known machine-learning techniques:

- Logistic regression (LR**)** assesses categorical issues, the binary results of the popular version of the LR model are true/false, yes/no, and others. LR is also offered with several outcomes in place of this multinomial. LR reads the input vector and maps it to the proper category using the logistic or sigmoid function. Due to its flexibility and robustness as a classification approach, we utilized LR in this article for assessment [21]
- Decision trees (DT) use recursive partitioning of all features in the training dataset to predict the final class. With nodes standing in for features, branches for decisions, and leaves for outcomes, the dataset is depicted as a tree. We started with the data as input and gradually divided it into smaller chunks until the output indicated whether the data was real or fake [22].
- Random forest (RF**)** combines several trees, which is why it's called a forest. It is applicable to use cases involving both regression and classification. RF uses ensemble learning to prevent over-fitting and merges numerous DTs to increase the model's accuracy. We employed this classifier to expedite our suggested model's training and learning process. Since this study focused on binary problems, the outcome of RF is determined by tallying the votes cast by each tree, which can be either 0 or 1. The highest number of votes determines the outcome of RF. Where $\hat{r}=\frac{1}{n}\sum_{i=1}^{n}\hat{r}(a)$, Here, "N" stands for the total number of trees in the forest, "i" for the current tree, and "a" for the training data. $\hat{r}$ is the tree prediction [23].
- Naive Bayes (NB) uses maximum conditional probability to classify news as authentic or phony. The foundation of it is "Bayes' Theorem.

$P(X|Y) = \frac{P(Y|X) * P(X)}{P(Y)}$ When X and Y represent two occurrences. Since the NB classifier is easy to use and reasonably priced to compute, we used it for text classification. Compared to other classifiers, NB requires less data for training [24].

- SVM Support Vector Machine is a supervised approach based on machine learning mostly applied to classification issues. The SVM algorithm aims to identify the best hyperplane in an N-dimensional space or optimal line in a 2D space for classifying the points. Such a hyperplane is used to maximize the margin between the classes to discriminate between the two classes of points. A data point's location on either side of the plane can be categorized into several classifications. The primary goal of the SVM approach is to increase the distance between the data points and the hyperplane. The hinge loss function is the loss function that maximizes the margin. The hyperplane's equation is

$w^x + b = 0$ (1)

[25].

Where the bias (b) and weight vector (w) are expressed.

$l(w) = \sum(max(0,1 - y_i[w^t x_i + |b|]) + \lambda||w||^2$ (2) [26]

.

Any errors resulting from data points closer to the categorization boundary than the margin are calculated using the loss function, which is the first term. The second term, the regularization function, is employed to prevent overfitting.

- Gradient Boosting: Super Gradient Boosting (XGBoost) is a boosting classifier designed for supervised machine learning methods. To create a model that is more reliable and accurate

This approach involves many weak learners, such as decision trees. Its basic idea is to train multiple models one by one with reinforcement. Each model attempts to handle errors.

Learners are combined to create a final prediction. Thus, it provides high accuracy and robustness to improve grading results. XG and SF comprise the two properties of this classifier. The first characteristic was used in this study [27].

- KNN: K-nearest neighbor (K-NN): This approach, which comes from supervised machine learning techniques, is known for its simplicity and ease of use. It was initially applied to resolving regression and classification issues in the early 1970s. The similarity between neighbors' ideas serves as the foundation for this method. Cases are categorized based on the majority of votes cast by their neighbors, as the similarity principle is dependent on the value of K. This happens due to nearby, comparable occurrences [28] [29].

• AdaBoost: techniques combine several weak classifiers to create a robust classifier. The suggested approach combined many decision tree classifiers with various hyperparameters using AdaBoost to produce a more reliable and accurate classification.
Weak classifiers are iteratively trained using AdaBoost, and their weights are modified based on their performance on the training set. The weak classifiers are then combined using a weighted majority vote to create the final classifier [30].
• XGBoost: This is well-liked by rivals in the machine learning space and practitioners in the field. Gradient-boosting decision trees are implemented using XGBoost, which offers speed and performance. Y.
For binary and multiclass classification tasks, it does rather well. XGBoost is employed in the two regressions.
For our classification task, classification issues are predicted to perform well. Resolution XGBoost
Due to its gradient-boosting nature, it typically performs better than random forests [31].

IV. METHODOLOGY

This study used different ML techniques to identify fake news, including SVM, LR, RF, Naive Bayes Gradient Boosting, AdaBoost, KNN, DT, and XGBoos. Data was collected, pre-cleaning, as well as extracting important features. Figure 4 shows the method of partitioning the training data.

Fig. 4. Data partition visualization

The proposed model combines several machine-learning techniques to classify whether the news is fake or real. The block diagram of the proposed model is shown in Figure 5.

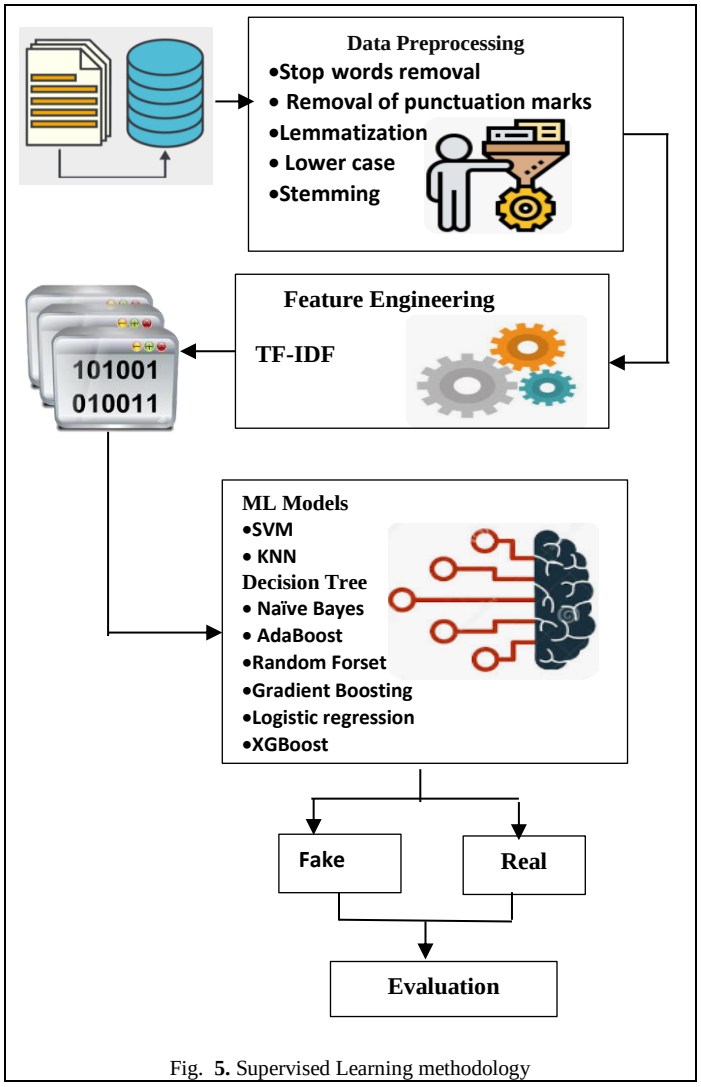

Fig. **5.** Supervised Learning methodology

V. PERFORMANCE EVALUATION MATRIC

We use accuracy (A), precision (P), recall (R), and F1-score (F) as assessment measures to compare and evaluate our model. Equations 1 calculate accuracy, and 2 and 3 calculate precision and Recall. On the other hand, as stated in equation 4, the F1-score is the harmonic mean for recall and precision.

1)Accuracy: The classification accuracy is the number of samples in the source data set that were properly matched with each sample in the data set. A "TP" result is an output from the classifier that has a positive forecast and a real class. The (TN) is what a classifier gives you when the expected class is also a true negative. When the classifier guesses a positive result when *the* real result is a negative one, this is called a false positive (FP) or false negative (FN) classification error [32]. Eq. (3)Explains the accuracy calculation.

$$A = \frac{TP+TN}{TP+FN+FP+TN} \quad (3)$$

2)Precision: Precision is the ratio of correctly categorized positive class values to the sum of incorrectly classified

positive class values. It provides information about the model's factual accuracy [33].

$$P=\frac{True\ Positive}{True\ Positive + False\ Positive} \quad (4)$$

3)Recall: The ratio of correctly categorized positive class values to the total of correctly classified positive class values and incorrectly classified negative class values is known as the recall rate. It provides information on the model's completeness [33].

$$R= \frac{True\ Positive}{TruePositive + FalseNegative} \quad (5)$$

4)F1-Measure: The F1-score harmonically represents recall and precision [32]. Score F1 is explained in Equation (6).

$$F1= 2* \frac{precision \cdot recall}{precision + recall} \quad (6).$$

VI. Results and Discussion

In this section, the final results of this study are presented. Several different machine learning classifiers were used in combination with the TF-IDF method. As shown in Table 1. The most effective classifiers are XGBoost, Random Forest (RF), and Gradient Boosting, all of which earn accuracy ratings higher than 0.99. SVM has exceptional performance, achieving an accuracy of 0.992 and demonstrating great results in other measures. - The SVM's capacity to process feature spaces with many dimensions and identify the most effective decision boundaries is likely a critical factor in its success in detecting fake news.

Logistic Regression and Decision Tree classifiers have favorable outcomes, achieving accuracy ratings exceeding 0.98 and 0.99, respectively. Naive Bayes and K-Nearest Neighbors (KNN) demonstrate the poorest performance compared to the other classifiers tested. The assumption of feature independence made by Naive Bayes may not sufficiently capture the intricate relationships present in fake news data, resulting in its relatively weaker performance. - The KNN classifier's accuracy of 0.7127 indicates that the proximity-based approach may not be suitable for effectively detecting fake news.

Table 1: Classification Result of Different Machine Learning Methods.

Classifier	Evaluation metrics				
	Accuracy	*Loss*	*Recall*	*Precision*	*F-score*
LR	0.9846	0.015364	0.985605	0.981523	0.9835
SVM	0.9920	0.007906	0.992642	0.990425	0.9915
RF	0.9964	0.003580	0.997761	0.994579	0.9961
Naive Bayes	0.9279	0.072047	0.91170	0.93228	0.9218
Gradient Boosting	0.9953	0.004624	0.998401	0.99173	0.9950
KNN	0.7127	0.287291	0.420026	0.92075	0.5768
DT	0.9956	0.004326	0.995521	0.99520	0.9953
AdaBoost	0.9947	0.005221	0.995841	0.99298	0.9944
XGBoost	0.9967	0.003282	0.99840	0.99458	0.9964

The Python code that runs the algorithm code on the Anaconda platform automatically obtains the confusion matrix using the cognitive learning module.

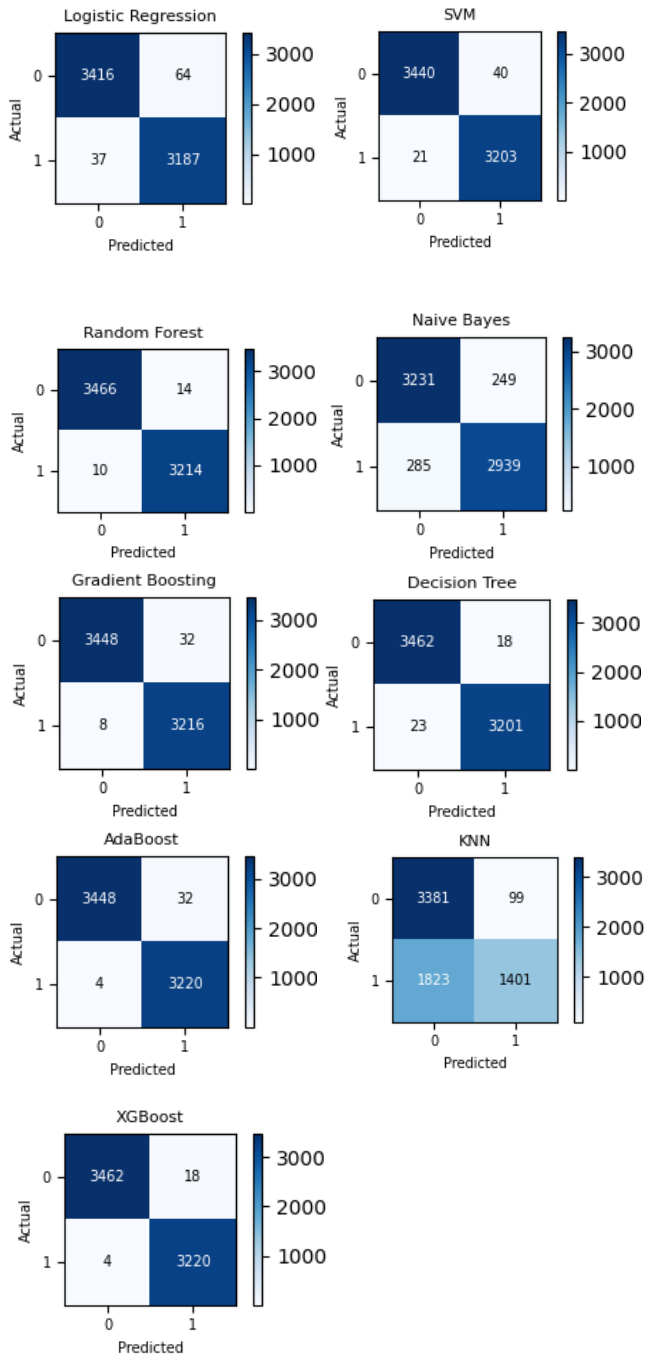

Fig. 6. Results of confusion matrices for technique

Figure (7) (8) shows the accuracy and loss results of all techniques. The results showed that the best technique for this study, which used data from Kaggel, was the XGBOOST technique. Figure (9) shows all the measurements for this technique.

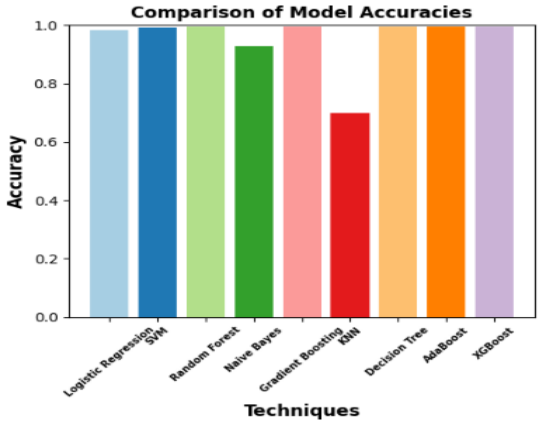

Fig. 7. Results accuracy of all techniques

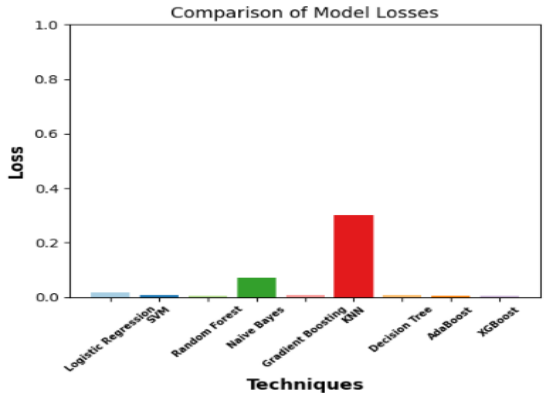

Fig. 8. Results in the loss for different techniques

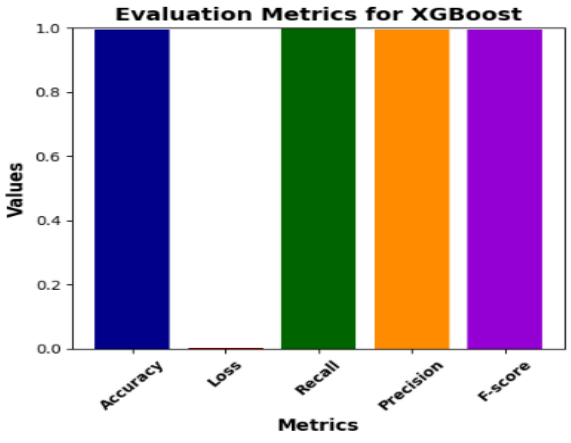

Fig. 9. Evaluation metrics XGBoost

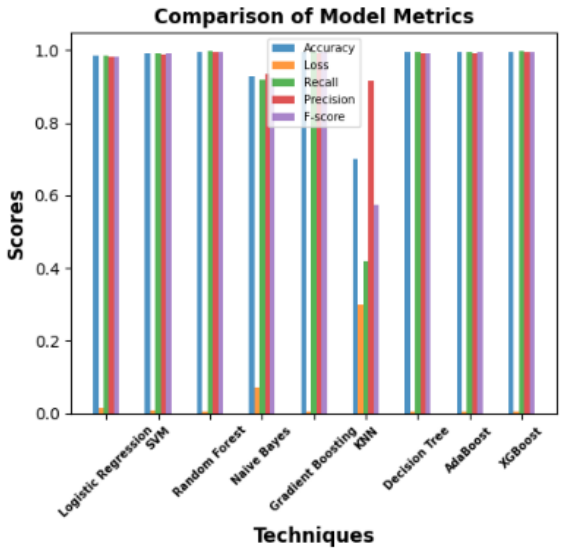

Fig. 10. Comparing different approaches to statistics and accuracy.

VII. Conclusion and Future Works

This paper compares various machine learning algorithms, such as Support Vector Machines (SVM), Random Forest (RF), logistic regression, naive Bayes, Gradient Boosting, AdaBoost, K-nearest neighbors (KNN), Decision Trees (DT), Extreme Gradient Boosting (XGBoost), for identifying fake news using textual content. The results indicate that the model with the highest performance attained an accuracy rate of 0.9967 and the lowest loss of 0.003, confirming the proposed approach's efficacy. The study utilizes a range of approaches, including natural language processing (NLP), machine learning, and data mining, to enhance the effectiveness of text processing and enhance the precision of false news identification. The automated detection of false information can be a time-saving and efficient solution, particularly considering the vast volume of online content. It can also help combat the quick dissemination of inaccurate information on the Internet and social media platforms.

The study incorporates a comparison analysis of various machine learning algorithms, a widely used method in the academic field to choose the most efficient models for a specific task. The investigation's utilization of natural language processing, machine learning, and data mining strategies aligns with the techniques employed in previous research on the detection of fake news. An essential addition to the work is the focus on how the proposed approach can efficiently and effectively identify fake news, particularly given the abundance of internet content.

Exploring the integration of supplementary data sources, such as user engagement metrics, social network analysis, or multimodal content (e.g., text, images, videos). Investigating the advancement of increasingly intricate machine learning models, such as deep learning architectures or advanced natural language processing approaches. Tackling the issues related to the scalability, real-time detection, and interpretability of the false news detection system.

References

[1] Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, and P. S. Yu, "TI-CNN: Convolutional Neural Networks for Fake News Detection," 2018, [Online]. Available: http://arxiv.org/abs/1806.00749

[2] S. Gundapu and R. Mamidi, "Transformer-based Automatic COVID-19 Fake News Detection System," pp. 1–12, 2021, [Online]. Available: http://arxiv.org/abs/2101.00180

[3] G. Shrivastava, P. Kumar, R. P. Ojha, P. K. Srivastava, S. Mohan, and G. Srivastava, "Defensive modeling of fake news through online social networks," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 5, pp. 1159–1167, 2020, doi: 10.1109/TCSS.2020.3014135.

[4] Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science. 2018 Mar 9;359(6380):1146-51.

[5] A. Roy, K. Basak, A. Ekbal, and P. Bhattacharyya, "A Deep Ensemble Framework for Fake News Detection and Classification," 2018, [Online]. Available: http://arxiv.org/abs/1811.04670

[6] A. Al Mamun Sardar, S. A. Salma, M. S. Islam, M. A. Hasan, and T. Bhuiyan, "Team sigmoid at CheckThat!2021 Task 3a: Multiclass fake news detection with Machine Learning," *CEUR Workshop Proc.*, vol. 2936, pp. 612–618, 2021.

[7] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, "Fake News Detection on Social Media using Geometric Deep Learning," pp. 1–15, 2019, [Online]. Available: http://arxiv.org/abs/1902.06673

[8] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," no. i, 2017, [Online]. Available: http://arxiv.org/abs/1708.01967

[9] O. Ngada and B. Haskins, "Fake News Detection Using Content-Based Features and Machine Learning," *2020 IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2020*, 2020, doi: 10.1109/CSDE50874.2020.9411638.

[10] H. Reddy, N. Raj, M. Gala, and A. Basava, "Text-mining-based Fake News Detection Using Ensemble Methods," *Int. J. Autom. Comput.*, vol. 17, no. 2, pp. 210–221, 2020, doi: 10.1007/s11633-019-1216-5.

[11] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani, "Fake News Early Detection: An Interdisciplinary Study," vol. 1, no. 1, 2019, [Online]. Available: http://arxiv.org/abs/1904.11679

[12] K. Shu, D. Mahudeswaran, and H. Liu, "FakeNewsTracker: a tool for fake news collection, detection, and visualization," *Comput. Math. Organ. Theory*, vol. 25, no. 1, pp. 60–71, 2019, doi: 10.1007/s10588-018-09280-3.

[13] H. E. Wynne and Z. Z. Wint, "Content-based fake news detection using N-gram models," *ACM Int. Conf. Proceeding Ser.*, 2019, doi: 10.1145/3366030.3366116.

[14] G. Gravanis, A. Vakali, K. Diamantaras, and P. Karadais, "Behind the cues: A benchmarking study for fake news detection," *Expert Syst. Appl.*, vol. 128, pp. 201–213, 2019, doi: 10.1016/j.eswa.2019.03.036.

[15] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic detection of fake news," *COLING 2018 - 27th Int. Conf. Comput. Linguist. Proc.*, pp. 3391–3401, 2018.

[16] A. M. P. Braşoveanu and R. Andonie, "Semantic Fake News Detection: A Machine Learning Perspective," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11506 LNCS, pp. 656–667, 2019, doi: 10.1007/978-3-030-20521-8_54.

[17] A. M. P. Braşoveanu and R. Andonie, "Integrating Machine Learning Techniques in Semantic Fake News Detection," *Neural Process. Lett.*, vol. 53, no. 5, pp. 3055–3072, 2021, doi: 10.1007/s11063-020-10365-x.

[18] N. Ahmed and M. Rawat, "Identification of Fake News using Machine Learning and Deep Learning," *2023 Int. Conf. IoT, Commun. Autom. Technol. ICICAT 2023*, 2023, doi: 10.1109/ICICAT57735.2023.10263681.

[19] S. Ghosh and M. S. Desarkar, "Class Specific TF-IDF Boosting for Short-text Classification," pp. 1629–1637, 2018, doi: 10.1145/3184558.3191621.

[20] K. Li, "HAHA at FakeDeS 2021 : A Fake News Detection Method Based on TF-IDF and Ensemble Machine Learning," no. September, 2021.

[21] P. K. Verma, P. Agrawal, V. Madaan, and R. Prodan, "MCred: multi-modal message credibility for fake news detection using BERT and CNN," *J. Ambient Intell. Humans. Comput.*, vol. 14, no. 8, pp. 10617–10629, 2023, doi: 10.1007/s12652-022-04338-2.

[22] R. K. Kaliyar, A. Goswami, and P. Narang, "Multiclass Fake News Detection using Ensemble Machine Learning," *Proc. 2019 IEEE 9th Int. Conf. Adv. Comput. IACC 2019*, pp. 103–107, 2019, doi: 10.1109/IACC48062.2019.8971579.

[23] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, "Fake News Detection Using Machine Learning Approaches," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1099, no. 1, p. 12040, 2021, doi: 10.1088/1757-899x/1099/1/012040.

[24] M. Villagracia Octaviano, "Fake News Detection Using Machine Learning," *ACM Int. Conf. Proceeding Ser.*, pp. 177–180, 2021, doi: 10.1145/3485768.3485774.

[25] N. Ruchansky, S. Seo, and Y. Liu, "CSI: A hybrid deep model for fake news detection," *Int. Conf. Inf. Knowl. Manag. Proc.*, vol. Part F131841, pp. 797–806, 2017, doi: 10.1145/3132847.3132877.

[26] S. Bajaj, "' The Pope Has a New Baby !' Fake News Detection Using Deep Learning," *Cs 224N*, pp. 1–8, 2017, [Online]. Available: https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1174/reports/2710385.pdf

[27] M. A. Taha, H. D. A. Jabar, and W. K. Mohammed, "Fake News Detection Model Basing on Machine Learning Algorithms," *Baghdad Sci. J.*, 2024, doi: 10.21123/bsj.2024.8710.

[28] K. Alkhatib, H. Najadat, I. Hmeidi, and M. K. A. Shatnawi, "Stock Price Prediction Using K-Nearest

Neighbor Algorithm," *Int. J. Business, Humanit. Technol.*, vol. 3, no. 3, pp. 32–44, 2013, [Online]. Available: https://www.ijbhtnet.com/journals/Vol_3_No_3_March_2013/4.pdf

[29] A. Sharma, I. Singh, and V. Rai, "Fake News Detection on Social Media," *2022 2nd Int. Conf. Adv. Comput. Innov. Technol. Eng. ICACITE 2022*, pp. 803–807, 2022, doi: 10.1109/ICACITE53722.2022.9823660.

[30] K. Anuradha, S. K. P, N. P. E, and M. Vignesh, "Fake News Detection Using Decision Tree and Adaboost," *Eurchembull*, vol. 12, no. 12, pp. 570–582, 2022, doi: 10.31838/ECB/2023.12.s3.065.

[31] R. S. Utsha, M. Keya, M. A. Hasan, and M. S. Islam, "Qword at CheckThat! 2021: An extreme gradient boosting approach for multiclass fake news detection," *CEUR Workshop Proc.*, vol. 2936, pp. 619–627, 2021.

[32] S. A. Abdul Kareem and Z. F. Rasheed, "A Machine Learning Model for Cancer Disease Diagnosis using Gene Expression Data," *J. Kufa Math. Comput.*, vol. 10, no. 2, pp. 179–185, 2023, doi: 10.31642/jokmc/2018/100227.

[33] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, and G. S. Choi, "Learning Architecture (CNN-LSTM)," vol. 4, pp. 1–13, 2020, doi: 10.1109/ACCESS.2020.3019735.