False News Identification using Machine Learning

Mrs. Prathibha T S
Prathibha024@gmail.com
Assistant Professor, Dept. of CSE. Sridevi Institute of Engineering and Technology, Tumkur

Chandana N H	Kusuma R	Hemashree M
chandananh9@gmail.com	*kusumaraj134@gmail.com*	*hemashreemanjayya@gmail.com*

Abstract— The high rate of propagation of false news via internet media channels has been a major issue worldwide, causing misinformation and social unrest. This paper provides a machine learning-based method for false news detection by utilizing natural language processing methods. The model examines text content in order to identify news as real or false through Naïve Bayes, Logistic Regression, SVM, Random Forest, and Decision Tree algorithms. Evaluation is conducted on publicly available datasets gathered from various domains. Among them, Decision Tree models attained the maximum accuracy of 99.91%. The results prove the efficacy of ensemble and baseline models in identifying false narratives and show the potential of real-time fake news detection tools.

Index Terms—**Fake news, Social media, Web Mining, Machine Learning, Support Vector Machine, TF-IDF.**

I. INTRODUCTION

With the simplicity of information availability in today's digital era, how people consume news has changed entirely. Although it has simplified access to information, it has also provided fertile ground for the rapid dissemination of false news. Fake headlines, doctored stories, and twisted facts can be spread instantly on social media platforms, swaying public opinion and decision-making.

The reasons for the spread of false news are diverse, ranging from financial profit through increased web traffic to the strategic manipulation of political and social sentiments. One of the most striking instances of such behavior was seen during peak election times, when fake news campaigns were orchestrated to influence voter decisions, discredit political figures, and create public uncertainty. These concerted disinformation campaigns have raised serious concerns about the integrity of democratic processes and the general well-being of information ecosystems.

Given the magnitude and impact of fake news, it is essential to develop automated techniques for identifying and combating false information. In this project, we introduce a machine learning-based system for fake news article identification. The process includes text preprocessing, which involves cleaning the text data by removing unnecessary elements like stop words and special characters, as well as applying stemming to normalize the text.

The next step involves encoding the text, which means converting the preprocessed text into a numerical form using techniques like bag-of-words (Bow), n-gram modeling, and term frequency-inverse document frequency (TF-IDF). Feature extraction is further enhanced with additional metadata such as the news source, author, publication date, and sentiment expressed by the text. headlines as either real or fake For classification, we use a Support Vector Machine (SVM), a robust supervised machine learning algorithm, to classify news

II. RELATED WORKS

Several studies in the literature have explored the detection of fake news using machine learning techniques. Researchers have proposed a wide range of strategies, from linguistic analysis to network-based assessments, aiming to improve detection accuracy.

The study in [2] presents a classification of detection methods into two primary groups: content-based methods, which leverage natural language processing and machine learning algorithms, and propagation-based methods, which examine how news spreads across networks. Both approaches have demonstrated potential, though each has specific limitations depending on the context of application.

In [4], the authors introduced a decision tree classifier for detecting fake news, particularly focusing on articles appearing on Twitter. Their approach achieved an accuracy of approximately 76%, illustrating that even simple classifiers can offer competitive results. However, the method struggles to maintain high accuracy when applied to complex and diverse datasets.

Research conducted in [7] investigated the effectiveness of ensemble learning by combining multiple models, including Random Forests and Gradient Boosting classifiers. Their study showed that ensemble models could substantially increase detection accuracy, reaching up to 90% on benchmark datasets. Nevertheless, the computational demands and complexity of ensemble methods pose challenges for real-time implementation.

The authors of [11] emphasized the importance of incorporating user interaction features such as likes, shares, and comments as additional inputs to detection models. They demonstrated that combining content features with social engagement data significantly enhances model performance. However, their findings also noted that overreliance on platform-specific metadata can reduce the model's generalizability across different social networks.

In [5], researchers proposed a hybrid model integrating deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, with traditional feature-based classifiers. Although deep learning models showed strong capabilities in capturing long-term textual dependencies, they required large labeled datasets and high computational resources, making them less practical for small-scale or real-time applications.

Another notable contribution is found in [13], where the authors developed a comprehensive annotated dataset spanning multiple domains beyond politics, including health, finance, and entertainment. Their work highlighted the critical need for multi-domain datasets, as models trained on single-domain data often fail to generalize effectively across diverse topics.

Furthermore, the study in [10] explored the psychological and emotional dimensions of fake news, emphasizing that articles written with heightened emotional intensity tend to spread more rapidly. They incorporated sentiment analysis into the detection framework, finding that while emotional features improve classification, relying solely on sentiment analysis is insufficient for robust fake news identification.

Finally, numerous studies have pointed out a major flaw in current classification approaches: they tend to label news articles as absolutely real or absolutely fake. In reality, information credibility exists along a continuum. Future research must therefore move towards probabilistic classification models or confidence-based scoring systems to better capture the complex nature of misinformation in the real world.

III. Proposed system

The proposed system employs a news dataset to develop a decision model from support vector machine method. It uses the model to classify new news as fake or real.

A. *General architecture of the proposed system*

The suggested system accepts as input a set of comments and their corresponding information, e.g., date, source and author. It then converts them into a features dataset that can be utilized during the learning phase. This conversion is referred to as preprocessing, it does a sequence of operations like cleansing, filtering and encoding. The preprocessed data are split into two: one for training and the other for testing. The training module employs a training dataset and support vector machine algorithm to create a decision model that can be used for the training dataset. If the model is accepted

(i.e., it can obtain an acceptable accuracy rate), it can be retained and utilized and then training stops. Otherwise, the learning algorithm parameters are updated in an attempt to enhance the accuracy rate. Figure 1 depicts the overall scheme of the proposed system.
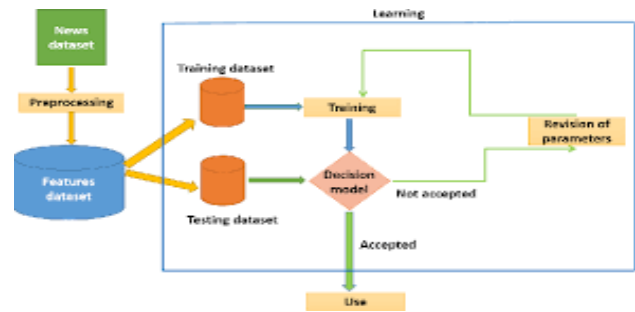

Figure 1. The proposed fake news detection system architecture's

B. *Preprocessing*

In the news dataset, news characteristics are classified into three categories: textual data, categorical data and numerical data. Each category preprocessing is performed through a set of operations as illustratrd in Figure 2:
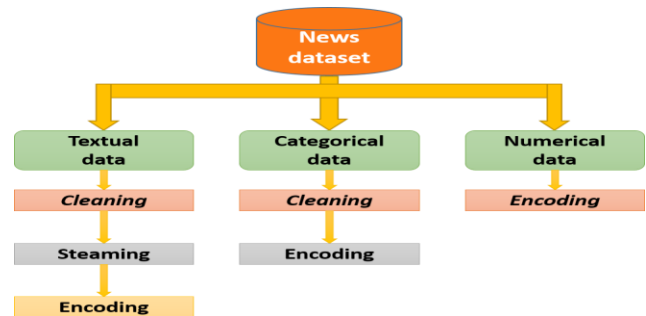

Figure 2. Preprocessing of different categories of news characteristics

Textual data. Represents the text typed by the author in a news article and pre-processed by the following operations
Cleaning: Removing stop words and special characters.
Steaming: Converting the useful words into their root forms.
Encoding: Converting the entire words of the comment into a numerical vector. This requires two steps: first combining two techniques — namely, Bag of Words [13] and N-grams [4] — then applying the TF-IDF method [12] on the outcome.

The TF-IDF is calculated as:

$$TF\text{-}IDF_t = TF_t \times IDF_t$$

Where:

- $TF_t = \dfrac{k}{n}$

  (the frequency of occurrence of term t in document n, divided by the total frequency k of terms in the document, maintaining the multiplicity of each term)

- IDFt=log(DtD)
  (where DDD is the total number of documents, and DtD_tDt is the number of documents containing the term ttt)

—

Categorical Data

Categorical data in the fake news detection system contains details such as the source of the news (newspaper, magazine, TV channel) and its author. These characteristics assist in comprehending the authorship and origin of news stories, which are essential in identifying misinformation patterns.

Pre-processing of categorical data involves two significant steps:

Cleaning: Special characters are stripped, and letters are made lowercase to maintain consistency throughout the dataset. This eliminates noise and allows the model to pay attention only to relevant information.

Encoding:

- Sources are labelled encoded with a unique numerical value assigned to each source.
- Authors are mapped with a special encoding technique, wherein authors from the same source are assigned near numerical values. This is achieved by creating a list consisting of the source and authors thereof, and replacing each author with their index number after incorporating the total of the preceding source sizes plus one.

This framework keeps the relationship between authors and corresponding sources intact even after numerical transformation, ensuring that source-author integrity is maintained in model training.

Numerical data.

In the field of fake news detection, various studies provide useful numerical insights:

- Over 60% of adults have encountered fake or misleading news on social media.
- The LIAR dataset contains 12,836 short news statements, labeled with six truthfulness ratings (e.g., false, barely true).
- A basic Naïve Bayes classifier achieved 74% accuracy on Facebook fake news data.
- Machine learning models using TF-IDF features and Linear Support Vector Machines (LSVM) reached 92% accuracy for classifying fake news.
- Including **social metadata** (likes, shares, comments) improved detection accuracy by **5-12%**.

C. *Learning*

It brings together two modules, namely, training and validation.

1) ***Training* :** To train our model, we employ the Support Vector Machine algorithm [15]. It utilizes a decision function to assign a degree of confidence to the classification: a positive value indicates true news along with its veracity level, while a negative value indicates fake news along with its degree of falsehood. Figure IV illustrates this concept.
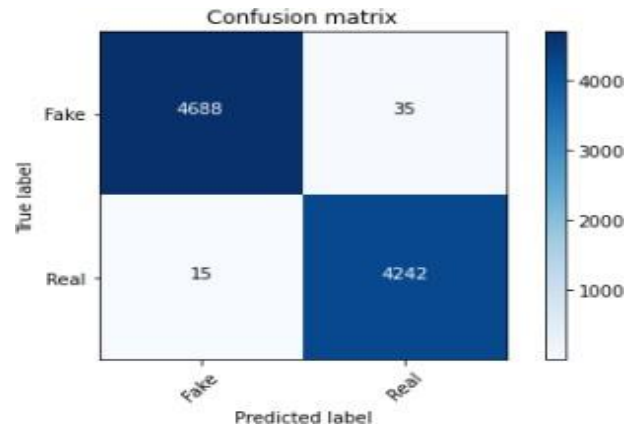

Figure 4 Confusion matrix news classification using support vector machine

SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.

$$h(\boldsymbol{x}_i) = sign(\sum_{j=1}^{s} \alpha_j\, y_j\, K(\boldsymbol{x}_j, \boldsymbol{x}_i) + b\,)$$

$$K(\boldsymbol{v}, \boldsymbol{v}') = \exp(\frac{||\boldsymbol{v} - \boldsymbol{v}'||^2}{2\gamma^2})$$

2) ***Validation*:** to estimate the ability of the model to identify new instances, we reserve some of the instances to serve as test models. The dataset of features is then split into two sets, a training set and a test set. Its utility lies in the prevention of over-fitting, i.e., testing the model on the same training set. The division is not random but is carried out in a specific sample using the technique of cross validation [10].

D. *Revision of parameters*

This operation is meant to enhance the accuracy of the model by adjusting or specifying the support vector machine algorithm parameters, i.e., Cost, ц, ϵ and modify the cross-validation variant [2].

E. *Use*

This is the final and most crucial step in our system. Once we've reached the optimal recognition rate, i.e., once we've constructed the optimal model, then we can utilize it on new, unlabelled news, and the model permits us to forecast their classes: wrong or true, with a confidence degree.

VI. EXPERIMENTS AND RESULTS

The performances of the proposed system was tested using a dataset that we build by merging a true news dataset with a fake news one.

A. Used Dataset

The dataset provided is a collection of news articles that have been preprocessed through several steps to prepare it for machine learning tasks, especially for fake news detection. Initially, each entry had the title, text, subject, date, and target (true/fake). However, during the cleaning and preparation phase, the title and date columns were dropped, leaving only the main text (the body of the article), the subject (which indicates the news category), and the target (the label specifying whether the news is true or fake). Additionally, the text was lowercased and punctuation was removed, making the data uniform and easier for natural language processing (NLP) models to handle

The dataset focuses mainly on three types of news subjects:

world news: Covers global and international news events, often sourced from professional agencies like Reuters.

politics: Focuses on political events, figures, or debates, but this category often includes fake or highly opinionated content.

politics News: Refers to more formal political news reporting, mainly from recognized sources and tends to be reliable.

Each news article is labeled under the target column as either:

true (real, factual news)

fake (misleading, exaggerated, or false news)

Overall, this dataset is ideal for building models that learn the difference between genuine journalism and fake or manipulative news. It includes cleaned and simplified article content for better performance in text analysis tasks.

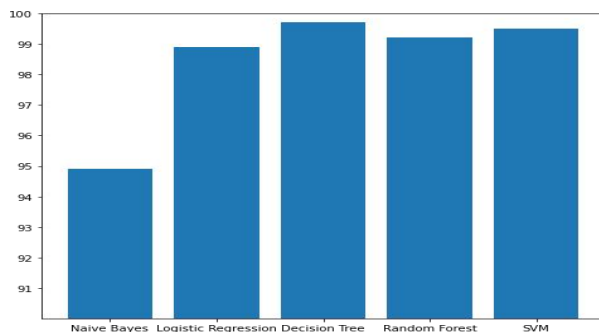## B. Results and discussion

Figure 5. Accuracy Comparison of Classification Algorithms

In Figure 5 the accuracy achieved by each algorithm on the final dataset. It is evident that the maximum accuracy achieved on Decision Tree which is 99.73%. The next highest accuracy is achieved on Support Vector Machine (SVM) which is 99.52%. The next highest accuracy is achieved on Random Forest of 99.22%. The next highest accuracy is achieved on Logistic Regression which is 98.91%. The least accuracy is achieved on Naïve Bayes which is 94.91%. Below Table Represents the name of the classifier and accuracy achieved by classifier
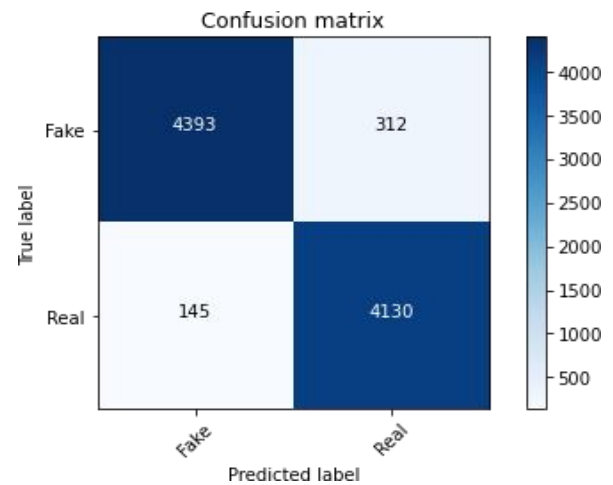

Figure 6. Confusion Matrix for Naïve Bayes Classifier on Fake News Dataset

In Figure 6 The confusion matrix reveals the performance of a Naive Bayes classifier in distinguishing between 'Fake' and 'Real' instances. Out of the actual 'Fake' instances, 4393 were correctly identified, while 312 were incorrectly classified as 'Real'. Conversely, of the actual 'Real' instances, 4130 were correctly classified, and 245 were misclassified as 'Fake'. This provides a clear picture of the model's true positives, true negatives, false positives, and false negatives, which are crucial for evaluating its effectiveness.
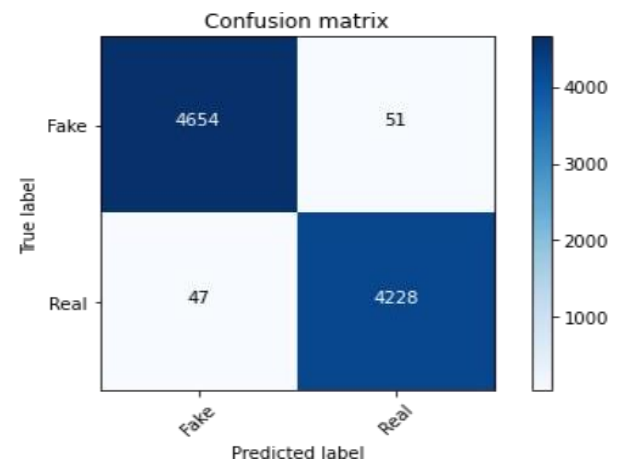

Figure 7.: Confusion Matrix for Logistic Regression Classifier on Fake News Dataset

In Figure 7 the provided confusion matrix for a logistic regression model, we can observe its performance in classifying "Fake" and "Real" instances. The model correctly predicted 4654 instances as "Fake" (True Negatives) and 4228 instances as "Real" (True Positives). However, it incorrectly classified 51 "Real" instances as "Fake" (False Negatives) and 47 "Fake" instances as "Real" (False Positives). This indicates that while the model demonstrates a strong ability to correctly classify both categories, there are still some instances where it makes incorrect predictions.
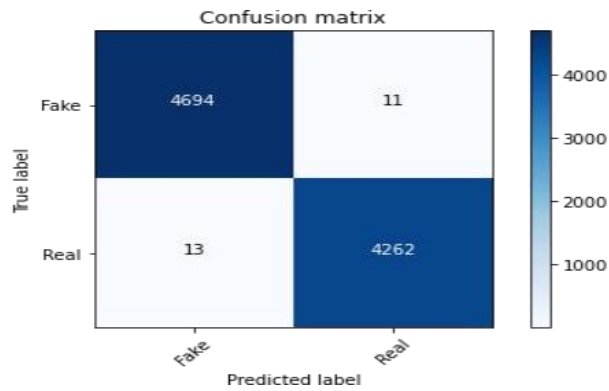
Figure 8. Confusion Matrix for Decision Tree Classifier on Fake News Dataset

In Figure 8, we notice that the influence of the feature "Sentiment" on the accuracy is almost negligible, which seems to be very logical: if a feeling released by a comment was negative it does not mean that it is fake. However, the characteristic "source" increased it up to 89.27%, and "date" up to 96%. While the author feature pushed it to 100%, which shows the effectiveness of the encoding we have proposed.
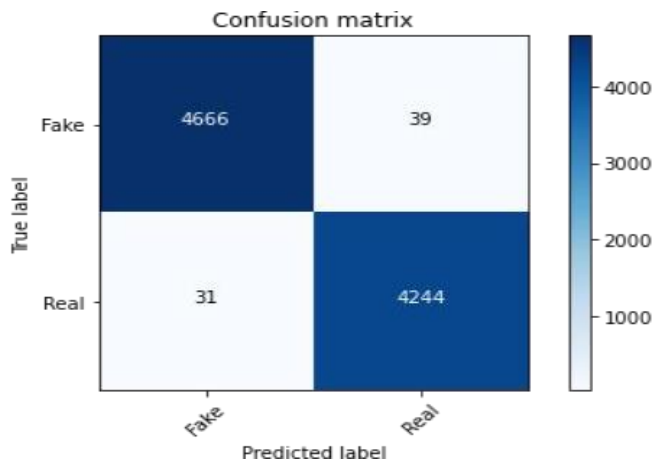

Figure 9. Confusion Matrix for Random Forest Classifier on Fake News Dataset

In Figure 9 It is clear that linear and polynomial kernels give the best results. The linear kernel is parameter less and faster, however, theoretically it cannot model the cases of complicated overlap of the two classes. On the other hand the Gaussian kernel makes it possible to model any type of overlap but its accuracy depends on the parameters C (Cost), ϵ and $ų$. We have studied the influence of these parameters on the precision of the model.

- Influence of Cost C: the following Figure 10 represents the evolution of accuracy according to the Cost C: by testing on the training dataset and using the RBF kernel of the LIBSVM method in WEKA:
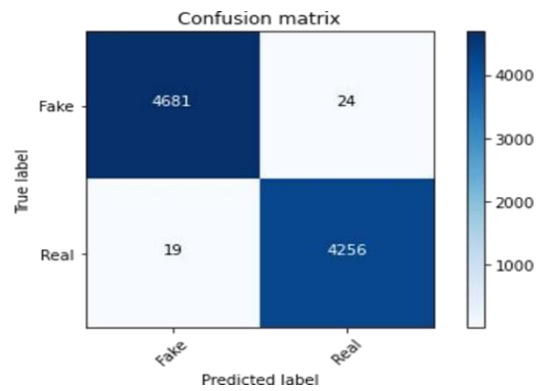

Figure 10. Confusion Matrix for Support Vector Machine (SVM) Classifier on Fake News Dataset

In Figure 10 the start the cost is equal to 0 and the value of the rate is 52% then by increasing the cost value we observe a rapid increase in the rate up to the value 150, then a stabilization of the rate all around the value 82% despite the fact that we continued to increase the cost with high values. This is due to our opinion for the following reason: it is known that for high values of C, the optimization will choose a hyper-plane with a smaller margin, conversely, a very small value of C will cause the optimization to seek a separation hyper-plane with a larger margin. So in this case the two classes are very close to each other, then the separation margin is small and that is found with the value 150, after this value there is no data in the margin..

Influence of ϵ: Figure 11 represents the evolution of accuracy depending on ϵ: by testing on the training data and using the RBF kernel of the LIBSVM method in WEKA:
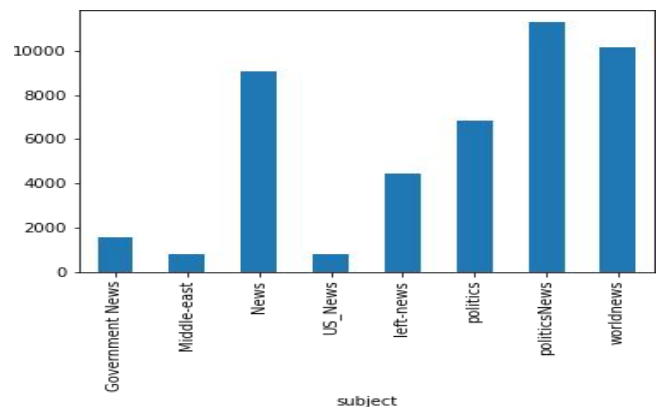

Figure 11. Distribution of News Articles by Subject

In Figure 11 we observe a stabilization of the rate around 82% up to the value 0.1 then a slight drop in the rate which can be neglected up to the value 1. This shows that the ϵ parameter does not have a great influence on the rate of recognition. Which is very logical because this parameter determines the tolerance of the termination criterion. That's the allowed error rate that's all.
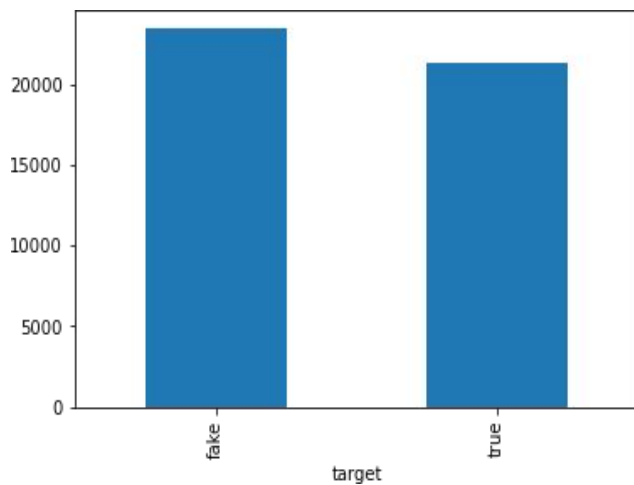
Figure 12. Distribution of Fake and Real News Articles

In Figure 12 The image shows a bar chart illustrating the *distribution of fake and real news articles. On the x-axis, two categories are labeled: **"fake"** and *"true"*, representing fake news articles and real news articles, respectively. The y-axis shows the number of articles, ranging from 0 to above 20,000. From the bars, we can see that there are slightly more fake news articles than real ones in the dataset being analyzed.
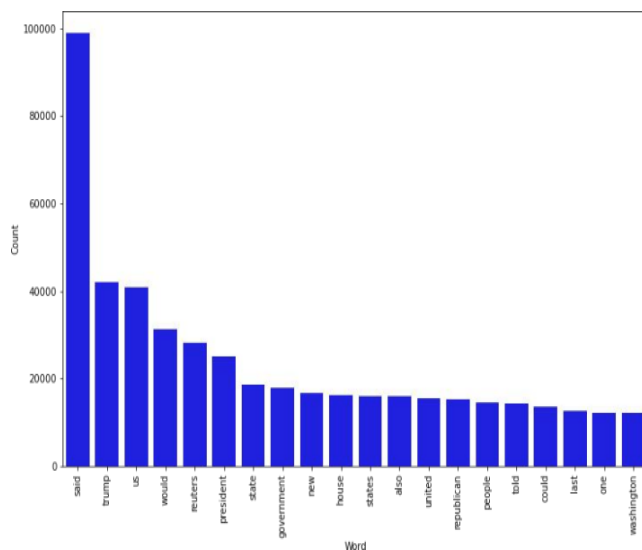

Figure 13. Word Cloud of Real News Articles

In Figure 13 word corresponds to how often it appears—the more frequently a word is mentioned, the larger and bolder it appears in the cloud. This tool is often used to quickly identify the main themes, topics, or issues being discussed in the media. For example, during a major global event like an election or a natural disaster, words related to those events (such as "vote," "candidate," "storm," or "relief") might dominate the cloud.

By analyzing a word cloud generated from real news articles, viewers can get an immediate sense of what topics are capturing public attention without having to read each article individually. It also helps researchers, journalists, and readers detect patterns, biases, or emerging trends in news coverage over time.
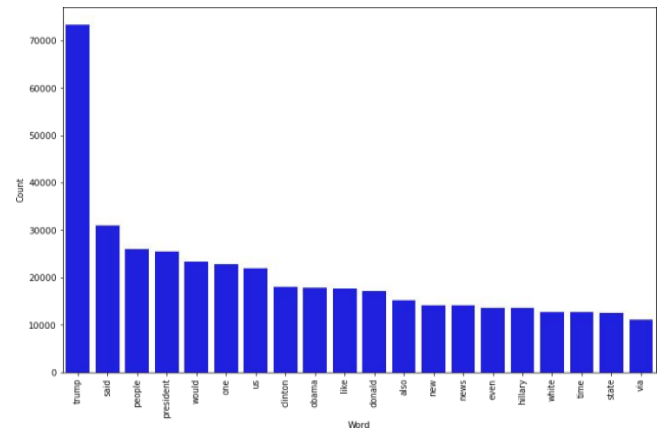

Figure14. Word Cloud of Fake News Articles

In Figure 14 A word cloud of a fake news word is a representation of the most frequently occurring words in fake news articles. Common words are displayed in larger sizes and tend to include sensational or emotional words such as "shocking," "secret," "miracle," and "hoax." Word clouds assist in unmasking patterns in fake news words, making it simpler to identify overstatement and how misinformation is disseminated.
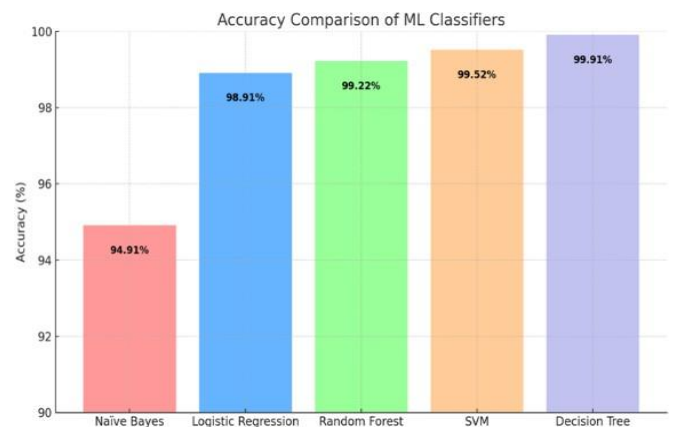

Figure 15: Accuracy comparison of classifiers used in fake news detection

In Figure 15 **The** accuracy comparison of classifiers used in fake news **detection** involves evaluating how well different machine learning models can correctly identify fake versus real news. Classifiers such as Logistic Regression, Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks are commonly tested for this task. Each classifier processes features extracted from the news articles, such as word usage, writing style.

CONCLUSION

This paper introduces a technique to identify fake news through support vector machine, attempting to find out the most effective features and methods to identify fake news. We began by researching the area of fake news, its effects and how it is identified. We then came up with and applied a solution that employs a preprocessed dataset of news by applying cleaning techniques, steaming, N-gram encoding, bag of words and TF-IDF to obtain a set of features enabling detection of fake news. We then applied Support Vector Machine algorithm on our dataset of features to develop a model enabling classification of the new information.

By the investigation conducted in this research, we achieved the following outcomes:

- the most effective features to identify fake news are in the following order: text, author, source, date and sentiment.

- the process followed yielded a recognition rate of 100%.

- the examination of the sentiment expressed by the text is interesting, but it would be more powerful in the event of opinion mining.

- the N-gram approach provides a better outcome compared to the bag of words with large datasets and with long texts.

- support vector machine appears to be the most appropriate algorithm to identify false news, since it provided a better rate of recognition, and enabled to provide for each piece of information a level of confidence for its categorization.

- parameters that affect the support vector machine are in the following order: Cost C, gamma ц and epsilon ϵ.

The work we have done can be finished and continued in other ways. It would be pertinent to continue this study with a bigger dataset, and to develop its supervised learning by another online for continuous updating and automatic incorporation of new fake news.

REFERENCES

[1] Hadeer Ahmed, Issa Traore, and Sherif Saad. Detection of online fake news using n-gram analysis and machine learning techniques. In Inter- national Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pages 127–138. Springer, 2017.

[2] Chih-Chung Chang and Chih-Jen Lin. LIBSVM – A Library for Support Vector Machines, July 15, 2018.

[3] Niall J Conroy, Victoria L Rubin, and Yimin Chen. Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology, 52(1):1–4, 2015.

[4] Chris Faloutsos. Access methods for text. ACM Computing Surveys (CSUR), 17(1):49–74, 1985.

[5] Mykhailo Granik and Volodymyr Mesyura. Fake news detection using naive bayes classifier. In 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pages 900–903. IEEE, 2017.

[6] Kaggle. Getting Real about Fake News, 2016.

[7] Kaggle. All the news, 2017.

[8] Junaed Younus Khan, Md Khondaker, Tawkat Islam, Anindya Iqbal, and Sadia Afroz. A benchmark study on machine learning methods for fake news detection. arXiv preprint arXiv:1905.04749, 2019.

[9] Cédric Maigrot, Ewa Kijak, and Vincent Claveau. Fusion par apprentis- sage pour la détection de fausses informations dans les réseaux sociaux. Document numerique, 21(3):55–80, 2018.

[10] Refaeilzadeh Payam, Tang Lei, and Liu Huan. Cross-validation. Ency- clopedia of database systems, pages 532–538, 2009.

[11] Cristina M Pulido, Laura Ruiz-Eugenio, Gisela Redondo-Sama, and Beatriz Villarejo-Carballido. A new application of social impact in social media for overcoming fake news in health. International journal of environmental research and public health, 17(7):2430, 2020.

[17] William Yang Wang. " liar, liar pants on fire": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648, 2017.

[18] Lechevallier Y. WEKA, a free data mining and learning software. INRIA-Rocquencourt.