Pattern Recognition - Final Exam Report

Fake News Classification for SDG 16
(Peace, Justice, and Strong Institutions)

Rabbani Nur Kumoro
(21/472599/PA/20310)

DEPARTMENT OF COMPUTER SCIENCE AND ELECTRONICS

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

UNIVERSITAS GADJAH MADA

YOGYAKARTA

2023

I. Introduction

The proliferation of fake news in today's digital age has raised concerns about the dissemination of misinformation and its impact on society. With the advent of social media and the ease of content sharing, misinformation can quickly reach a wide audience, leading to potential social, political, and economic consequences [**1**]. Fake news has the potential to manipulate public opinion, undermine trust in institutions, and disrupt democratic processes [**2**]. Recognizing the scope of this problem, this project aligns with Sustainable Development Goal 16 (SDG 16) - Peace, Justice, and Strong Institutions, which emphasizes the importance of promoting transparent and accountable societies. SDG 16 aims to promote peaceful and inclusive societies for sustainable development, provide access to justice for all, and build effective, accountable, and inclusive institutions at all levels [3]. However, this SDG has attracted relatively limited public attention, with noticeable changes only becoming evident quite recently. This is reflected by the relative scarcity of recent news on the topic shown in **Figure 1** [**4**].

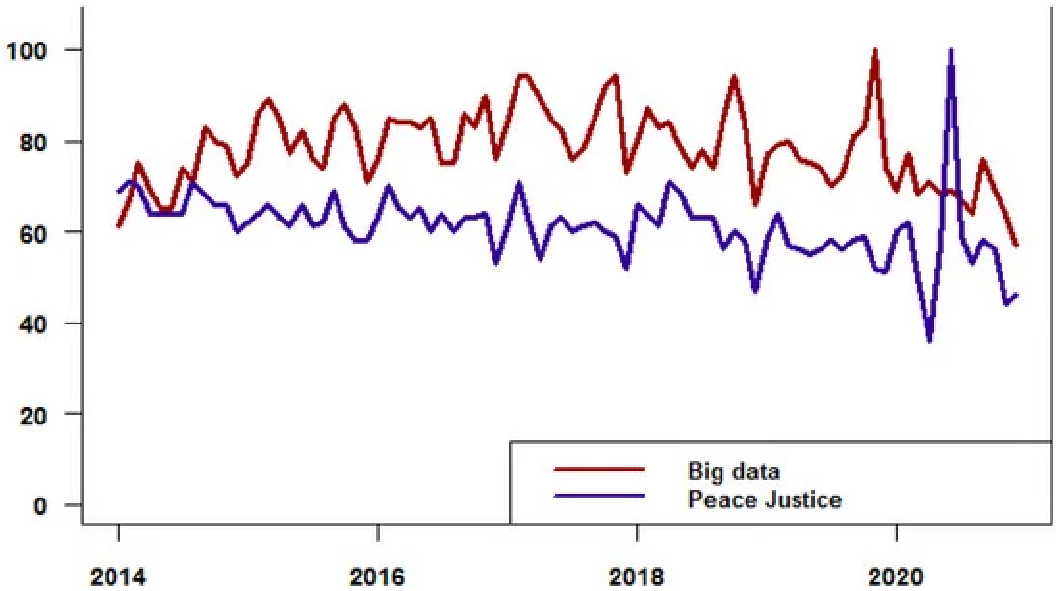


Figure 1. SDG 16 and Big Data Global Google Trends from 2014

The importance of tackling the issue of fake news lies in its potential to undermine the principles of SDG 16 and jeopardize social stability. Fake news can erode trust in institutions, breed misinformation, and hinder the establishment of justice and accountable systems. The proliferation of fake news has emerged as a critical concern, particularly during impactful events such as the presidential elections and the COVID-19 pandemic, posing significant risks to societies worldwide [**5**]. The World Economic Forum (WEF) has recognized the dissemination of fake news and misinformation online as one of the top threats to society [**6**]. Instances of fake news influencing public opinion, hindering public health responses, and eroding trust in democratic institutions have been observed globally. For instance, during the COVID-19 pandemic, false information spread through social media platforms impeded effective public health measures [**7**]. In

Indonesia, fake news has fueled social unrest and influenced election outcomes. Notable events, such as the 2016 United States presidential election, have demonstrated how the spread of fake news can influence public opinion and create divisions within society [**8**]. These examples underscore the urgent need for media literacy and critical thinking to counteract the detrimental effects of fake news on public health and democratic processes. Therefore, by developing a robust classification model, this project seeks to contribute to the broader efforts of promoting media literacy, strengthening democratic institutions, and combatting the harmful effects of misinformation.

The primary objective of this project is to create an effective and reliable classification model that can contribute to the fight against fake news. By accurately identifying and categorizing fake news articles, the project's goal is to empower individuals, organizations, and institutions with the ability to make informed decisions based on trustworthy information. Furthermore, by safeguarding the integrity of news and information sources, the project aims to promote the principles of justice, ensure peace, and strengthen democratic processes.

To undertake this project, a comprehensive dataset consisting of 18,285 articles has been collected from Kaggle through the University of Tennessee, Knoxville Machine Learning Club [**9**]. The dataset primarily focuses on the 2016 United States politics, which is a domain highly susceptible to fake news. The data is structured in a tabular format, with each article associated with specific features such as title, author, text content, and a label indicating whether it is real or fake news. However, it is important to note that the dataset contains noise and null values, necessitating thorough preprocessing steps to ensure data quality and reliability. These preprocessing steps will involve handling missing values, cleaning the text data, and transforming it into a suitable format for subsequent analysis and classification tasks.

By adopting a pattern recognition approach and NLP techniques, the outcomes of this project can have implications for media literacy initiatives, the development of robust fact-checking mechanisms, and the promotion of public awareness regarding the credibility of news sources. By combatting misinformation and advancing the principles of SDG 16, this project can support and foster peace, justice, and strong institutions in the face of the challenges posed by fake news.

II. Related Work

Fake news is a pervasive issue rooted in the dissemination of false information across the internet, encompassing non-existent news, hoaxes, and sensationalized articles through social media platforms. Detecting and labeling fake news remains a challenge as researchers explore various approaches. While previous studies have delved into machine learning and deep learning, research specifically focusing on sentiment analysis and sentiment information is scarce. Ahmed et al. [**10**] utilized term frequency and TF-IDF models to extract fake content but found that nonlinear machine learning models performed less effectively compared to linear models, especially for higher n-gram applications. Bondielli and Marcelloni [**11**] examined different methods for detecting fake news, highlighting the difficulty in gathering relevant and accurately written information. However, their study faced limitations in terms of the assortment of important information and the lower performance of machine learning models.

In the realm of NLP and ML, Bali et al. [**12**] conducted a study on fake news detection using three representative datasets, and their findings showed that gradient boosting outperformed other classifiers. However, the accuracy and F1 scores of seven alternative machine learning algorithms remained below 90%. Shaikh and Patil [**13**] employed TF-IDF features to detect fake news resources but were limited by the size of their datasets. Their study achieved 95% accuracy using the passive-aggressive classifier and SVM model, despite the minimal dataset samples available. Additionally, Conroy et al. [**14**] presented an overview of strategies for identifying fake news, specifically focusing on linguistic methodologies and network approaches. These approaches involved analyzing language patterns and utilizing network data to measure deception. There is potential for innovative hybrid methods that combine semantic signals, artificial intelligence, and network-based social information to address the challenges of fake news detection.

Pattern recognition techniques have demonstrated their effectiveness in identifying fake news by detecting distinct patterns and characteristics in fake news articles. A widely employed technique, TF-IDF, is used to extract and represent textual features that capture the unique attributes of fake news [**15**, **16**]. TF-IDF assigns weights to terms based on their frequency and rarity, allowing for the identification of informative and discriminative features that can indicate the presence of fake news. This approach provides valuable insights into the linguistic cues and content characteristics that set fake news apart from genuine news. In addition to pattern recognition techniques, machine learning algorithms such as Decision Trees and Naïve Bayes have been extensively utilized for the classification of fake news [**17**, **18**]. These models leverage the extracted features to construct accurate classifiers that effectively differentiate between fake and real news articles.

III. Methodologies

In this project, a range of methodologies were employed, as illustrated in **Figure 2**. The steps include acquiring the dataset from Kaggle, data pre-processing, exploratory data analysis (EDA), feature extraction, modeling, and classification report. The process started with the dataset, which focused on US politics, obtained from Kaggle and it require pre-processing to handle issues such as noise and null values. EDA was conducted to gain insights into the data, while feature extraction techniques like feature selection and TF-IDF were utilized to identify informative features. Machine learning models such as Decision Trees and Naïve Bayes were trained and evaluated for the classification of fake news. The performance of these models was assessed using metrics such as accuracy, precision, recall, and F1-score, and the results were summarized in a comprehensive classification report.

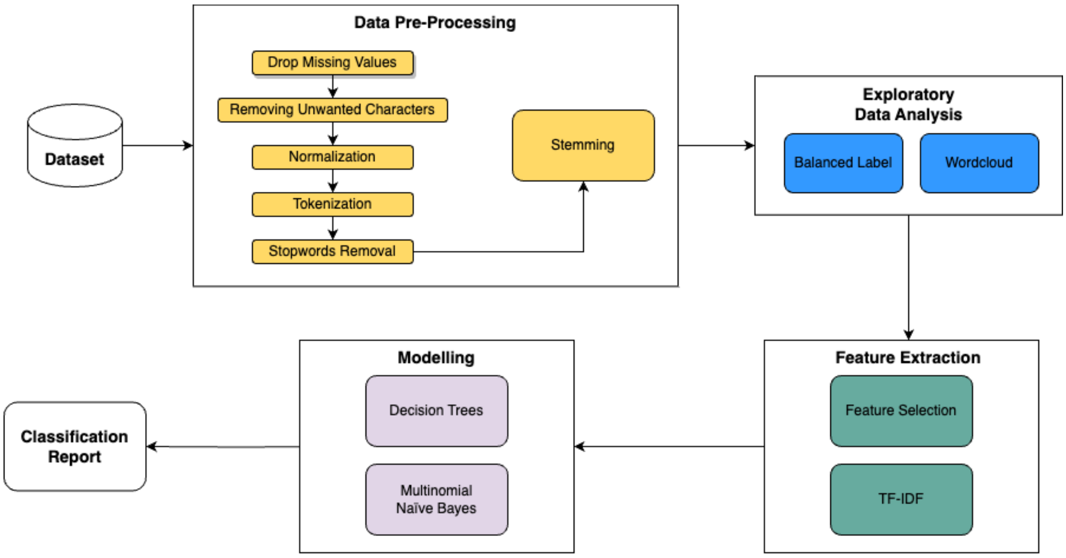


Figure 2. Flowchart Diagram

a. Dataset

The dataset used in this project contains news articles with key features such as unique identifiers ("id"), article titles ("title"), authors ("author"), and the article content ("text"), which will be leveraged to develop a classification model. **Figure 3** illustrates the four essential features of the dataset. Starting from the "id" which serves as a unique identifier for each news article. The "title" represents the article's headline, providing a concise summary or description of its content. The "author" feature specifies the individual responsible for creating the article, which could be a journalist or a corporate news company. Lastly, the "text" feature contains the actual body of the article, providing a comprehensive view of the information being conveyed. Additionally, each article is assigned a "label" indicating its reliability, with a value of 0 denoting a reliable source and 1 representing an unreliable source.

	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

Figure 3. Example of the First Five Rows of the Dataset

b. Data Pre-Processing

In the pre-processing stage, several methods were applied to clean and prepare the text data. Missing values were dropped to ensure data integrity. Unwanted characters were removed to eliminate noise and enhance the quality of the text. Normalization techniques were employed to standardize the text, such as converting all text to lowercase. Tokenization was performed to break down the text into individual words or tokens. Stop words, common words with little semantic value were removed to focus on more meaningful content. Lastly, stemming was applied to reduce words to their root form, enabling better analysis of word frequency and patterns. These pre-processing methods were crucial in transforming the raw text data into a more structured and suitable format for subsequent analysis and modeling.

id	0
title	558
author	1957
text	39
label	0

(a)

id	0
title	0
author	0
text	0
label	0

(b)

Figure 4. (a.) Before Cleaning; **(b)** After Cleaning

The missing values in the dataset were handled through a data cleaning process, as depicted in **Figure 4**. The initial dataset contained missing values in the columns for "title," "author," and "text." These missing values were addressed by dropping the corresponding instances from the dataset. As a result, the cleaned dataset showed no missing values, as indicated by the value "0" for each respective column.

To ensure data cleanliness and reduce noise, the text undergoes a comprehensive cleaning process. This involves removing unwanted elements such as capital letters, special characters, digits, URLs, and excessive whitespace. By eliminating these unnecessary components, the data becomes more refined and ready for focused analysis. The subsequent step is normalization, which standardizes the entire corpus. This includes converting all text to lowercase, ensuring a consistent representation throughout the dataset, and facilitating further analysis. Moving forward, the text is tokenized, where it is systematically divided into individual tokens or words. This essential step establishes a solid foundation for subsequent analysis, enabling the extraction of deeper insights from the textual data.

To enhance the quality of the text data, the elimination of stop words becomes crucial. These frequently used words, such as "and," "the," and "is," hold little significance in sentiment analysis. By discarding these stop words, the dataset's dimensionality is reduced, enabling a focus on more meaningful and contextually relevant words. Additionally, stemming, the subsequent step, is employed to further refine the remaining words. Stemming involves reducing words to their base or root form by removing prefixes and suffixes. This standardization process captures the essence of words, facilitating the discovery of underlying patterns and a more comprehensive understanding of the conveyed sentiment. Finally, the stemmed words are reassembled into sentences, preserving the contextual information for improved analysis, as depicted in **Figure 5**.

house dem aide we didnt even see comeys letter...
flynn hillary clinton big woman campus breitba...
why truth might get you fired consortiumnewsco...
15 civilians killed in single us airstrike hav...
iranian woman jailed fictional unpublished sto...

Figure 5. Stemming Results

c. Exploratory Data Analysis (EDA)

In the EDA conducted for this project, the dataset was carefully balanced to include a proportional representation of fake and real news labels. This ensured a fair distribution of data for both categories. To gain insights into the textual content, word cloud visualizations were employed to showcase the most frequent words associated with real and fake news articles. These word clouds provided a comprehensive overview of the key terms and themes prevalent in each category, helping to identify distinct patterns and significant features that differentiate between real and fake news.

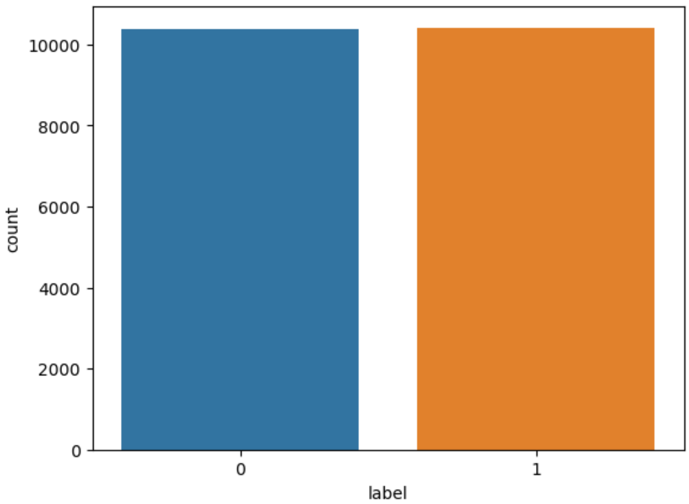
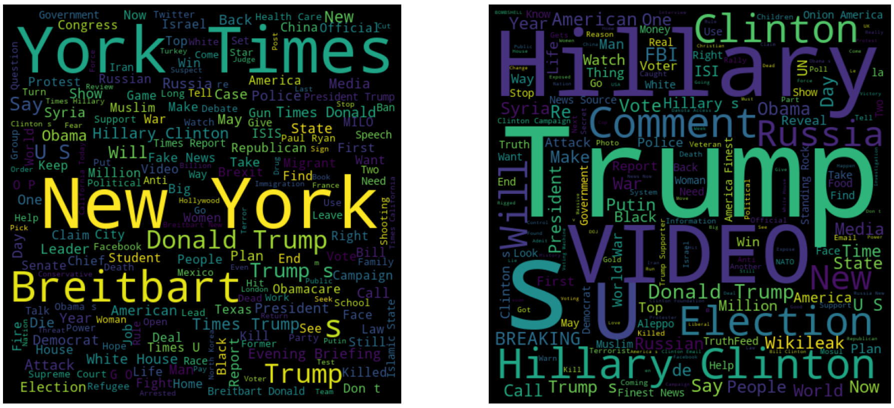


Figure 6. Real and Fake News Label

In the phase of the EDA process, the distribution of labeled real and fake news was analyzed, as depicted in the bar chart **Figure 6**. The chart reveals a balanced distribution between the two classes, indicating that no additional class balancing techniques are required. This finding is crucial, as it ensures a fair representation of both types of news in the subsequent analyses.

Subsequently, the word cloud we generate will illustrate the most commonly occurring words in a given text or collection of texts. These words will be visually arranged in a cloud-like pattern, where the size of each word corresponds to its frequency of occurrence in the text. Less frequently used terms will appear in smaller font sizes, while the most frequently used words will be presented in larger font sizes. Word clouds serve as effective tools for summarizing and presenting large volumes of text data in text analysis and data visualization. They facilitate the identification of frequently used words and phrases, as well as quickly identifying key ideas or topics within a document.



(c) **(d)**

Figure 7. (c) Frequent Real Words; **(d)** Frequent Fake Words

The word clouds were plotted for the fake and real news samples, revealing the essential terms in the datasets. The word clouds in **Figure 7 (c)** showcased the keywords for real news, with phrases like "Breitbart", "New York", and "York Times"standing out. These findings suggest that Breitbart News and New York Times are prominent and reliable news sources. In **Figure 7 (d)**, the word cloud for fake news samples emphasized terms such as "Trump" and "Hillary" from the article titles, indicating a potential connection with fake news articles. The word cloud analysis highlighted political figures as the most frequently mentioned in fake news.

d. Feature Extraction

In the feature extraction process, two features were combined to create a single feature for analysis. One of the key techniques employed was TF-IDF (Term Frequency-Inverse Document Frequency). TF-IDF assigns weights to terms based on their frequency in a document and their rarity across the entire dataset. This approach allows for the identification of informative and discriminating features that can distinguish between different types of news articles. By combining features and utilizing TF-IDF, the feature extraction phase aimed to capture relevant and meaningful information that would contribute to the subsequent analysis and classification of news articles.

	content
id	
0	Darrell Lucus house dem aide we didnt even see comeys letter...
1	Daniel J. Flynn flynn hillary clinton big woman campus breitba...
2	Consortiumnews.com why truth might get you fired consortiumnewsco...
3	Jessica Purkiss 15 civilians killed in single us airstrike hav...
4	Howard Portnoy iranian woman jailed fictional unpublished sto...

Figure 8. New Feature named ‘content’

To enhance the feature representation, a combination of two features, namely 'author' and 'title', was performed by merging them into a single feature called 'content', as illustrated in **Figure 8**. By merging these features, the resulting 'content' feature incorporates information from both the author and title fields, providing a more comprehensive representation of the news articles. This combined feature allows for a holistic analysis of the textual content, capturing relevant information that may contribute to the subsequent analysis and classification tasks. By consolidating these two features into one, the feature space becomes enriched, potentially leading to improved performance in fake news classification.

TF-IDF (Term Frequency-Inverse Document Frequency) stands as a powerful feature in the domain of fake news classification. This feature plays a crucial role by filtering out commonly occurring words and assigning lower weights to such terms, enabling a focus on more meaningful and discriminatory terms. By giving higher weights to important yet relatively rare terms, TF-IDF effectively captures the significance of terms within each document, aiding in the identification of potentially fake news content [**2**].

$$\mathrm{TF}\,(t,d) = \frac{\textit{Number of times t occurs in a document 'd'}}{\textit{Total word count of document 'd'}}$$

$$\mathrm{IDF}\,(t) = \log_e\left(\frac{\textit{Total number of documents}}{\textit{Number of documents with term t in it}}\right)$$

$$\mathrm{TF\text{-}IDF}\,(t,d) = \mathrm{TF}\,(t,d) * \mathrm{IDF}\,(t)$$

Figure 9. The formulas for Calculating TF-IDF

The method of TF-IDF vectors represents a term's relative significance in the record or as a whole. The next factor of this method is term frequency (TF), which represents the frequency of a word occurring in the dataset. Another important aspect to ensure the proper functioning of the model is IDF, standing for inverse document frequency, used to measure the prominence of a term in the entire dataset. Additionally, the TF-IDF integrates the inverse document frequency with the term frequency [**19**]. The formulas for calculating TF, IDF, and TF-IDF can be seen in Figure 9.

Moreover, TF-IDF contributes to dimensionality reduction as it represents documents using sparse vectors of TF-IDF values, depicted in **Figure 10**. This dimensionality reduction not only improves computational efficiency but also enhances feature selection, enabling the identification of the most informative and distinguishing terms. Another advantage of TF-IDF is its flexibility, making it compatible with various machine learning libraries and easily integrated into the overall pipeline of fake news classification tasks. These qualities make TF-IDF a valuable tool for extracting relevant and discriminative features in the fight against fake news.

```
(0, 19097)    0.27315635150958634
(0, 16473)    0.23676064517956455
(0, 11072)    0.33384522056560495
(0, 10747)    0.26822209263186303
(0, 9692)     0.22757176689298134
(0, 8832)     0.20534182453318398
(0, 6433)     0.21422587910261737
(0, 5256)     0.27468869329117757
(0, 4995)     0.2512923264945339
(0, 4763)     0.33044571153796654
(0, 3952)     0.2266469969205269
(0, 3403)     0.33756896138985654
(0, 809)      0.3646500188253278
(1, 20416)    0.29951908908156866
(1, 8608)     0.19815023888659125
(1, 7101)     0.711483310803025
(1, 4728)     0.26268668599849243
(1, 3778)     0.19062686807106288
(1, 3100)     0.3870784468942128
(1, 2713)     0.15460118725006144
(1, 2258)     0.2928176012009572
(2, 19015)    0.41491113753784553
(2, 11878)    0.49151393723208897
(2, 7650)     0.34605253138342823
(2, 6968)     0.39293503470255664
```

Figure 10. TF-IDF Weight Results

e. Modelling

In the modeling phase, a range of machine learning algorithms will be utilized to classify and differentiate between real and fake news articles. One such algorithm that will be explored is Naïve Bayes, a popular probabilistic classifier that works based on the assumption of independence between features. Naïve Bayes has been widely used in text classification tasks and has shown promising results in distinguishing between genuine and deceptive news content [**20**]. Additionally, Decision Trees, another powerful algorithm, will be employed. Decision Trees create a hierarchical structure of decision rules based on the features of the dataset, making them interpretable and intuitive for understanding the classification process [**21**]. By evaluating the performance of these algorithms and potentially other classifiers, insights will be gained into their effectiveness in accurately identifying fake news and contributing to the development of a reliable and robust fake news model.

The Naïve Bayes classifier is a widely utilized machine learning algorithm that is particularly effective for solving binary classification and multi-class classification problems, as it simplifies the estimation of probabilities for each possibility, making the measurements more manageable [**2**]. However, it is important to note that the model has a couple of disadvantages. Firstly, it assumes that the variables are independent of each other. Secondly, the order in which variables are presented does not influence the classification decision. Naïve Bayes is an appealing choice for fake news classification due to its simplicity, computational efficiency, and ability to handle high-dimensional feature spaces and large datasets [**22**]. Despite its assumption of independence, which is often reasonable for text classification tasks, Naïve Bayes demonstrates effectiveness in detecting fake news [**20**].

Naïve Bayes utilizes the TF-IDF features extracted from the content of the news to calculate conditional probabilities, determining the likelihood of a news article belonging to the real or fake class. By considering the probability of observing specific feature values of the TF-IDF weights given the class label, along with the prior probabilities of the classes, Naïve Bayes makes predictions and classifies new, unseen articles as either real or fake. By assuming conditional independence of TF-IDF features given the class label, Naïve Bayes calculates the conditional probabilities for each term in a news article. These probabilities are then used to make predictions based on the joint probabilities of the observed features. The algorithm computes the likelihood of a news article falling into the real or fake class and assigns the predicted label based on the class with the highest probability. The visualization presented in **Figure 11** offers a visual depiction of the separation between the two classes based on the calculated probabilities. In this illustration, the blue dots represent fake news articles, while the red dots represent real news articles. Naïve Bayes employs probabilistic calculations to classify news articles into their respective classes.

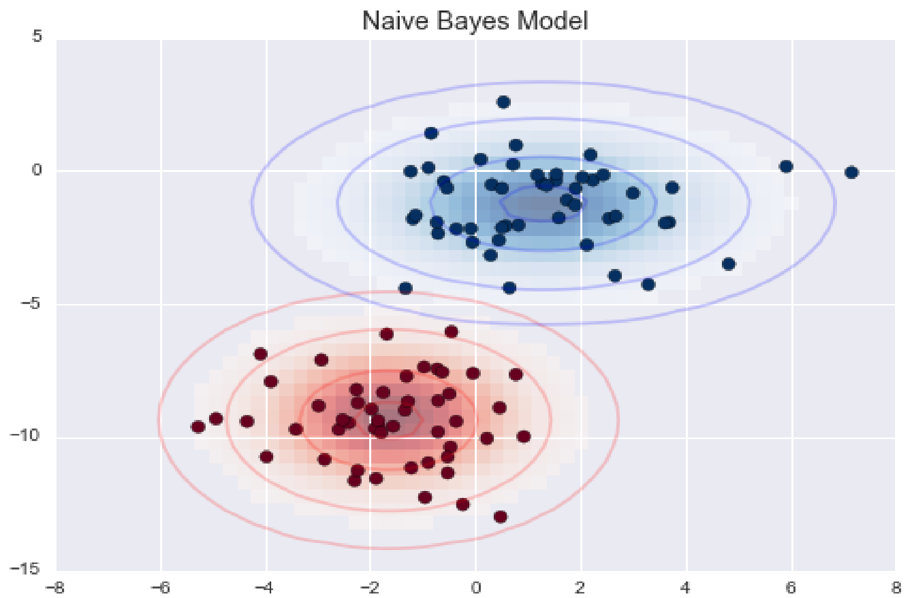


Figure 11. Naïve Bayes Visualization

The Decision Trees is a widely used supervised learning model that is applicable for both classification and regression tasks [**16**]. It is a powerful nonparametric approach that partitions the dataset into subsets based on attribute value tests. This recursive process continues until all the nodes within a subset share the same variable, resulting in the creation of decision nodes and leaf nodes. Decision nodes represent categorization or decision points, while leaf nodes represent the final outcomes. The algorithm is capable of handling both categorical and numerical data, making it versatile in various domains [**23**].

Decision Trees can be effectively utilized for fake news classification by leveraging TF-IDF features. Decision Trees are a popular machine learning algorithm that can handle both categorical and numerical data, making them suitable for processing text-based features derived from TF-IDF representation [**24**]. One advantage of Decision Trees is their ability to capture complex relationships between features. By recursively splitting the dataset based on different feature values, Decision Trees can identify informative thresholds that separate real and fake news articles effectively. Another benefit of Decision Trees is their inherent feature importance estimation. Feature importance analysis performed by Decision Trees helps to gain insights into the linguistic cues and indicators of fake news by assigning higher importance to discriminative features. This analysis reveals which specific TF-IDF features contribute significantly to the classification task. Moreover, Decision Trees excel in handling non-linear relationships between features and the target variable, which is valuable in fake news classification [**25**].

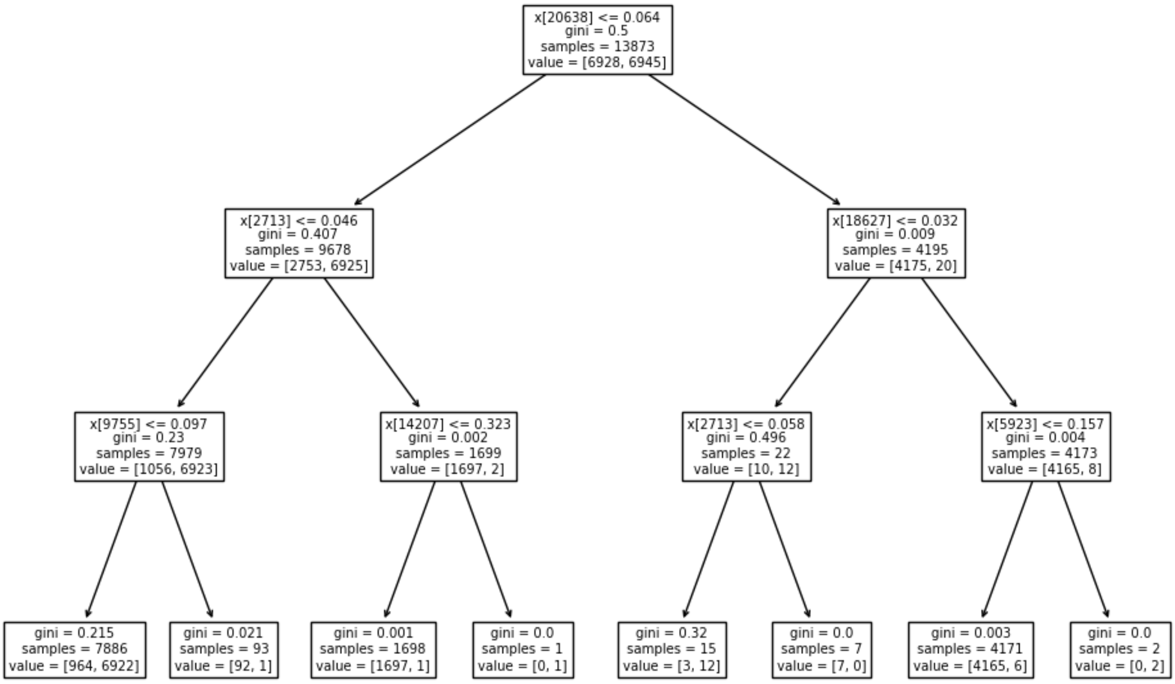


Figure 12. Decision Trees Visualization

In this context, where certain combinations of words or phrases strongly suggest fake news, Decision Trees can capture these non-linear relationships and accurately predict based on the given TF-IDF features. The provided visualization in **Figure 12** showcases the application of Decision Trees in classifying fake news. It is worth noting that the depicted Decision Trees have a limited height of three levels for clarity and visualization purposes, but in practice, Decision Trees can be much deeper, capturing a wider range of features and patterns in the data.

f. Performance Metrics Evaluation

The classification report provides an overview of the model's performance, assessing key metrics such as accuracy, precision, recall, and F1 score. These metrics are crucial in evaluating the effectiveness of the classification model in distinguishing between real and fake news [**1**]. Accuracy measures the overall correctness of the predictions, while precision quantifies the proportion of correctly classified instances among the predicted positive cases. Recall, also known as sensitivity or true positive rate, evaluates the model's ability to correctly identify positive instances. The F1 score is a harmonic mean of precision and recall, offering a balanced measure that considers both metrics [**26**]. By examining the classification report, one can gain insights into the model's performance and make informed decisions about its reliability and effectiveness in the context of fake news classification.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\text{-}score = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

Figure 12. Evaluation Metrics Formula

In order to compute the evaluation metrics, it is necessary to consider the counts of true positives (*TP*), true negatives (*TN*), false positives (*FP*), and false negatives (*FN*) in relation to the ground truth. A true positive is recorded for each article that is correctly predicted as fake news and is also annotated as fake in the test set. Conversely, a false positive is tallied for each article predicted as fake news but annotated as true news in the test set. Similarly, a true negative is counted for each article predicted as true news and annotated as true in the test set, while a false negative is noted for each article predicted as true news but annotated as fake in the test set [**1**].

Figure 12 provides the formulas for each evaluation metric. Accuracy represents the percentage of articles that have been correctly labeled, whether as fake or true news. Precision signifies the percentage of predicted fake news articles that correspond to real fake news. Recall indicates the percentage of total fake news items in the test set that are successfully recognized as fake. F1 score combines precision and recall by taking their harmonic mean. Higher values for all these measures indicate better performance of the classification system.

IV. Experimental Results

The outcomes of the model were obtained by applying it to the test dataset, and a comprehensive analysis of the results was conducted. Evaluation metrics such as accuracy, precision, recall, and F1 score were utilized to assess the model's predictive abilities and overall performance. Additionally, the confusion matrix provided a detailed representation of the model's classification results, allowing for a deeper understanding of its ability to accurately classify instances of real and fake news. By considering these metrics and analyzing the confusion matrix, a thorough assessment of the model's performance was achieved, providing valuable insights for further improvements and optimizations.

The classification results of the Decision Trees and Naïve Bayes algorithms are presented in **Table 1**. The Decision Tree algorithm achieved an impressive accuracy of 99.32% and exhibited high precision, recall, and F1 score values of 0.99. These findings indicate its strong capability to accurately classify both real and fake news articles. In contrast, the Naïve Bayes algorithm achieved an accuracy of 95.52% with precision, recall, and F1 score values of 0.95, indicating its effectiveness in distinguishing between genuine and fake news. Overall, the Decision Tree algorithm surpassed Naïve Bayes in performance, suggesting its potential as an effective approach for the precise classification of fake news.

Table 1. Classification Report

Models	Accuracy	Precision	Recall	F1-Score
Decision Trees	**99.32%**	**0.99**	**0.99**	**0.99**
Naïve Bayes	95.52%	0.96	0.96	0.96

The confusion matrices in **Figure 13** and **Figure 14** provide a visual representation of the algorithm's performances. Upon analyzing these matrices, excellent overall performance was observed for both algorithms. However, a notable difference emerged in terms of false negatives. The Decision Tree algorithm displayed significantly fewer false negatives compared to Naïve Bayes, achieving approximately 8 times better performance in this aspect. This indicates that the Decision Tree algorithm excelled in correctly identifying instances of fake news that were initially misclassified by Naïve Bayes. Based on these findings, both algorithms can be considered successful in effectively classifying fake news. Nevertheless, the Decision Tree algorithm holds a significant advantage in reducing false negatives, highlighting its superior capability in accurately identifying deceptive information.

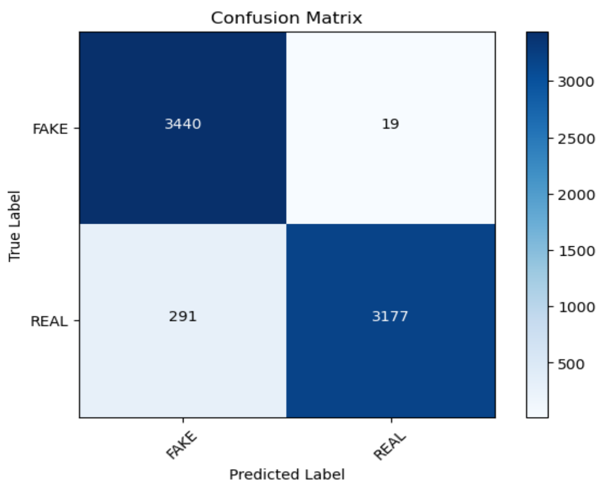


Figure 13. Naïve Bayes Confusion Matrix

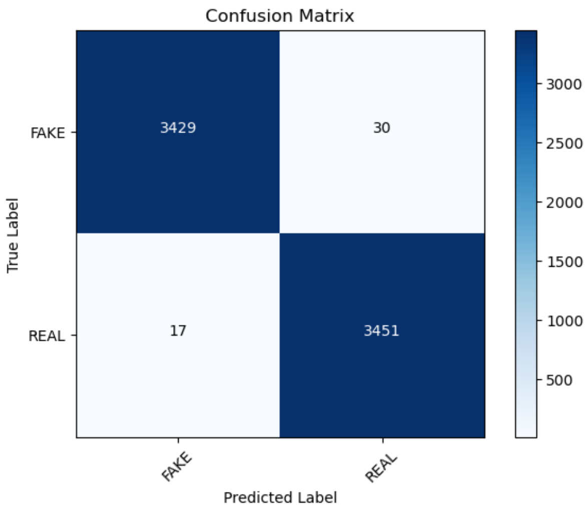


Figure 14. Decision Trees Confusion Matrix

V. Discussion and Analysis

Naïve Bayes, while simple and efficient, can struggle with predicting false negatives due to its assumption of conditional independence among features. In the context of fake news classification, where feature dependencies and interactions are prevalent, this assumption can lead to higher false negatives. Additionally, Naïve Bayes assigns equal weight to all features, potentially overlooking strong discriminatory features crucial for accurate classification.

On the other hand, Decision Trees have the advantage of explicitly capturing discriminatory features during the learning process. They can handle complex relationships and provide feature importance, making them suitable for analyzing fake news articles. [**1**] However, it is important to consider the trade-offs between model performance and interpretability when choosing the appropriate algorithm for fake news classification tasks. Both Naïve Bayes and Decision Trees have their strengths and weaknesses, the selection should be based on the specific requirements and characteristics of the fake news problem at hand.

VI. Conclusion and Future Works

In summary, the TF-IDF feature extraction technique has played a significant role in the analysis of fake news classification by capturing important characteristics of news articles and representing them as weighted terms. This technique allows for the identification of real and fake features of the news content. It is important to note, however, that the superiority of Decision Trees over Naïve Bayes is not absolute and depends on various cases. Decision Trees have shown higher accuracy, with 99% compared to Naïve Bayes' 95%.

In the future, there are several avenues for further improvement and exploration in the classification of fake news. Firstly, obtaining a dataset that is more sequential and can be continuously updated with recent news would greatly enhance the model's ability to handle real-time scenarios. This would involve establishing a pipeline that can collect and process news articles as they are published, allowing for timely analysis and adaptation to emerging trends and techniques used by misinformation spreaders. Additionally, it would be beneficial to extend the scope of the classification beyond textual content alone. Fake news can be disseminated through various mediums such as audio clips, images, and videos. By incorporating these different content types into the dataset and developing computational models capable of detecting fake news across multiple modalities, we can improve the model's effectiveness in identifying and combating misinformation that may not be conveyed through text alone.

To further enhance the performance of the model, exploring the utilization of ensemble algorithms and deep learning approaches is essential. Ensemble algorithms, which combine multiple models, can leverage the strengths of individual models and produce more accurate and robust predictions. Furthermore, deep learning techniques like Long Short-Term Memory (LSTM) networks, known for their ability to capture temporal dependencies in sequential data, can be employed to analyze news articles over time and improve the model's understanding of the evolving nature of fake news. Incorporating advanced language models such as Bidirectional Encoder Representations from Transformers (BERT) can provide a more nuanced understanding of textual content and enhance the model's ability to capture subtle contextual cues. Furthermore, exploring techniques like LSTM sequences to sequences, which extend LSTM models to handle longer sequences, and incorporating bigrams and trigrams in traditional machine learning and neural network models can also contribute to improved performance in classifying fake news [**2**].

References

[**1**] M. J. Awan et al., “Fake news data exploration and analytics,” Electronics, vol. 10, no. 19, p. 2326, 2021. doi:10.3390/electronics10192326

[**2**] C.-M. Lai, M.-H. Chen, E. Kristiani, V. K. Verma, and C.-T. Yang, “Fake news classification based on content level features,” *Applied Sciences*, vol. 12, no. 3, p. 1116, 2022. doi:10.3390/app12031116

[**3**] “Goal 16 | Department of Economic and Social Affairs,” United Nations, https://sdgs.un.org/goals/goal16 (accessed Jun. 15, 2023).

[**4**] H. Hassani, X. Huang, S. MacFeely, and M. R. Entezarian, “Big Data and the United Nations Sustainable Development Goals (UN SDGS) at a glance,” Big Data and Cognitive Computing, vol. 5, no. 3, p. 28, 2021. doi:10.3390/bdcc5030028

[**5**] F. Fifita, J. Smith, M. B. Hanzsek-Brill, X. Li, and M. Zhou, “Machine learning-based identifications of COVID-19 fake news using biomedical information extraction,” Big Data and Cognitive Computing, vol. 7, no. 1, p. 46, 2023. doi:10.3390/bdcc7010046

[**6**] M. Del Vicario et al., “The spreading of misinformation online,” Proceedings of the National Academy of Sciences, vol. 113, no. 3, pp. 554–559, 2016. doi:10.1073/pnas.1517441113

[**7**] S. Bin Naeem and M. N. Kamel Boulos, “Covid-19 misinformation online and Health Literacy: A brief overview,” International Journal of Environmental Research and Public Health, vol. 18, no. 15, p. 8091, 2021. doi:10.3390/ijerph18158091

[**8**] F. Olan, U. Jayawickrama, E. O. Arakpogun, J. Suklan, and S. Liu, “Fake news on social media: The impact on society,” Information Systems Frontiers, 2022. doi:10.1007/s10796-022-10242-z

[**9**] “Fake news,” Kaggle, https://www.kaggle.com/competitions/fake-news/overview (accessed Jun. 15, 2023).

[**10**] H. Ahmed, I. Traore, and S. Saad, “Detecting opinion spams and fake news using text classification,” Security and Privacy, vol. 1, no. 1, 2017. doi:10.1002/spy2.9

[**11**] A. Bondielli and F. Marcelloni, “A survey on fake news and rumour detection techniques,” Information Sciences, vol. 497, pp. 38–55, 2019. doi:10.1016/j.ins.2019.05.035

[**12**] A. P. Bali, M. Fernandes, S. Choubey, and M. Goel, “Comparative performance of machine learning algorithms for fake news detection,” Communications in Computer and Information Science, pp. 420–430, 2019. doi:10.1007/978-981-13-9942-8_40

[**13**] J. Shaikh and R. Patil, “Fake news detection using machine learning,” 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), 2020. doi:10.1109/isssc50941.2020.9358890

[**14**] N. K. Conroy, V. L. Rubin, and Y. Chen, “Automatic deception detection: Methods for finding fake news,” Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, 2015. doi:10.1002/pra2.2015.145052010082

[**15**] H. Ahmed, I. Traore, and S. Saad, “Detection of online fake news using N-gram analysis and Machine Learning Techniques,” Lecture Notes in Computer Science, pp. 127–138, 2017. doi:10.1007/978-3-319-69155-8_9

[**16**] J. Alghamdi, Y. Lin, and S. Luo, “A comparative study of machine learning and Deep Learning techniques for fake news detection,” Information, vol. 13, no. 12, p. 576, 2022. doi:10.3390/info13120576

[**17**] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, “Fake news detection using machine learning approaches,” IOP Conference Series: Materials Science and Engineering, vol. 1099, no. 1, p. 012040, 2021. doi:10.1088/1757-899x/1099/1/012040

[**18**] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, “Fake news detection using machine learning approaches,” IOP Conference Series: Materials Science and Engineering, vol. 1099, no. 1, p. 012040, 2021. doi:10.1088/1757-899x/1099/1/012040

[**19**] A. K. Singh and M. Shashi, “Vectorization of text documents for identifying unifiable news articles,” International Journal of Advanced Computer Science and Applications, vol. 10, no. 7, 2019. doi:10.14569/ijacsa.2019.0100742

[**20**] H. Zhang and D. Li, “Naïve Bayes text classifier,” 2007 IEEE International Conference on Granular Computing (GRC 2007), 2007. doi:10.1109/grc.2007.40

[**21**] Y.-L. Chen, H.-W. Hu, and K. Tang, “Constructing a decision tree from data with hierarchical class labels,” Expert Systems with Applications, vol. 36, no. 3, pp. 4838–4847, 2009. doi:10.1016/j.eswa.2008.05.044

[**22**] L. Jiang, C. Li, S. Wang, and L. Zhang, “Deep feature weighting for naive Bayes and its application to text classification,” Engineering Applications of Artificial Intelligence, vol. 52, pp. 26–39, 2016. doi:10.1016/j.engappai.2016.02.002

[**23**] S. Mishra, P. Shukla, and R. Agarwal, “Analyzing machine learning enabled fake news detection techniques for diversified datasets,” Wireless Communications and Mobile Computing, vol. 2022, pp. 1–18, 2022. doi:10.1155/2022/1575365

[**24**] K. Wong, “A Language, not a Letter: Learning Statistics in R Intro R-programming,” Chapter 24: Decision trees, https://ademos.people.uic.edu/Chapter24.html (accessed Jun. 16, 2023).

[**25**] D. Rohera et al., “A taxonomy of fake news classification techniques: Survey and implementation aspects,” IEEE Access, vol. 10, pp. 30367–30394, 2022. doi:10.1109/access.2022.3159651

[**26**] D. Papakostas, G. Stavropoulos, and D. Katsaros, “Evaluation of machine learning methods for fake news detection,” Studies in Computational Intelligence, pp. 163–183, 2021. doi:10.1007/978-3-030-90087-8_8