Title

Leveraging Random Forrest for Fake News Detection on Social Media

MSc Research Project
Programme Name
MSC in AI for Business (MSCAIBUS)

Raja Muhammad Naveed
Student ID: x23117311

School of Computing
National College of Ireland

Supervisor: Devanshu Anand

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Raja Muhammad Naveed
Student ID:	x23117311
Programme:	MSC in AI for Business **Year:** 2024
Module:	MSC Research Project
Lecturer:	Devanshu Anand
Submission Due Date:	12-08-2024
Project Title:	Leveraging Random Forrest for Fake News Detection on Social Media
Word Count:	5920 words
Page Count:	20 Pages

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Raja Muhammad Naveed

Date: 12-08-2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Table of Contents

Abstract

The rise of technology and increasing access to the Internet and social media (SM) platforms have made it easier to disseminate information globally. Social media platforms have become a primary source of spreading news. Individuals, governments, and organizations are using these platforms to increase their reach to target audiences. While it is an advantage of the internet and social media, many entities are abusing the facility to spread misinformation, resulting in chaos and disorder. This research project aims to determine fake news in SM posts to enhance the reliability of online information. The research intends to use a dataset comprising real and fake news articles from Kaggle. Machine learning techniques will classify the articles.

The research process will leverage Random Forest as a primary model for this project whereas Naive Bayes & Decision Tree will serve as baseline models. The models were trained and evaluated based on precision, accuracy, recall, and F1 score. This research will highlight the effectiveness of machine learning models in determining between fake and real news content with maximum accuracy and other parameters. The results provide a foundational framework for developing reliable tools to combat misinformation and contribute toward minimizing the spread of fake news in digital media.

Keywords: Fake News, Machine Learning, Random Forest, Social Media, Decision Tree, Naive Bayes, Misinformation, Fake News Detection.

1. Introduction

False information and fake news have always been a global problem. The rise of the internet escalated the challenge tenfold. Fake news or misinformation is when someone forges and shares a piece of information to mislead the audience for their own benefit. Social media and web portals have now provided a platform where entities can easily share fake news. Some outlets on social media spread wrong information to increase readers and followers. The clickbait benefits the culprits through profits in their revenue. Examples of these clickbaits are flashy headlines. When readers get attracted and click, the brand, individual, or organization gets an increase in its overall revenue and reach (Aldwairi and Alwahedi, 2018).

Fake news usually covers different fields, and the language within such materials is designed to persuade people. Sometimes, people spread false information by portraying the actual news in a twisted way to claim or support an agenda that is not factual (Feng, Banerjee, and Choi, 2012). The advancement of social media has made the spread of information fast, extensive in reach, and inexpensive. Therefore, SM platforms have become the primary source of sharing news and information with a wide range of audiences. Although social media is different from traditional news media in many ways, many news agencies use it to spread the news theatrically while posing that it is spontaneous (Shabani and Sokhn, 2018).

The concern about fake news is rising on a global level because it is becoming an ordeal to distinguish between actual and fake news. The main reason is the language of fake news which is often similar to the language of real news. Therefore, it has become highly challenging to detect fake news.

The above challenge motivates the proposed research. The proposed research aims to impact the societal norms that individuals and even organizations are spoiling through the spread of misinformation for their personal and economic gains. The broadcast of fake news significantly impacts global society, including everyone to some extent.

A significant example of fake news and its impact is when a piece of false news was disseminated during the US presidential election in 2016 (Shabani and Sokhn, 2018). Deceitful news of such nature severely damages the reputation of the targeted entity. Afterward, it becomes a challenge to restore the victim's position. So, this research aims to highlight the severity of the fake news and propose ways to detect it through Machine-learning techniques. The research intends to use Random Forrest with an aim of 85% to 90% accuracy.

As we aim to conduct this research on a real-world scenario, we will gather the dataset from an open-source Kaggle. This research intends to answer the following question.

1.1 Research Question:

"To what extent can a Machine learning Model such as Random Forrest detect fake news with maximum accuracy on social media?"

1.2 Research Objective:

The research objective is to develop an accurate yet low latent machine-learning model to find the answers to research questions. Our questions will address the achievement of maximum accuracy with low time complexity. We have selected The Random Forest algorithm, which is best due to its accuracy in results and high efficiency. Here are the defined objectives for the research:

- Evaluate the performance of Random Forest in detecting fake news.
- Compare its effectiveness against Naive Bayes and Decision Tree models.
- Analyze the implications of using Random Forest for real-time fake news detection.

2. Related Work

Tyagi et al., (2019) addressed the issue of misinformation on social media. Their study effectively executed machine-learning approaches to distinguish fake news on Twitter. Their study used Decision Trees and Naïve Bayes to analyze the veracity of tweets. The mentioned method is linked with the current research objectives of utilizing machine learning, particularly the Random Forest algorithm. These methods are effective in distinguishing between genuine and misleading information. We can conveniently consider this paper as a significant torch to address the challenges that my research questions outline.

However, a potential limitation of their approach is its simplicity. There is a high possibility that their approach might not be sufficient to handle the complexities of modern fake news. This is particularly due to the high-dimensional and unstructured data types that exist in today's digital media landscape.

(Gupta et al., 2018), conducted research, "A Comparative Study of Spam SMS Detection using Machine Learning Classifiers" and evaluated different machine-learning algorithms for their efficiency in recognizing spam SMS. Their objective is very similar to the challenge of detecting social media fake news - the segregation of text-based data. Gupta et al., compared standard machine-learning classifiers, including Support Vector Machines and Naïve Bayes with deep learning techniques such as CNNs (Convolutional Neural Networks). Their research offers a significant understanding of precision, accuracy, and recall metrics, which are essential in assessing any text classification model.

CNNs impressively does the image processing to classify textual data. It means that these kinds of advanced models could be equally efficient when detecting fake news. Hence, this paper supports our proposed research. The findings in their paper highlight the benefits of using sophisticated machine-learning techniques to improve detection capabilities.

On the other hand, there is also a limitation in the conclusion as its application is only for SMS content. The methods and findings may not directly solve more complex forms of communication, such as SM posts or news articles. The reason is that these modes of communication significantly involve context and multimedia elements. It emphasizes the need to explore more significant methods to detect fake news.

(Fang et al., 2019) conducted a study to present a polished model to detect fake news. The model was the combination of CNN (convolutional neural networks) with a mechanism of self-multi-head to detect fake news solely based on content. The result came out to be with a high precision rate of 95.5%. The recall rate was observed at 95.6% under 5-fold cross-validation. The mechanism of self-multihead attention allowed a better understanding of the text's internal relationships. It will remarkably enhance the effectiveness of this model in identifying deceptive information. On the other hand, traditional CNNs can neglect deceptive information as they are not able to handle long-range dependencies between words. The success of this model on a public dataset highlights its relevance to the current project. The paper achieved a high accuracy with Random Forest algorithm which we have used in our approach. The existing gap that my paper would cover is the complexity of the model which can lead to high computational expenses and longer training times. Additionally, large datasets for training can limit its application in situations where we have limited data sets.

(Jang, Kim and Kim, 2019) examined the combination of Word2vec and CNNs and found its importance in categorizing relevant and irrelevant textual content. They evaluated the efficacy of the CBOW (Continuous Bag-of-Words) and Skip-gram Word2vec models. The result showed that both methods significantly and positively impact the accuracy of CNNs. This reflects CBOW as a more suitable approach for handling news articles. On the other hand, Skip-gram came out to be more effective for tweets. The relevancy to our project is because this study highlights the benefits of Word2vec in improving text classification systems. Hence, we know its importance while dealing with the complex and diverse nature of news content on social media.

However, the model of this research needs improvement when it comes to detecting subtle fake news. The advanced language of the broadcast can trick the semantic scope of Word2vec. This is because it primarily understands the context within a specified window of words.

(Altunbey Ozbay and Alatas, 2019) developed a new method to detect social media fake news by using two metaheuristic algorithms. The two algorithms they used are GWO (Grey Wolf Optimization) and SSO (Salp Swarm Optimization). The method developed robust fake news detection in three stages: data preprocessing, GWO application, and SSO algorithms. The authors then tested this model with real-world data. The study tests the efficiency of these algorithms against conventional supervised AI algorithms throughout multiple datasets. The findings indicate that GWO outperforms other methods in accuracy, precision, and recall. It proves the superiority of GWO over other conventional methods.

This study is relevant to my research as it utilizes advanced machine-learning techniques, such as Random Forest algorithm. Moreover, it also indicates innovation through the use of metaheuristic optimization algorithms such as GWO and SSO. The limitation is that these algorithms can face the issues of convergence in high-dimensional areas. Also, they are sensitive to the initial parameter settings, which could impact the overall robustness and reliability of the fake news detection system.

(Helmstetter and Paulheim, 2018) used weakly supervised learning to detect fake news on Twitter. This method created a large-scale training dataset by labeling tweets according to the reliability of their sources, instead of their integrity. This approach is essential in collecting large, accurately labeled datasets. As a result, it addressed a common limitation in fake news detection.

Even though it gave false positives and negatives, the model achieved an F1 score of up to 0.9. It offered a practical solution to the limitations of smaller, meticulously assembled datasets.

Weak supervision can negatively impact fake news detection. And our research focuses on the same scalability to ensure more reliable fake news detection with efficient computational resources.

The weakly supervised learning process heavily relies on the source accuracy for initial labeling. It can show biases if the assessment of source credibility needs to be revised. Moreover, this method may require additional support in case of unseen data sources or rapidly evolving news topics where we may still require established credibility metrics.

(Abu-Salih et al., 2018), offered an all-inclusive framework for the evaluation of the users' credibility across various big social data domains. The CredSaT framework in this study combines semantic analysis to reduce the doubt in user-generated content. And all this is done with practical implications. It was a crucial step to learn and verify the context of information, which is important to detect fake news. Moreover, a temporal dimension allowed the model to adapt to changes over time and become more effective for the dynamic nature of fake news.

All in all, CredSaT came out to be an effective method of assessing user credibility through semantic and temporal analysis. However, it lacks an understanding of the complexities of users' behavior patterns across different platforms and their frequently changing algorithms.

(Aphiwongsophon and Chongstitvatana, 2018) applied machine learning techniques to identify and address the spread of fake news. They analyzed various algorithms such as decision trees, support vector machines, and neural networks. They assessed how much the entities were effective in parsing and classifying data for its authenticity. They found feature selection and model tuning to be important for improving detection accuracy. This makes the finding very relevant to our project. The paper from 2018 is a comprehensive analysis of different machine-learning strategies. The analysis provides a critical foundation to understand the strengths and limitations informatically different ways in which various other papers solved similar problems Therefore we can understand the optimization and development of our detection model to reach high levels of accuracy with reliability based on this information.

(Baarir and Djeffal, 2021) analyze fake news detection with a Support Vector Machine (SVM) and TF-IDP for feature extraction. The research discusses a system that helps increase the accuracy of sorting fake from real news. The methodology included very deep text preprocessing and encoding using a simple bag of words and N-grams. It also processed multi-dimensions in feature extraction in the textual source along with publication date, author, sentiment.

The insights from (Baarir and Djeffal, 2021) are important because of their successful dataset creation and SVM application. It highlights the efficacy of detailed preprocessing and integration of SVM features into Random Forest modeling.

Another analysis by (Jain and Kasbe, 2018) centered on the fake news detection through a Naive Bayes classification model. It classified fake and real news articles on Facebook. Their methodology used web scraping to update the datasets with real-time, reliable news for ongoing machine learning training. They kept updating the model with newer data to help increase its accuracy. They tried out several changes, such as varying the length of articles used for training and implementation in practice of web scraping to keep the dataset current.

(Jain and Kasbe, 2018) use Naive Bayes in the current project. This could be a comparative analysis for the implementation of adaptive learning techniques and proposes potential areas within a Random Forest framework for such implementation.

(Jain and Kasbe, 2018), however, depended on Naive Bayes which might not effectively handle the complexity and subtleties of detection against lies. The model may therefore be unable to generalize effectively across different types of fake news content.

In their work on machine learning techniques for detecting falsehoods in social media, (Raja and Raj, 2022) conducted a comprehensive study. They suggest a model that includes feature extraction methods such as Count Vectorizer, N-gram, as well as TF-IDF Vectorizer. Additionally, they used machine learning classifiers such as SVM, Naive Bayes, Logistic Regression, and Random Forests. According to their findings, the highest accuracy is realized by SVM classifier with TF-IDF feature extraction reaching up to 93%. This demonstrates how important more advanced feature extractions are in order to improve the performance of classification algorithms in detecting fake news.

(Raja and Raj, 2022) study successfully applies the TF-IDF and SVM, proving the effectiveness of combining refined text analysis techniques. Moreover, the study provides significant practical evidence supporting the use of machine learning in detecting fake news. This can help us in making further development and refinement of the proposed model. However, the model adaptation to different kinds of fake news content can be a limitation.

(Reis et al., (2019) explore the utilization of explainable machine learning models. The study focuses mainly on XGB (extreme gradient boosting machines) to detect fake news on social media. Their work evaluates the effectiveness of almost 200 diverse features and examines their discriminative power by developing and evaluating a plethora of models. The uniqueness of this study lies in its unbiased model generation approach that randomly selects features for model testing. It leads to a detailed analysis of a feature's impact on model performance. The results indicate the benefits of integrating a wide range of features and model explainability in fake news detection. These factors are significant to refine the effectiveness of Random Forest models. Moreover, their innovative methodology provides a rigorous framework that could enhance model transparency and feature selection in my proposed project. But, the vast number of models generated in this study is a potential limitation. While the number makes the process thorough, it can affect the organization of the computation process. Such a detailed method also makes it difficult to apply in real-world scenarios where we often have limited time and resources. This limitation calls for more targeted feature selection strategies in future research.

(Waikhom and Goswami, 2019) examine the use of ensemble machine-learning techniques in their research. Their aim was to enhance the detection accuracy of fake news using the LIAR dataset. The authors enable labeling from multi-label to binary so as to handle the problem of fake news on social media. Also, they focus on a more pragmatic method that can be adopted for real world situations. Essential steps include feature extraction; advanced pre-processing such as TFIDF vectorization, N-grams and cosine similarity measures.

In this 2019 study, ensemble methods were used that included Bagging classifiers and XGBoost which help in understanding how integration of multiple models can improve detection accuracy. This approach aligns with our Random Forest framework to improve classification results.

However, their study primarily focuses on textual features that can ignore other significant attributes such as network patterns and user behavior. These attributes could contribute to a more holistic fake news detection

system. This shows the importance of integrating diverse data dimensions for the accuracy and strength of detection models. Moreover, the dataset in this paper is not significantly detailed and results could vary if we add more detailed data into it.

In summary, the reviewed studies reflect different machine-learning approaches that we can use to address the issue of fake news, particularly on social media. Several of the above researchers focus on Naïve Bayes and Decision Trees for Twitter. These models provide a standard to evaluate the model performance. Spam SMS detection through classifiers such as neural networks and SVM sheds light on text classification challenges relevant to fake news (Gupta et al., 2018). (Altunbey Ozbay and Alatas, 2019) explored the application of metaheuristic optimization algorithms to improve fake news classification.

All in all, these studies show a torch to approach our methodology where we aim to use Random Forest algorithm to accurately detect the fake news. The findings of these studies support us in developing an enhanced machine-learning model that can improve the reliability of fake news detection systems.

3. Research Methodology:

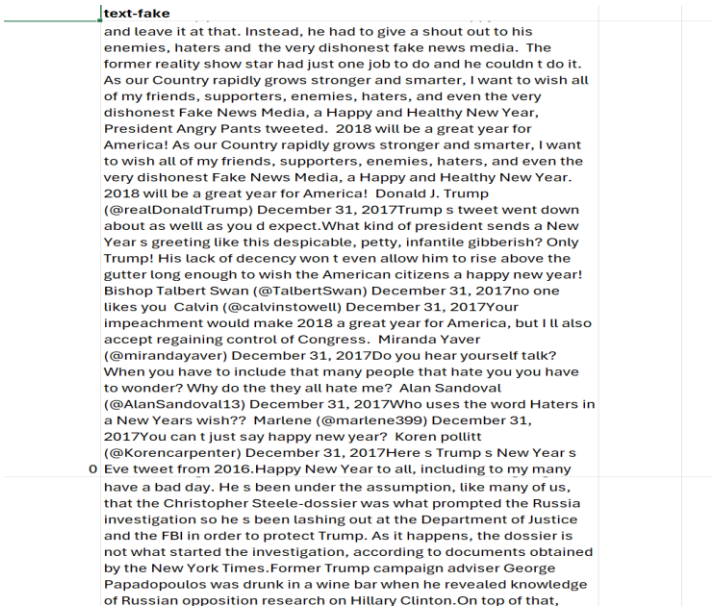
In the research methodology section, we have discussed description of procedures, techniques and statistical methods used in the study in detail. This step-by-step approach will help us understand the results in a better way.

3.1 Research Procedure:

1-Data Collection:

The data used in this study is based on news articles which are categorized as Real or Fake articles. This data was collected and gathered from publicly available resource Kaggle.com and consist of two CSV files i.e. **DataSet_Misinfo_TRUE.csv** for real news and **DataSet_Misinfo_FAKE.csv** for fake news articles data.

Datasets illustrated below:



	text-fake
0	and leave it at that. Instead, he had to give a shout out to his enemies, haters and the very dishonest fake news media. The former reality show star had just one job to do and he couldn t do it. As our Country rapidly grows stronger and smarter, I want to wish all of my friends, supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year, President Angry Pants tweeted. 2018 will be a great year for America! As our Country rapidly grows stronger and smarter, I want to wish all of my friends, supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year. 2018 will be a great year for America! Donald J. Trump (@realDonaldTrump) December 31, 2017Trump s tweet went down about as welll as you d expect.What kind of president sends a New Year s greeting like this despicable, petty, infantile gibberish? Only Trump! His lack of decency won t even allow him to rise above the gutter long enough to wish the American citizens a happy new year! Bishop Talbert Swan (@TalbertSwan) December 31, 2017no one likes you Calvin (@calvinstowell) December 31, 2017Your impeachment would make 2018 a great year for America, but I ll also accept regaining control of Congress. Miranda Yaver (@mirandayaver) December 31, 2017Do you hear yourself talk? When you have to include that many people that hate you you have to wonder? Why do the they all hate me? Alan Sandoval (@AlanSandoval13) December 31, 2017Who uses the word Haters in a New Years wish?? Marlene (@marlene399) December 31, 2017You can t just say happy new year? Koren pollitt (@Korencarpenter) December 31, 2017Here s Trump s New Year s Eve tweet from 2016.Happy New Year to all, including to my many have a bad day. He s been under the assumption, like many of us, that the Christopher Steele-dossier was what prompted the Russia investigation so he s been lashing out at the Department of Justice and the FBI in order to protect Trump. As it happens, the dossier is not what started the investigation, according to documents obtained by the New York Times.Former Trump campaign adviser George Papadopoulos was drunk in a wine bar when he revealed knowledge of Russian opposition research on Hillary Clinton.On top of that,

Fake News Dataset

	text-true
0	who voted this month for a huge expansion of the national debt to pay for tax cuts, called himself a â€œfiscal conservativeâ€ on Sunday and urged budget restraint in 2018. In keeping with a sharp pivot under way among Republicans, U.S. Representative Mark Meadows, speaking on CBSâ€™ â€œFace the Nation,â€ drew a hard line on federal spending, which lawmakers are bracing to do battle over in January. When they return from the holidays on Wednesday, lawmakers will begin trying to pass a federal budget in a fight likely to be linked to other issues, such as immigration policy, even as the November congressional election campaigns approach in which Republicans will seek to keep control of Congress. President Donald Trump and his Republicans want a big budget increase in military spending, while Democrats also want proportional increases for non-defense â€œdiscretionaryâ€ spending on programs that support education, scientific research, infrastructure, public health and environmental protection. â€œThe (Trump) administration has already been willing to say: â€¯Weâ€™re going to increase non-defense discretionary spending ... by about 7 percent,â€™â€ Meadows, chairman of the small but influential House Freedom Caucus, said on the program. â€œNow, Democrats are saying thatâ€™s not enough, we need to give the government a pay raise of 10 to 11 percent. For a fiscal conservative, I donâ€™t see where the rationale is. ... Eventually you run out of other peopleâ€™s money,â€ he said. Meadows was among Republicans who voted in late December for their partyâ€™s debt-financed tax overhaul, which is expected to balloon the federal budget deficit and add about $1.5 trillion over 10 years to the $20 trillion national debt. â€œItâ€™s interesting to hear Mark talk about fiscal responsibility,â€ Democratic U.S. Representative Joseph Crowley said on CBS. Crowley said the

True News Dataset

2-Data Preparation:

Before we train and test the data for evaluation purposes, we must prepare the data to make sure it is in actual required condition for the required tests to be performed. Following steps are taken to prepare the data.

i- Loading Data:

Both CSV files were uploaded in Panda Data Frames. To facilitate more comprehensive results and analysis, the Data Frames are merged into a single Data Frame.

ii- Data Preprocessing and Cleaning:

The data used for the model testing consist of two datasets which were combined into one for cleaning purposes into the single DataFrame. Missing values in the **text** column of data were handled by filling them with empty strings. Further, a new column was added which would represent the length of each article named as **text_length.**

	Unnamed: 0	text	label
0	0	The head of a conservative Republican faction ...	TRUE
1	1	Transgender people will be allowed for the fir...	TRUE
2	2	The special counsel investigation of links bet...	TRUE
3	3	Trump campaign adviser George Papadopoulos tol...	TRUE
4	4	President Donald Trump called on the U.S. Post...	TRUE

	text_length
0	4635
1	2537
2	2765
3	2437
4	5172

Figure 1: New Article Column -text_length

3- Exploratory Data Analysis (EDA):

EDA was conducted to understand the characteristics and structure of the dataset. To prevent extreme values from skewing the result, outliers' detection and removal was performed on the data.

Statistical Visuals: To understand the distribution of text lengths and the frequency of articles in each category, descriptive statistics were generated to summarize the stats.

i- Word Clouds:

To visualize the data in this phase, Word clouds were created to visualize the most frequently used words in Real and Fake news articles. This would help us to identify the common themes and terms used within the articles. To understand how the data is spread, distribution of article lengths is plotted and further to get an understanding of most frequently occurring terms, word clouds for real and fake news has been generated.

Figure 2: Word clouds

ii- Distribution of Article Lengths after Outlier Removal:

Distribution of article lengths were plotted which will compare the length of articles in two categories. With the help of outliers extremely short or long articles can be considered outliers when they are compared to the majority. The outliers were successfully detected and removed based on text_length. With the help of detection, it can be assured that extreme values do not skew the performance of model.

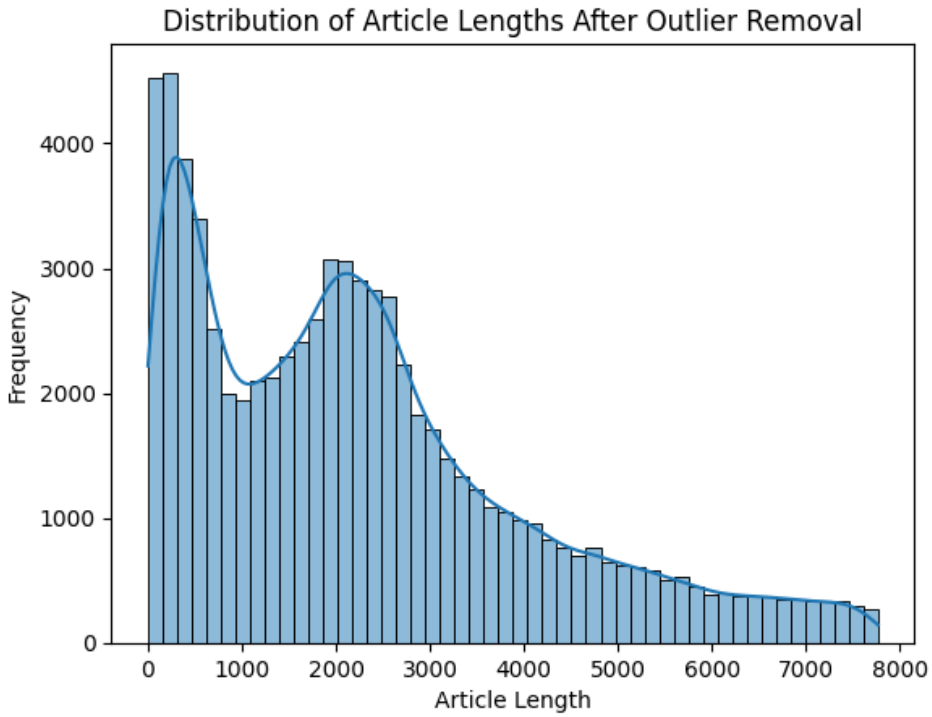


Figure 3: Distribution of Article lengths after Outlier Removal

3.2 Equipment and Tools:

To prepare the data and run the required tests accurately following tools were used:

Python:
Python has been used as a primary programming language for analysis and data processing.

Google Colab:
Google colab has been used for development, testing and execution of the code including the results with visuals to better understand the analysis against the available data.

Python Libraries:

- **Panda-** library was used for manipulation of the data.
- **Numpy-** Used to manage mathematical operations and handle arrays.
- **'Matplotlib'** and **'Seaborn'** was used visualization of data.
- **Wordclouds -** was used to generate the word clouds.
- **Scikit-learn -** was used for evaluation matrix and machine learning models.
- **Textblob-** Used to perform sentiment analysis on text.
- **Textstat-** Used to assess readability scores of data.
- **Nltk-(Natural language toolkit)-** Used for text tokenization and NLP tasks.

- **Confusion matrix display-** Used to plot confusion matrices and display.

3.3 Applied Techniques:

Before we could start the tests for analysis, several machine learning techniques were applied on the datasets:

1-Feature Engineering:
In this step, TF-IDF (Term Frequency- Inverse Document Frequency) Vectorizer was used to convert the data which consisted of texts into numerical features. Additionally, sentiment scores and readability indices were also developed which helped in improving the models` ability to distinguish between real and fake news.

2- Statistical Techniques:
To evaluate the models for required results, following metrics were used:

Accuracy:
Accuracy is the ratio of instances which are correctly predicted to the total instances.
Precision:
The ratio of observations correctly predicted to the total predicted positives.
Recall:
The ratio of observations correctly predicted to all the observations in the actual class.
F1- Score:
It is the weighted average of Prediction and Recall.
Confusion Matrix:
It is a table used for evaluation of classification model performance showing the actual versus predicted classifications.

4. Design Specification:

In this section the architecture and techniques used for implementation are summarized. This section will determine and explain the framework, important elements, and proposed models along with their functionalities.

4.1. Framework:

The research framework uses machine learning techniques to systematically process and classify news articles into fake or real categories. The framework consists of following components:

1. Data Preprocessing and Loading:

- **Loading data:** The first step involves loading datasets consisting of fake and real news articles. The data is stored in CSV files which are then read into Panda Data frames for easy manipulation.
- **Data Merging:** Both separate datasets for fake and real news are merged into one Data frame to facilitate comprehensive and detailed analysis.
- **Handling Missing Values:** If there are any missing values in the text column of the data, they are filled with empty strings to ensure consistency in data processing phase.

2. Feature Engineering:

The Term Frequency-Inverse Document Frequency (TF-IDF) Vectorizer converts the textual data into numerical features. This technique transformed the text into a matrix of TF-IDF features, representing the importance of words in the documents relative to the entire corpus. The TF-IDF Vectorizer helped reduce the impact of commonly used words and highlighted the significance of rarer terms. Furthermore, an additional feature, text_length, representing the length of each article, is computed and included in the dataset. Further, readability indices and sentiment scores were also created to help enhance the models` ability.

3. Model Training:

The framework involves the training of three different machine learning models i.e., Random Forest, Naive Bayes, and Decision Tree. Each model is trained on the pre-processed data, continuously learning and improving to discern the patterns that distinguish real news from fake news. **Random Forest** utilizes an ensemble of decision trees to enhance classification performance and decrease overfitting. **Naive Bayes** is considered as a probabilistic classifier based on Bayes' theorem, suitable for text classification. **Decision Tree** is a tree-like model which splits data into subsets based on its feature values, which then results into a decision about the class label.

4. Model Evaluation:

All three models are evaluated using different metrics to assess their performance in classification of news articles. To get a comprehensive evaluation of these models' metrics including accuracy, precision, recall and F1-score are calculated.

To visualize the performance of each model, a confusion matrix and ROC Curve were generated. The matrix vividly presents the performance of models in terms of true positives, false positives, true negatives and false negatives which helped in a deeper understanding of their performance.
To ensure the robustness of the models and to analyze their performance across different subsects of the data, the cross-validation test was also performed for these models. Lastly, the learning curve is plotted to visualize the performance improvement of models with increased data.

4.2. Algorithms Description and Architecture:

The system architecture of this research involves different steps from data ingestion, data pre-processing, model training, model evaluation and visualizations. In this study, the focus was to apply well-established machine learning algorithms to classify news articles. While we did not develop a new model in this process, the implementation and combinations of these models provided valuable insights into their comparative performance. By comparing the performance of Random Forest, Naïve Bayes and Decision Tree, we were able to identify the strengths and weaknesses of each method. This comparative analysis will not only serve as a basis for our study but also serve as a basis for future research where hybrid models of different classifiers could be analyzed to further enhance classification performance.

5. Implementation:

The implementation of this research project was initiated from data collection, preprocessing, training of model and with final evaluation. To implement the code python 3.8.8 was utilized and the dataset comprising of one data CSV file was used after combining two different data files containing Fake and True datasets. Multiple libraries such as Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn and Wordcloud were used for visualization and modeling of data. For training and evaluation purposes a local machine was used with 16GB RAM and Intel Core i7 processor.

5.1 Output Produced:

1- Transformed Data:

The raw textual data gathered from the news articles was transformed into the numerical features using the TD-IDF vectorizer. This transformation process enabled the models to effectively interpret and process the data. Further, a new feature defining the length of each article was created to provide additional insight into the characteristics of data.

2- Code Written:

In this phase, python scripts were created to automate the workflow of the project starting from data loading and preprocessing to the training of model and evaluation. These scripts were categorized into certain modules to handle tasks such as data cleaning, feature extraction, training of model and evaluation of performance. Data preprocessing scripts consisted of steps to clean the data by creating the new features after removing missing values. Feature extraction scripts used the TF-IDF vectorizer to convert the data into numerical form.

```
# Initialize the TF-IDF Vectorizer
TFIDF_VECTORIZER = TfidfVectorizer(max_features=5000)

# Fit and transform the text data into numerical features
X_FEATURES = TFIDF_VECTORIZER.fit_transform(ALL_NEWS_DF['text']).toarray()

# Convert labels to numerical values: 1 for 'TRUE', 0 for 'FAKE'
Y_LABELS = ALL_NEWS_DF['label'].apply(lambda x: 1 if x == 'TRUE' else 0).values
```

Figure 4: Feature Extraction using TF-IDF

For each machine learning model, separate scripts were written for training purposes. Each script included steps to split the data into training and testing sets that would help to evaluate the performance of these models.

5.2 Developed Models:

1- Random Forst Classifier: An ensemble model that incorporates numerous decision trees to enhance the accuracy of classification and reduce the overfitting risk. This model was selected for its robustness in handling complex datasets and high-performance level.
2- Naïve Bayes Classifier: It is a probabilistic model based on Bayes` theorem. This is very well suited for text-based classification tasks. It was selected due to its ability to handle large datasets efficiently and with simplicity.
3- Decision Tree Classifier: It is a model that employs a tree-like structure to make decisions based on the characteristics of data. This model was included for its ease of visualization and interpretation capabilities.

5.3 Visualizations:

To provide clear insights into the data and performance of models, different visualizations were developed. These visualizations would visualize the performance of models after running the required tests.

1-Confusion Matrix:

In this study, we employed the confusion matrix as a critical evaluation metric to assess the performance of our machine-learning models. The confusion matrix delivers a detailed analysis of the model's predictions, which highlights the number of true positives, true negatives, false positives, and false negatives. By visualizing these matrices, I better understood how well each model distinguished between real and fake news articles, allowing me to identify areas of strength and potential improvement in our classification approach.

2- ROC Curve and AUC Score:
The ROC curve is a visual plot which shows the diagnostic abilities of binary classifier, and the AUC score provides a single scalar value to summarize the performance of the model. With the help of this visualization, it further helped me in showing the trade-off between recall and fall-out of different points. Also, AUC allowed me to measure the ability of the model to differentiate between classes which indicate better performance through higher values.

The ROC curve comparison indicates that the Random Forest classifier surpasses the Naive Bayes and Decision Tree classifiers in differentiating between real and fake news articles. The AUC scores confirm these findings, with the Random Forest reaching an AUC of 0.99, Naive Bayes 0.92, and Decision Tree 0.89. This further illustrates the effectiveness of the Random Forest classifier for the task, which makes it the most suitable model based on the given data and evaluation metrics.

3- Classification Report:
With the help of classification report I received a detailed summary of precision, recall, F1-score and support for each class. It provided a broad summary on the performance of model in each class which further explained how well the model is performing in terms of these components.

6. Evaluation:

This research project was conducted and evaluated with three different tests where each model was tested with the key focus towards the metrics such as confusion matrix, accuracy, recall, precision and F1-score to evaluate the performance. In this research, based on the type of dataset and information within the data I considered using the machine learning models i.e. Random Forest, Naïve Bayes and Decision tree as these models are considered to be very effective against such data. In this research, the proposed machine learning model is Random Forest whereas Naïve Bayes and Decision tree are used as baseline models. In this evaluation session, I have shown the results and the comparison between these models for better understanding.

6.1 Experiment 1: Random Forest Classifier:

In this experiment, the preprocessed and vectorized dataset was split into training and testing datasets. The test size for this was 0.2. In the next step, the Random Forest classifier was then trained on the training data. The final evaluation on the test data gave the following results.

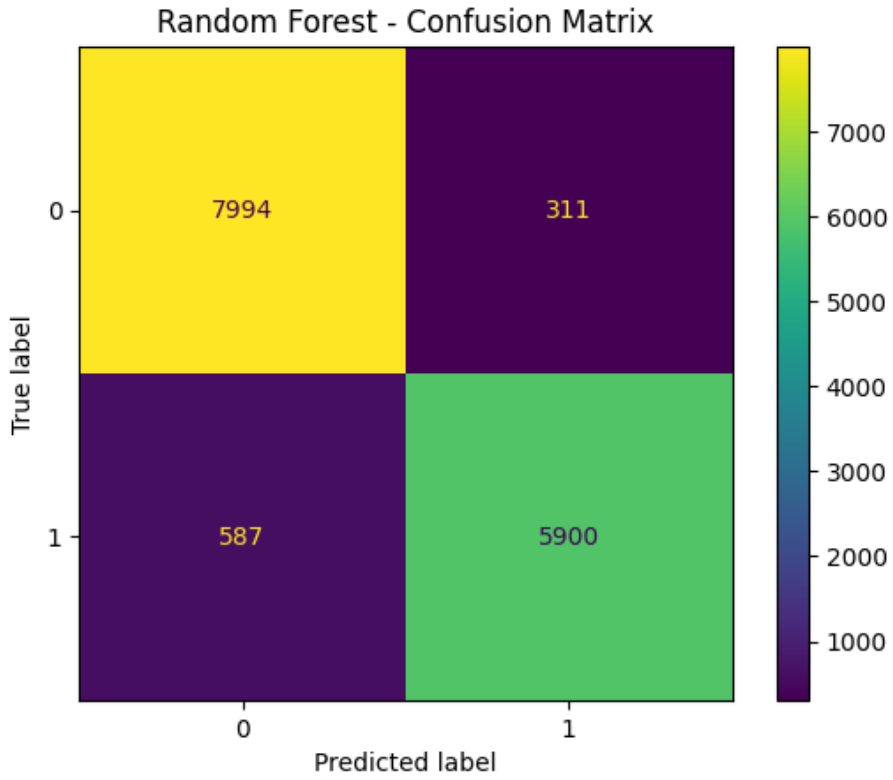


Figure 5: Confusion Matrix of Random Forest

The confusion matrix of the Random Forest classifier provides a visual illustration of its performance by showing the number of True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). The results show that the Random Forest classifier achieved an accuracy of 94% which means that it predicts the 94% of instances correctly. The 95% precision means that the instances it predicted as true were true. The recall of 91% shows that the 90% percent of true instances identified by the model were correct. Finally, the 93% of F1-score shows the balance between precision and recall.

6.2 Experiment 2: Naïve Bayes Classifier:

After preprocessing and splitting the vectorized data into training and testing data with the same size of 0.2, the Naïve Bayes classifier was also trained on the training data. Final evaluation provided an with the following results.

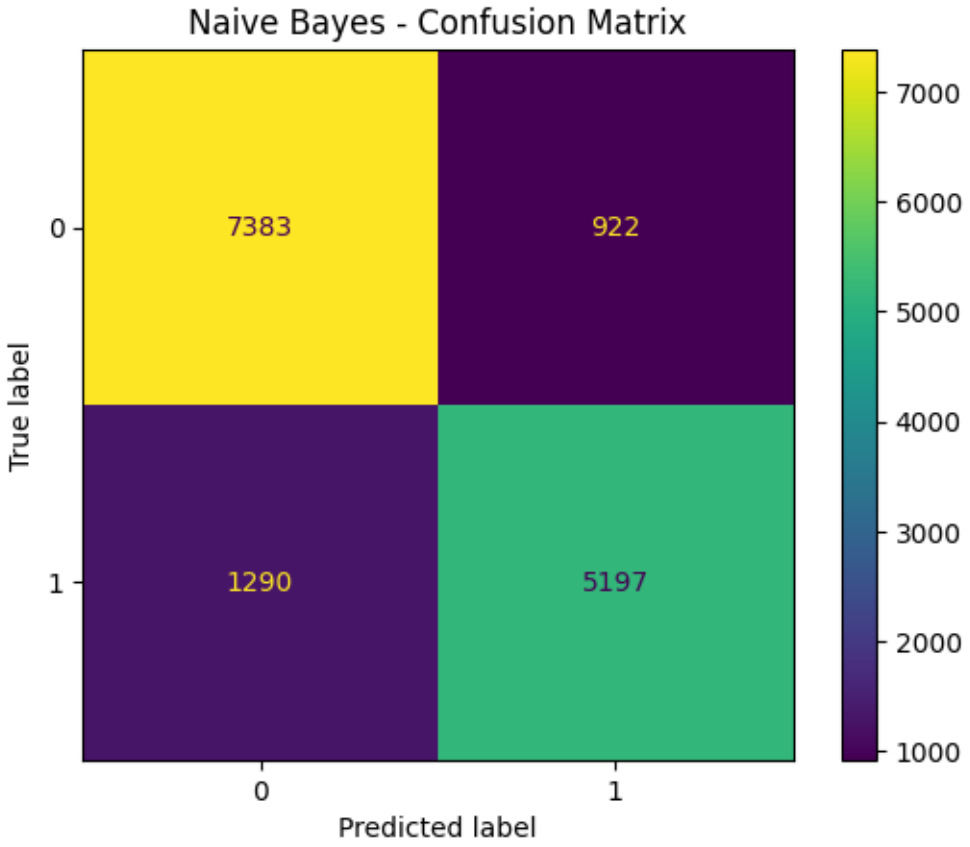


Figure 6: Confusion Matrix of Naïve Bayes

The confusion matrix of Naïve Bayes shows the visual illustration of the model`s performance towards the provided dataset. The Naïve Bayes classifier shows the accuracy of 85% which means that it predicted 85% of the instances correctly. The precision of 85% shows that 85% of instances predicted as true were true in actual. The recall of 80% indicates that the 80% true instances predicted by the model were correctly done. The F1- Score of 82% provides a balance between precision and recall.

6.3 Experiment 3: Decision Tree Classifier:

In the third experiment, the preprocessed and vectorized dataset was again split into training and testing datasets with the size of 0.2. After that, the Decision Tree classifier was trained on the training data. On completion of the tests the following results were obtained.

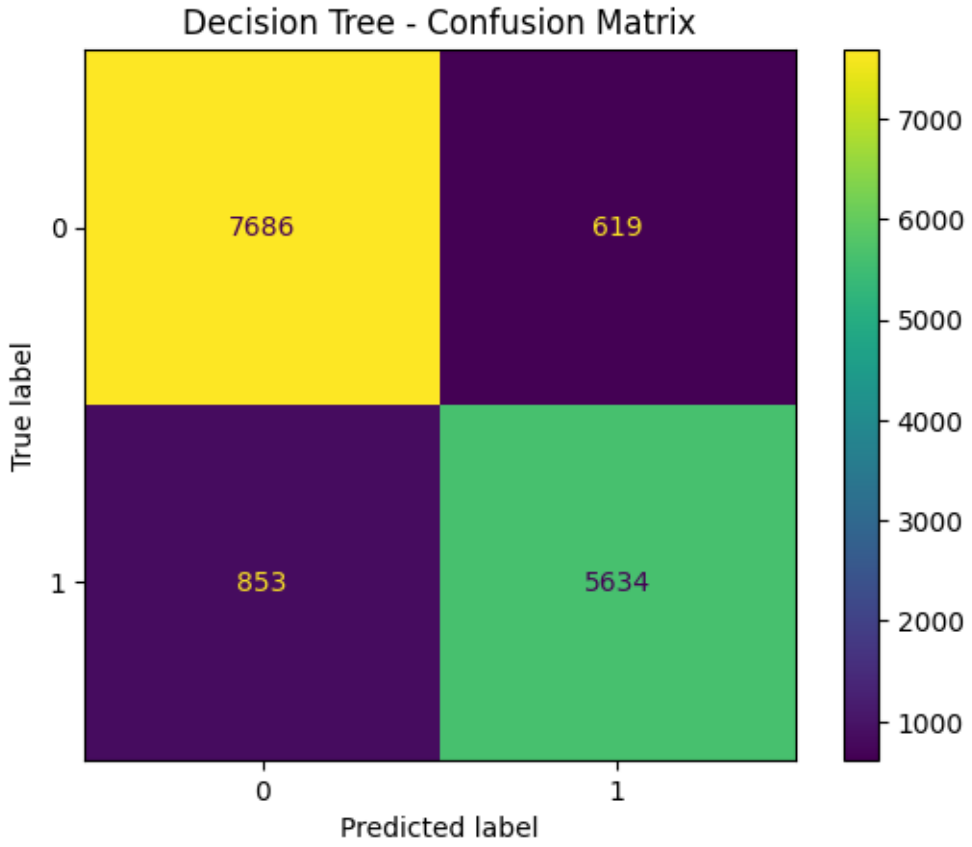


Figure 7: Confusion Matrix of Decision Tree

The confusion matrix of Decision Tree provides a visual representation of performance of the model. The Decision Tree classifier achieved an accuracy of 90% which shows that it correctly predicted 90% of instances. Precision score was 90% meaning that 90% of instances predicted as true were actually true. The recall of 87% shows that true instances of 87% were correctly identified by the model. The F1-score of 88% shows the balance between precision and recall.

6.4 Experiment 4: ROC Curve and AUC Score:

The ROC curve was utilized to evaluate the trade-off between sensitivity (true positive rate) and particularity (false positive rate) across different threshold settings for each model. By plotting the ROC curve, I could visually assess our classifiers' overall performance, with the Curve (AUC) providing a single scalar value that reflects the model's ability to differentiate between classes. This analysis helped me to identify the most effective model for accurately classifying real and fake news.

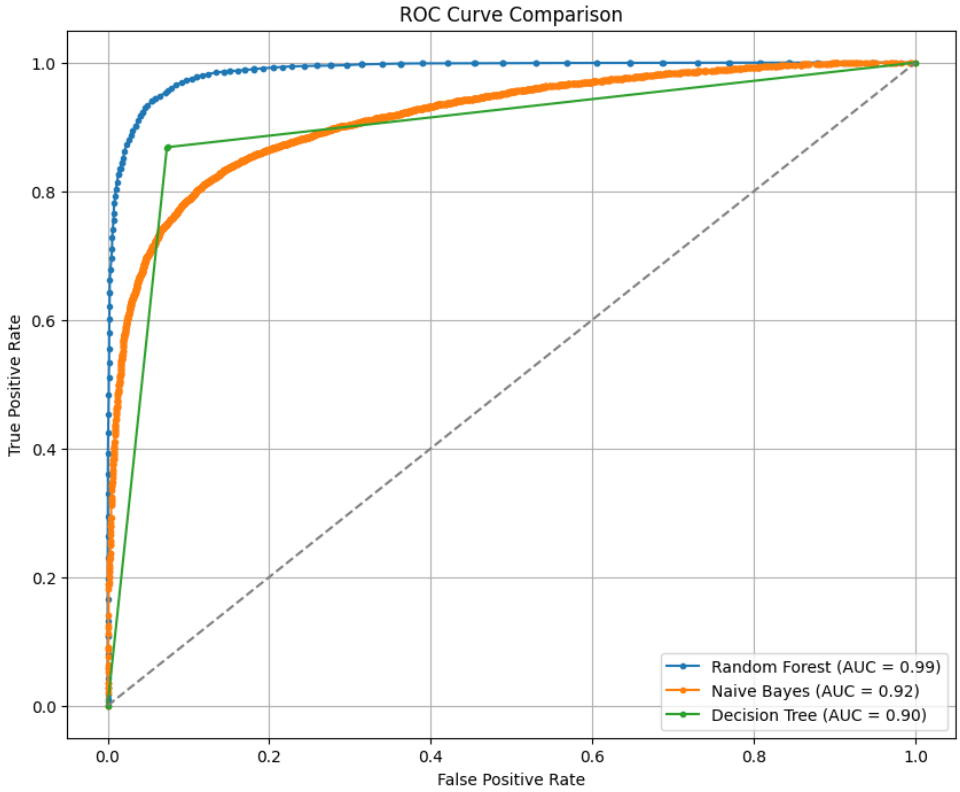


Figure 8: ROC Curve and AUC Score Comparison

6.5 Experiment 4: Classification Report:

Random Forest - Classification Report				
	precision	**recall**	**F1-score**	**support**
0	0.93	0.96	0.95	8305
1	0.95	0.91	0.93	6487
accuracy			0. 94	14792
macro avg	0. 94	0.94	0. 94	14792
weighted avg	0. 94	0.94	0. 94	14792

Naive Bayes - Classification Report				
	precision	recall	F1-score	support
0	0.85	0.89	0.87	83Ø5
1	0.85	0. 83	0.82	6487
accuracy			0.85	14792
macro avg	0.85	0.85	0.85	14792
weighted avg	0.85	0.85	0.85	14792

Decision Tree - Classification Report				
	precision	recall	F1-score	support
0	0.90	0.93	0.91	8365
1	0.90	0.87	0. 88	6487
accuracy			0. 96	14792
avg	0.90	0. 96	0. 96	14792
weighted avg	0.90	0. 96	0. 96	14792

Figure 9: Classification Report of three models

6.6 Discussion:

These results stress the importance of model selection appropriately based on the specific requirements of the classification task. The Random Forest classifier proved to be more dependable and robust for classification of news articles as real or fake through its excellent approach.

7. Conclusion and Future Work:

The performance of all three models is evaluated on the basis of confusion matrix and key metrics results. The performance of Random Forest classifier surpassed the Naïve Bayes and Decision Tree classifiers in all key metrics such as accuracy, precision, recall and F1-score. Although, it is efficient, the Naïve Bayes had a lower recall which means that it skipped and missed the true instances more than the other models. The

Decision Tree classifier delivered a satisfactory balance, but it did not perform as well as the Random Forest classifier.

This project focused on developing and assessing suitable and effective machine learning models which can identify the fake and real news in this current age of misinformation on digital media. Using the dataset based on real and fake news articles, after preprocessing and using TF-IDF for vectorizing the data, three methods were used to evaluate it i.e. Random Forest, Naïve Bayes and Decision Tree.

Based on the evaluation results, the Random Forest classifier showed highest performance with its accuracy as 94%, AUC of 99%, and of recall of 96 % for fake news and 91% for real news, precision of 93% for fake news and 95% for real news, and F1-score of 95% for fake news and 93% for real news. In comparison, Naïve Bayes model showed accuracy of 85% and AUC score of 92%. precision at 85% for fake and real news, the recall of 89% for fake news and 80% for real news, which resulted in an F1-score of 87% for fake news and 82% for real news. While the Decision tree model ended up with accuracy of 90%, and AUC of 90%, precision of 90% for both fake and real news, recall of 93% for fake news and 87% for real news, and F1-score of 91% for fake news and 88% for real news. This shows that the Random Forest classifier was not just most effective, it also showed excellent balance between recall and precision through its robust performance.

Models	**Accuracy**
Random Forest	0.94
Naive Bayes	0.85
Decision Tree	0.90

The implication of this research approach to integrate a valuable model into digital platforms to help the detection of fake news which will result in reducing the misinformation. However, the performance of models may vary based on the available datasets with different variables. For future work, the research can be more focused on using more diverse and larger datasets combining the strengths of different classifiers. Real time testing of models can be even more effective in live environments and most importantly, the fairness of data and unbiased classification must be a key to get more suitable results in digital world.

Acknowledgement:

I would like to acknowledge and thank Prof. Devanshu Anand for all his support, guidance and help throughout the semester which allowed me to finish this report as required.

8. References

Abu-Salih, B., Wongthongtham, P., Chan, K.Y. and Zhu, D. (2018). CredSaT: Credibility ranking of users in big social data incorporating semantic analysis and temporal factor. Journal of Information Science, 45(2), pp.259–280. doi:https://doi.org/10.1177/0165551518790424.

Aldwairi, M. and Alwahedi, A. (2018). Detecting Fake News in Social Media Networks. *Procedia Computer Science*, [online] 141(141), pp.215–222. doi:https://doi.org/10.1016/j.procs.2018.10.171.

Altunbey Ozbay, F. and Alatas, B. (2019). A Novel Approach for Detection of Fake News on Social Media Using Metaheuristic Optimization Algorithms. *Elektronika ir Elektrotechnika*, 25(4), pp.62–67. doi:https://doi.org/10.5755/j01.eie.25.4.23972.

Aphiwongsophon, S. and Chongstitvatana, P. (2018). *Detecting Fake News with Machine Learning Method*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ECTICon.2018.8620051.

Fang, Y., Gao, J., Huang, C., Peng, H. and Wu, R. (2019). Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection. *PLOS ONE*, 14(9), p.e0222713. doi:https://doi.org/10.1371/journal.pone.0222713.

Feng, S., Banerjee, R. and Choi, Y. (2012). *Syntactic Stylometry for Deception Detection*. Association for Computational Linguistics, pp.171–175.

Gupta, M., Bakliwal, A., Agarwal, S. and Mehndiratta, P. (2018). A Comparative Study of Spam SMS Detection Using Machine Learning Classifiers. *2018 Eleventh International Conference on Contemporary Computing (IC3)*. [online] doi:https://doi.org/10.1109/ic3.2018.8530469.

Helmstetter, S. and Paulheim, H. (2018). *Weakly Supervised Learning for Fake News Detection on Twitter*.

Jang, B., Kim, I. and Kim, J.W. (2019). Word2vec convolutional neural networks for classification of news articles and tweets. *PLOS ONE*, 14(8), p.e0220976. doi:https://doi.org/10.1371/journal.pone.0220976.

Katsaros, D., Stavropoulos, G. and Papakostas, D. (2019). Which machine learning paradigm for fake news detection? *IEEE/WIC/ACM International Conference on Web Intelligence*. doi:https://doi.org/10.1145/3350546.3352552.

Klyuev, V. (2018). *Fake News Filtering: Semantic Approaches*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICRITO.2018.8748506.

Shabani, S. and Sokhn, M. (2018). *Hybrid Machine-Crowd Approach for Fake News Detection*. [online] IEEE Xplore. doi:https://doi.org/10.1109/CIC.2018.00048.

Tyagi, S., Pai, A., Pegado, J. and Kamath, A. (2019). *A Proposed Model for Preventing the spread of misinformation on Online Social Media using Machine Learning*.

Baarir, N.F. and Djeffal, A. (2021). Fake News detection Using Machine Learning. *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*. doi:https://doi.org/10.1109/ihsh51661.2021.9378748.

Jain, A. and Kasbe, A. (2018). *Fake News Detection*. [online] IEEE Xplore. doi:https://doi.org/10.1109/SCEECS.2018.8546944.

Raja, M.Senthil. and Raj, L.Arun. (2022). Fake news detection on social networks using Machine learning techniques. *Materials Today: Proceedings*. doi:https://doi.org/10.1016/j.matpr.2022.03.351.

Reis, J.C.S., Correia, A., Murai, F., Veloso, A. and Benevenuto, F. (2019). Explainable Machine Learning for Fake News Detection. *Proceedings of the 10th ACM Conference on Web Science - WebSci '19*. [online] doi:https://doi.org/10.1145/3292522.3326027.

Waikhom, L. and Goswami, R.S. (2019). *Fake News Detection Using Machine Learning*. [online] papers.ssrn.com. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3462938.