Fake News Detection Using Machine Learning

Mohiuddin

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

SUMMER 2025

APPROVAL

This thesis titled on "**Fake News Detection Using Machine Learning**", submitted by **Mohiuddin (ID: 211-35-713)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfilment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents**.**

BOARD OF EXAMINERS

Fazla Elahe

-- **Chairman**

Dr. Md. Fazla Elahe
Assistant Professor & Associate Head
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Marzia

-- **Internal Examiner 1**

Dr. Marzia Ahmed
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

13.09.2025

-- **Internal Examiner 2**

Dr. Shabnom Mustary
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

13.9.25

-- **External Examiner**

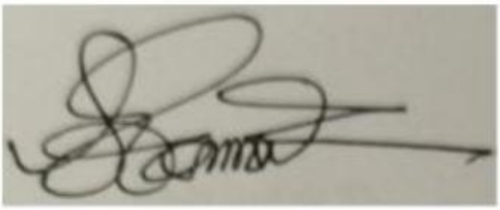
Mohammod Abul Kashem
Professor
Department of Computer Science and Engineering
Dhaka University of Engineering & Technology, Gazipur.

Declaration

I hereby declare that this thesis has been completed under the supervision of **Mr. Nuruzzaman Faruqui, Assistant Professor**, Department of Software Engineering, Daffodil International University. I also affirm that this thesis is my original work, submitted for the degree of B.Sc. in Software Engineering, and neither the entire work nor any portion has been previously submitted for another degree at this or any other university.

Mohiuddin
ID: 211-35-713
Department of Software Engineering
Daffodil International University

Certified By:



Nuruzzaman Faruqui
Assistant Professor
Department of Software Engineering
Daffodil International University

ACKNOWLEDGEMENTS

I would like to thank my supervisor for his wonderful supervision, constructive feedback & the required support throughout this research. I am also indebted to my family members and dear friends who supported me and helped me academically. This thesis would not have been possible without utilizing the resources and facilities of the Department of Software Engineering at Daffodil International University.

ABSTRACT

The domination of digital platforms for distribution of fake news traveling through social media outlets continues to be a pressing concern for sustained societal trust and the sanctity of information. This study focuses on evaluating the performances for different fake news detecting algorithms using machine learning techniques with the popular datasets “Fake.csv” and “True.csv”. In our study, we use of four algorithm models; LR, DTC, GBC and RFC. The data undergoes rigorous preprocessing stages, namely text cleansing, tokenization as well as TF-IDF vectorization. Furthermore, the data is then split into sets of training-test data, where model hyperparameter values are optimized utilizing the GridSearchCV framework. Accuracy, precision, recall and the F1-score calculated for these models prove the stated hypothesis where model Gradient Boosting Classifier outperforms the rest with the best metrics, while Logistic Regression showed good performance, which underlines the model’s usefulness in practice. The durability of DT and RF classifiers under rigorous tuning is noteworthy. The outcome of our study hold significant importance for the worlds of journalism, politics, and social media, stressing the significance of ML models in detecting fake news, thus enabling more public discourse and more informed debates. This study thus reiterates the need for ever-evolving techniques in machine learning to counter the issue of misinformation.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BERT	Bidirectional Encoder Representations from Transformers
CV	Cross Validation
NLP	Natural Language Processing
LE	Label Encoding
DTC	Decision Tree Classifier
GBC	Gradient Boosting Classifier
LR	Linear Regression
RFC	Random Forest Classifier
TF-IDF	Term Frequency-Inverse Document Frequency

CHAPTER 1

INTRODUCTION

1.1 Introduction

The explosion of internet along with the rapid growth of online communication platforms and social media has had a tremendous impact on information dissemination and consumption. Such technologies have led to a new form of journalism which makes access to global news a click away [1]. At the same time, a new concern has arisen: the widespread reach of inaccurate information and news. Fake news, meaning the false information disseminated with an intention to mislead readers, has become a quintessentially modern issue. The spread of fake news threatens democracy, public trust, and social cohesion.

Within politics, fake news can be and is being utilized to impact elections, influence and steer public opinion, and undermine democracy. As example, while the Presidential election of US was going on in 2016, numerous false news stories emerged which influenced public perception, arguably swinging the election [5]. In the case of natural calamities or public health emergencies, the dissemination of false information would create untold confusion and unmanageable repercussions. During the COVID pandemic, misinformation associated with the virus, its treatment, and available vaccines information emerged which severely hindered public health efforts [12]. In addition, the softening of trust in legitimate news sources due to fake news leads to challenges of discerning reality. There is a need to address the effective and scalable countermeasures to the ease and rapid dissemination of fake news online.

The increasing prominence of the problem of fake news has brought attention to the use of ML, which is a branch of AI. The capability of ML algorithms for processing large datasets, in order to identify similarities and make predictions is nothing short of extraordinary. ML algorithms are able to scrutinize the linguistic, stylistic, and other relevant attributes of a news piece, thus training the models to differentiate between genuine and fake news. In this way, ML possesses considerable ability of assisting in the automated and efficient filtration of false news and genuine news articles [2], [24].

This research aims to justify the accuracy, as well as, efficiency of some of the most prominent ML algorithms in identification of fake/false news. Specifically, our interest lies in the LR, DTC, GBC and RFC. These algorithms are quite well known for binary classification and are very effective with very complicated datasets. Through the comparison of the mentioned ML models, we hope to focus on the strengths and weaknesses of these algorithms in detecting fake/false news while establishing which one of them performs best.

The datasets used in this study - “Fake.csv” and “True.csv” - are well-known in the area of NLP as they contain news articles which are heavily annotated. Dataset “Fake.csv” consists of articles suspected to be false or misleading, and “True.csv” contains articles deemed to be true. These datasets prove to be a foundational asset in the training and benchmarking of the ML models as they seek to optimize the text classification tasks.

Resolving the problem follows the methodology which begins with data preprocessing steps. This step is critical in transforming the raw text data for analysis. In this case, the cleaning process consists of the extraction of HTML tags as well as the straining of punctuation and minimal meaning words, also known as stop words. Thereafter, text tokenization is done, which is the splitting of sentences into words or tokens and is followed by a vectorization step of TF-IDF, which is a statistical technique which evaluates any word's relevance in a given document as compared to larger sets of documents. In TF-IDF, words that are contextually important are assigned a greater value which aids in efficiently modelling training feature selection [4].

Post data preprocessing, we moved onto the set partitioning with training, as well as, testing sets with 70-30 ratio. The method helps in ensuring that repeated samples during testing are of a considerable quantity. Each machine learning algorithm is fitted on their matching training sets and hyperparameters are fine-tuned by means of GridSearchCV [3]. Hyperparameter optimization is critical to any model as it is the process that sets the model to work to a user set expectation. GridSearchCV finds the best set of hyperparameters for a model by assessing its various configurations to ensure maximum accuracy and generalization on validation data and unseen data.

For model evaluating and testing stage, each model's performances are analyzed with a set of commonly accepted evaluating criteria like- accuracy, precision, recall and the F1 score. Each of these metrics retrieves a value that is beneficial and informative. For example, accuracy retrieves the percent of correct samples classified out of the total sample population. Precision deals with accuracy of all correct guesses, while recall, or mentioned as sensitivity in many studies, is the proportion of truly positive guesses. F1 score, also defined as harmonic mean of afore mentioned precision with recall is useful where there is class imbalance. In a nutshell, these metrics as a whole present a reasonable and well-rounded evaluation of the performance of the models, which is useful in selecting the right ML algorithm with the mentioned problem [2].

This study is essential because it enhances the correctness, as well as, trustworthiness for the fake/false news detecting set-ups. Our work aids in developing sophisticated systems to combat misinformation by pinpointing the top-performing ML system. This research is significant for several stakeholders, such as journalists, politicians, and analysts of social media. Detecting misinformation is crucial in restoring public trust in the media, facilitating informed choices, and safeguarding democratic values. Furthermore, this research provides a basis for employing more advanced techniques in natural language processing, particularly within frameworks combining multiple methods, to improve detection accuracy and robustness.

This research is equally useful for understanding the accuracy for different ML models in detecting fake/false news, and it also reveals directions for new research. Some advanced NLP methods, such as DL models and transformer-based systems, could further improve detection accuracy. The application of LSTM and CNN frameworks have proven effective for various NLP tasks as they can comprehend deep contextual and pattern-based intricacies [6], [7], [16]. With techniques like- BERT, NLP is enhanced further as transformer-based models comprehend word context better within sentences [8], [9], [10].

The combination of other modalities such as images and videos with texts can strengthen systems advanced for detecting fake news due to their robustness. Text documents usually come with various multimedia elements. Their examination together with the text helps in fake news detection [22]. Members of the public can participate in more informed discussions as fake news detection is available [14]. More research is needed to explore the social impacts and ethical obligations of such technologies, prioritizing focus on the lack of clarity on the technologies employed. The application of ethical AI principles that safeguard fairness, responsibility, and openness are crucial to mitigate bias and unforeseen impacts when designing systems for propaganda and fake news detection [23].

1.2 Conclusion

In the attempt to mitigate the spread of disinformation, this very study seeks in assessing the effectiveness of different ML practices, thereby demonstrating the capability of ML in detecting fake news. It is essential to highlight that ML has all potential in mitigating the issues for fake/false news distribution and detection, and this capability will contribute toward ensuring an informed and adaptive society. In this regard, the study fills in the gap in the machine learning and fake news detection ecosystem and encourages innovation and further research towards machine learning designed for this purpose. One must be ready for the impending change in the digital landscape.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Research pertaining to the utilization of ML and NLP for fake news detection is abundant and rapidly developing. In view of the pervasive societal problem fake news poses, developing automated tools for the detection and curtailment of its spread has become imperative. Within this section, the focus will be the principal contributions, approaches, and findings relating to this issue.

Early attempts at combating fake news relied heavily on fact-checking; either done by individuals, or through the use of traditional media. However, the enormous volume of content generated on social media and news sites necessitated the development of automated techniques. This need brought about attempts to apply machine learning towards the classification of news articles in terms of their credibility. The method taken focused on the vocabulary and style that identifies and differentiates fake/false news from all the real news.

One of the first investigations in this field looked into the application of classical ML techniques like- Naive Bayes (NB), Support Vector Machines (SVM) and LR. All these models have been trained using textual data, including simple features such as word counts and part of speech tagging, as well as more advanced representations like Word2Vec for news articles [11]. Naive Bayes classifiers, as shown in [20] and [21], are amongst the earliest algorithms applied to fake news detection owing to their robustness in text classification. They operate by estimating the likelihood of a news article being fake or true, based on the occurrence of certain words and phrases. Along the same lines, SVMs, using text features of the articles noted for their proficiency in separating two classes, were employed to distinguish fake news from true news articles [2].

As the domain evolved, focus on advanced feature extraction methods. The development of TF-IDF combined with vector representations of words like TF-IDF, Word2Vec or GloVe enabled capturing the significance of wordings inside any document in relation to a group of documentations [15]. TF-IDF plays a vital role in artificial intelligence by identifying words that are crucial to the generation of fake news and, thus, enhancing model training [4]. Besides, n-gram models, which take into consideration n sequences of words, were developed to capture the context that word (unigram) models miss.

Alongside these changes, the beginning of DL opened new possibilities for advancements of fake news detection. Neural DL models, in particular, provided the capacity for automatic derivation of multi-level variables from all of the raw textual data [18]. Text classification problems were solved through adaptations of convolutional neural networks (CNNs) which were initially designed for image processing. As CNNs are capable of identifying local text patterns, they are able to recognize indicators of fake news at the phrase or word group level. LSTMs, which are a type of RNN, as well as some bidirectional models, are skilled at processing sequences of data and capturing long range dependencies in text. LSTMs excel at assessing the logical flow and coherence in news articles, which differentiates fake from true news, and is critical in fake/falsified news detecting [17].

The detecting with the classifying for fake/false news will be further advanced by applying BERT and other transformer models. BERT and its variants have excelled and set new benchmarks in several NLP disciplines. BERT comprehends meaning of words within the context of which they appear in text both to their left and to their right. Word embedding provides an additional aid in grasping the context of the given text. BERT based models and models based on BERT are pre-trained on large corpus and then they are refined on specific datasets which makes BERT and such models perform with accuracy in a range of tasks such as text classification and detection of fake news.

Also, amalgamation of multimodal data seeks to enhance advancement of fake/false news detecting system(s). As with many news stories, fake news often contains images, videos, and even infographics to make the content more appealing, which can also be deceptive or altered. Scholars have looked into methods of analyzing the multimedia facets together with the text. For example, through the application of some image evaluation methods along with text categorization algorithms, the identification of fake news can be achieved in a more sophisticated manner image analysis and text classification [22]. This approach enhances detection systems by using a variety of data types and, therefore, reinforcing the systems.

In addition, the use of fake news detection algorithms has its roots in feature engineering. In addition to the text, some researchers looked at the metadata with attention to the document's source, its publishing date, the author's information, and social interaction data with the document (like, share, and comment) [13], [19]. These features can justify the credibility or non-credibility of the news document and enhance its context. For instance, such news articles are more likely to be untrue along with obscure and newly established websites.

The literature has similarly placed significant emphasis on ensemble techniques. Ensemble learning is used for combining many techniques, in order to achieve improvement on a system's overall predictive accuracy. Variance and bias reduction have been accomplished through the output combination of several individual models using bagging and boosting techniques.

RF classifiers aggregate predictions of multiple decision trees (DTs) and perform well on text classification because of their robustness and the large number of features they can accommodate. The GB algorithms, which focus on progressively improving accuracy by capturing and correcting misclassifications of prior iterations, have also been very effective for detecting fake/false news.

Moreover, some scholars have considered the sparsely labeled datasets problem from the perspective of unsupervised and semi-supervised learning techniques. To facilitate pattern and anomaly detection in relation to the emergence of fake news, some scholars have attempted to cluster news articles utilizing k-means and hierarchical clustering. The semi-supervised learning approach, which employs both labeled and unlabeled data, has been implemented to increase model generalization and to reduce the reliance on extensively labeled datasets.

The scholarly literature discusses the ethical considerations related to the design of fake news detection systems. Concepts such as fairness, transparency, and bias in the algorithm are noted to be critical for the acceptable application of such systems. As an example, scholars have pointed out the importance of designing models which are devoid of bias and do not disproportionately target certain groups or ideologies. Users and stakeholders can be compensated through the model's demonstrating trust by providing transparency in the algorithm’s rationale for its decisions.

2.2 Conclusion

As discussed earlier, there is a plethora of published work regarding the application of ML and NLP for the detecting of fake/falsified news, featuring numerous lines of inquiry and diverging a wide variety of methodologies and approaches. Significant advancements have been achieved in the detecting of fake/falsified news and these improvements have been made using traditional machine learning algorithms, complicated DL models, and multimodal analyses. Moreover, the trustworthiness of the models is enhanced by feature engineering, ensemble techniques, and ethical designs, among others. In the case of the persistent issues posed by fake news, there is a critical need for the online world to adapt to the constantly shifting information landscape.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This very section goes in detailing the methodology, going through our study on fake news detection. It deeply focus on the used ML models employed in the process. Through the data preprocessing, feature extraction, model selection, training, hyperparameter tuning, and evaluation, the methodology section can be considered the heart of this study. Each individual step, here, contributes deeply to the construction of a reliable, accurate models for detecting fake news.

3.1.1 Dataset Collection

In our research, the NLP datasets “Fake.csv” and “True.csv” were selected for the classification. These are the prominent datasets in the NLP community and are important for training and testing the ML algorithms for the text classification tasks, particularly for the detection of fake news.

3.2 Characteristics of the Dataset

Here, the “Fake.csv” dataset includes news articles, which are vastly considered as misleading or false from different and diversified domains of politics, health and current affairs. Each entry includes the news article's text alongside metadata such as the title, author, and date of publication. In contrast, the “True.csv”, another dataset, contains articles verified as true news and includes text and metadata from various domains.

3.3 Sources of Dataset

The “Fake.csv” and “True.csv” datasets are developed from news articles collected from both prominent and obscure news websites. They make use of various repositories and projects aimed at addressing the misinformation problem. The texts included in “True.csv” dataset was gone through and checked by trusted entities, and therefore, are considered accurate and reliable. In contrast, the “Fake.csv” dataset consists of articles labelled as false by fact-checking agencies.

3.4 Data Preprocessing

The raw information seen in the text files requires much preprocessing for any ML model to be tested, to analyze and use them, especially in training ML algorithms. The following steps outline the preprocessing methods applied to both datasets:

- Data Cleaning: This process also consists of removing text fragments such as HTML tags, special symbols, punctuation, and even entire words like of the. The latter category is referred to as stop words, words that have minimal meaning and in the context of "noisy" data, they can easily be eliminated without causing any loss to the value.
- Tokenization: The process of transforming the text to be analyzed by a pattern of words or tokens, also popularly called tokenization. This preparatory phase is of great importance, as it will ensure that the models analyze the text and as a result interpret the output without complications in the subsequent intricate steps.
- Vectorization: After the text has been tokenized, it is converted into a numerical format using techniques such as TF-IDF. This evaluates the significance of a term in a particular context in comparison with an ensemble of documents. This is beneficial in accentuating the important distinguishing words that differentiate fake news articles from real ones.
- Lemmatization and Stemming: These methods aim to abstract words to their base forms. Context is important in meaningful word reduction, such as in lemmatization, while in stemming, a more uniform approach is applied by suffix removal. These methods assist in text uniformity by reducing the feature set complexity.
- Splitting the Dataset: In this case, the whole dataset is first split into one set used for training of the ML model and a set reserved for testing. This is done in a 70 to 30 ratio. The model is first trained by applying the mentioned data to be used for training, or training set, to the designated ML attributes, while the reserved set undergoes model evaluation after training. This method of separation ensures that the models constructed for learning are based on a representative sample of real data and are exposed to fresh, unseen data for evaluation to test their ability to generalize.
- Balancing the Dataset: In situations where there is an imbalance in the proportion of fake and real news articles, techniques like oversampling (increasing the amount or count of underrepresented samples) or under sampling (decreasing the amount or

count of overrepresented samples) may be implemented. In this case, dataset imbalance is mitigated to obstruct the used model from biasing accuracy toward the used majority class. In turn, this enhances the accuracy and reliability of the models constructed.

3.5 Feature Extraction

Text data is transformed into measurable features for easier comparison and evaluation during training machine learning models is known as feature extraction. For the purpose, the TF-IDF vectorization technique is primarily used. TF-IDF assigns a weighted score to each term, particularly to the documents where it appears most, while measuring the term's importance in comparison to the whole corpus. This technique assists in capturing salient features for differentiating the real and authentic news articles aside from the fake or false news. Along with TF-IDF, other textual features like n-grams, which can also be describes as the contiguous sequences of n number of words, are also used. N-grams are better than single word models (unigrams). For instance, two and three word sequences (bigrams and trigrams) add tremendous value to understanding the text's structure and meaning.

3.6 Model Selection

The research investigates four well-known machine learning algorithms that are effective in binary classification:

- **Logistic Regression (LR):** A linear estimation model predicts a binary outcome and is simple, text- classifier and well understood.
- **Decision Tree Classifier (DTC):** A non-linear algorithm performs classification by recursively partitioning a dataset into smaller, easier to classify portions, producing a tree-like structure. It is simple interpretable and captures sophisticated relationships remarkably well.
- **Gradient Boosting Classifier (GBC):** A technique of ensembled learning that is used to build a sequence of, so called, weak learners using DTs, with each of the new iteration correcting the mistakes or errors of its predecessor. It is remarkably powerful and often provides high accuracy.

- **Random Forest Classifier (RFC):** Also a method that uses ensemble learning which builds a set of DTs in parallel while it integrates each of the stage's predictions. It provides good accuracy and reduces overfitting.

3.7 Model Training

Each model is trained individually to the chosen algorithms using a common training set. The training procedure consists of:

- Model Initialization: Each model starts from the given default or baseline values.
- Training: The algorithms are fitted to the cleaned training set. In this stage, the models start learning the mapping between the features, which are the parts of the text extracted by the algorithms, and the labels which are classified as fake or true.
- Cross-Validation: Cross-validation techniques, like k-fold cross-validation, are implemented to assess how well a model performs across different segments of the training dataset.

This method alleviates the risk of overfitting a model to the training dataset, thus ensuring more robust model performance on unseen data.

3.8 Hyperparameter Optimization

Hyperparameter tuning is critical for improving model performance in machine learning. Each model's optimal hyperparameters are systematically analyzed using GridSearchCV. This process consists of the following steps:

- Defining the Hyperparameter Grid: For every hyperparameter, a range of possible values is defined. Taking the Random Forest Classifier as an example, relevant hyperparameters consist of 'n estimators', or the quantity of trees, 'max depth', the maximum depth of each tree, along with 'min samples split', the minimum number of samples necessary to split a node.
- Exhaustive Search:
    - Vectorization: GridSearchCV carries out an exhaustive search across all specified hyperparameters for a fixed model, training, as well as, validating each of the model for every set of hyperparameters.

- Selecting the Best Model: The combination of hyperparameters which have the best cross-validation score is taken and considered as the optimal model.

3.9 Evaluation Metrics

Comprehensive evaluation of the models involves several metrics calculation to obtain a holistic appraisal of model performance:

- Accuracy: The model accuracy metric can be said as the ratio of accurately predicted terms from the overall total instances, serving as a broad measure of model effectiveness.
- Precision: Evaluates accuracy for the positive or correct guesses by determining the proportion of true positive cases over the total positive predictions made (true positives and false positives).
- Recall (Sensitivity): Determines the ratio of real positive cases that are detected by the model. It measures the ratio of true positives against the total actual positives which comprises of true positives and false negatives.
- F1-Score: Combines both precision and recall in a single measure which is useful in cases of a class imbalance. It is derived as the harmonic mean for precision and recall.
- Confusion Matrix: A table that allows visualization of the results of the model's prediction by displaying the true positives, false positives, true negatives, and false negatives. Thus, it gives some information about the model's errors.

3.10 Model Evaluation

We assess the trained models against the testing set using the defined metrics. This involves:

- Predicting on the Testing Set: Assigning the label of fake or true prediction to the articles in the testing set is done by each model.
- Calculating Evaluation Metrics: Calculating the defined metrics based on the actual results for the given predictions of the class labels like accuracy, precision, recall, F1-score, and the confusion matrix.

- Comparative Analysis: The descriptive analytics of the models based on the defined metrics are used for comparison in order to determine the best model.

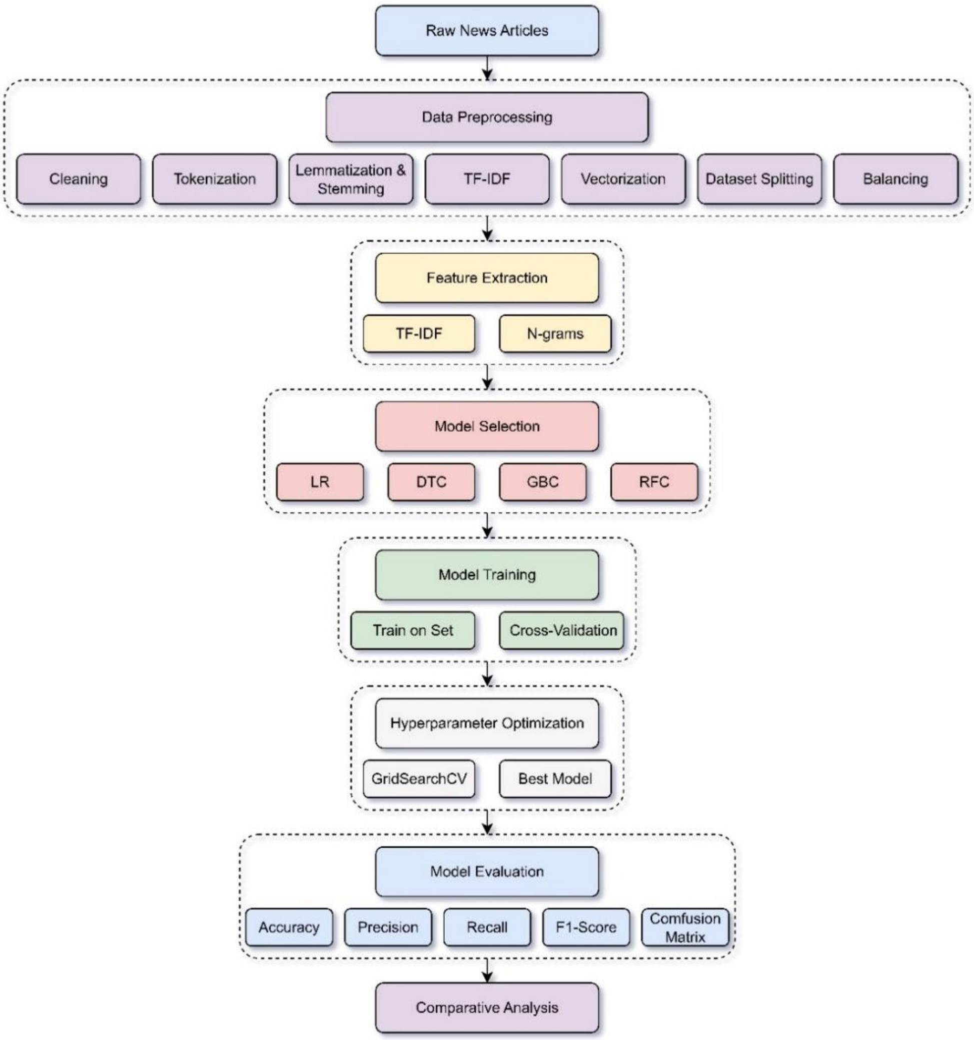


Fig. 1. Block diagram representation of the proposed architecture

Fig. 1 highlights the study's architecture. It provides an understanding of the workflow of the research and also helps in creating an illustrative framework.

3.11 Summary

As previously stated, feature extraction involves transforming the text based data to measurable attributes relevant in order to train ML models. For this purpose, TF-IDFV vectorization is most commonly utilized. TF-IDF scalar quantifies the significance of any word or text inside of a documentation and relation to any collection of documents, giving a score of importance to every word. This is useful in selecting crucial attributes that aids in distinguishing fake and true news articles. Other than TF-IDF, other text attributes, n-grams, cited as the contiguous sequence of n number of words, are also used. N-grams capture context and phrase-level semantics which is absent from unigram models. Like, bigrams (two-word sequences) and trigrams (three-word sequences) are extremely useful for the understanding of the textual and semantic structure of language.

CHAPTER 4

RESULT AND DISCUSSION

4.1 Introduction

This section summarizes and analyzes the machine learning techniques employed for detecting fake news and their results. We studied the results of models built using LR, DTC, GBC and RFC on an independent test set. Crucial evaluation metrics like accuracy, precision, recall, and F1-score were calculated, and the performances were evaluated. The results clearly show the effectiveness and usefulness of the proposed models to identify and separate the fake/falsified news from all the real news accurately.

4.2 Logistic Regression

The LR model performs well on the test set. This model is evaluated on metrics as follows on its test set:

- Precision: 0.99 for class 0 (true news) and 0.98 for class 1 (fake news).
- Recall: 0.99 for both classes.
- F1-Score: 0.99 for both classes.
- Support: 5829 for class 0 and 5391 for class 1.

The LR model achieved a total accuracy of one hundred percent which means it accurately categorized ninety nine percent of the news articles in the test set. Its macro average and weighted average scores for precision, recall as well as F1 score are all equal to point nine, which implies balanced performance for the class and category.

4.3 Decision Tree Classifier

The Decision Tree Classifier performed remarkably well achieving an accuracy score of one. Below is a summary for the classification report metrics:

- Precision: 1.00 for both class 0 and class 1.

- Recall: 1.00 for class 0 and 0.99 for class 1.
- F1-Score: 1.00 for both classes.
- Support: 5829 for class 0 and 5391 for class 1.

The overall performance of the Decision Tree Classifier is equal to one hundred percent which demonstrates its outstanding ability to classify and categorize the news articles with almost perfect accuracy. However, it underperforms in recall for class 1 reference to the fake news which indicates that the model has the tendency of overfitting.

4.4 Gradient Boosting Classifier

The Gradient Boosting Classifier achieved outstanding performance, and accuracy score of one, below are detailed metrics:

- Precision: 1.00 for class 0 and 0.99 for class 1.
- Recall: 0.99 for class 0 and 1.00 for class 1.
- F1-Score: 1.00 for both classes.
- Support: 5829 for class 0 and 5391 for class 1.

The model demonstrates outstanding performance with its perfect 100% accuracy that showcases its ability to differentiate fake news articles and real news articles, and in addition the macro and weighted averages for precision, recall and F1 score are all 1 which proves outstanding performance.

4.5 Random Forest Classifier

The Random Forest Classifier demonstrated good performance achieving an accuracy of 99% The classification report is as follows:

- Precision: 0.99 for both class 0 and class 1.
- Recall: 0.99 for both classes.
- F1-Score: 0.99 for both classes.
- Support: 5829 for class 0 and 5391 for class 1.

The Random Forest Classifier demonstrates noteworthy capability in the accurate classification of news articles, achieving an overall accuracy of 99%. Moreover, the macro and weighted averages for precision, recall, and F1-score all reach 0.99, showing consistency across both categories.

4.6 Comparative Analysis

The table below summarizes the performance of the four models based on the evaluation metrics:

Table 1.1 Performance metrics (part 1) of different machine learning models for fake news detection.

Model	**Accuracy**	**Precision (0)**	**Precision (1)**	**Support (0)**
LR	0.99	0.99	0.98	5829
DTC	1	1	1	5829
GBC	1	1	0.99	5829
RFC	0.99	0.99	0.99	5829

Table 1.2 Performance metrics (part 2) of different machine learning models for fake news detection.

Model	Recall (0)	Recall (1)	F1-Score (0)	F1-Score (1)	Support (1)
LR	0.99	0.99	0.99	0.99	5391
DTC	1	0.99	1	1	5391
GBC	0.99	1	1	1	5391
RFC	0.99	0.99	0.99	0.99	5391

4.7 Discussion

The provided results allow us to make the strong hypothesis that the fourth model also captures fake news accurately. and that these models achieved outstanding classification performance thanks to high model accuracy and well-optimized hyperparameters. Both Gradient Boosting and Decision Tree Classifier performed exceedingly well, separated only by a small decline in recall from one of the classes in the decision tree. Together with Random Forest and Logistic Regression, these models provided balanced results and confirmed strong TF-IDF performance and hyperparameter optimization, considerably aiding model performance by allowing the models to distinguish fake news.

These discoveries are important for automating systems for detecting fake/falsified news. Misinformation being clearly a severely outgrowing issue that machine learning is a potential solution for, especially considering automated systems can achieve balanced and accurate performance. If machine learning models are integrated into news aggregator services and social media platforms, devices could provide automated information protection and shielding services, in essence, fortifying personal information security and protecting citizens from the dangers related to misinformation.

BERT and other transformer models should be studied to grasp language's deeper subtleties. Furthermore, a more in-depth fake news analysis could be done by taking a multimodal approach, analyzing text while incorporating images or videos. These technologies also raise important social issues, such as bias, and the need for transparent and accountable design.

The study demonstrates that machine learning models are capable of analyzing and classifying fake news. It also suggests that there should be more machine learning techniques, especially advanced models like the Gradient Boosting and Decision Tree Classifiers. These models are proven to be high performing and reliable in real-world scenarios. The need for ongoing work in the digital space to counter misinformation is more critical than ever.

CHAPTER 5

CONCLUSION

5.1 INTRODUCTION

This study underscores the growing relevance of the problem of misinformation by focusing specifically on the classification of false information using datasets and machine learning models. The models LE, DTC, GBC and RFC were thoroughly trained and tested on the datasets “Fake.csv” and “True.csv” containing labeled datasets. The results were exceptionally accurate, as DTC and GBC not only performed exceptionally, but also reached the remarkable benchmark of 100% accuracy. This attests to the effectiveness of these models in the binary classification problem of fake and true news. In general, LR and RFC also performed remarkably in accuracy, and their precision, recall, F1-score metrics were well balanced. The findings of this study showcase the growing advancements through the world of ML with the attempts to automize the detection of fake news, which is critical in the mitigation of the spread of misinformation.

During the mentioned process of model training, the machine learning models underwent a series of preprocessing steps such as data cleansing, tokenization, lemmatization, stemming, and TF-IDF vectorization which transformed the raw text data into a form the models could process. Each individual model's performance was enhanced by careful selection of hyper-parameters for each model individually using GridSearchCV, which contributed greatly towards accuracy.

These findings impact many sectors such as journalism, politics, and social media. More effective fake/falsified news detecting processes will hugely assist and determine in regaining the public’s media trust, enabling informed choices, and protecting democratic systems. When applied to news aggregation services and social media, trust can be built in the information systems and people can be protected from the dangers of strategic misinformation and propaganda.

Further research needs to be conducted to incorporate more sophisticated natural language processing techniques, including BERT and other transformer models focused on complex language comprehension. In addition, the textual analysis of images and videos as multi-modal data can enable a more holistic approach to fake news detection. From an ethical standpoint, such technologies must also be safeguarded with social responsibility with regards to, for example, bias and transparency.

In essence, this research showcases the ever-growing need and potential and opportunities ML models present in assisting the identification of fake news. Emphasis on the effectiveness of the models proves their utility and uses in various scenarios, including the battle against disinformation. Dealing with the challenges posed by the pervasive dissemination of misinformation will require the modification of the existing countermeasures and the implementation of proactive study and progress in the information science field due to the ever-evolving nature of technology. The adoption of such technologies will greatly bolster the trust and resilience of the digital space in question.

REFERENCES

[1] S. T., ‘THE CHALLENGES AND ROLE OF MODERN JOURNALISM IN DIGITAL MEDIA ENTERPRISES’, *ShodhKosh: Journal of Visual and Performing Arts*, vol. 5, Jul. 2024, doi: 10.29121/shodhkosh.v5.i7.2024.1921.

[2] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, 2000.

[3] Y. Goldberg and O. Levy, "word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method," *arXiv preprint* arXiv:1402.3722, 2014.

[4] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22-36, 2017.

[5] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645-6649, 2013.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[7] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, pp. 5998-6008, 2017.

[8] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint* arXiv:1409.0473, 2014.

[9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv preprint* arXiv:1810.04805, 2018.

[10] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," *arXiv preprint* arXiv:1301.3781, 2013.

[11] K. Shu, S. Wang, and H. Liu, "Beyond news contents: The role of social context for fake news detection," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pp. 312-320, 2019.

[12] W. Y. Wang, "Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pp. 422-426, 2017.

[13] N. Lomas, "Google's Jigsaw launches new anti-fake news AI tool aimed at journalists," *TechCrunch*, 2018.

[14] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532-1543, 2014.

[15] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746-1751, 2014.

[16] M. Schuster and K. K. Paliwal, "Bidirectional Recurrent Neural Networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, 1997.

[17] Z. Zhang and M. H. Lee, "A neural network approach to ordinal regression," *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1514-1523, 2005.

[18] N. Diakopoulos and D. A. Shamma, "Characterizing debate performance via aggregated Twitter sentiment," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1195-1198, 2010.
[19] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, pp. 900-903, 2017.
[20] A. Kumar and M. M. Sastry, "An approach for fake news detection using machine learning," in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pp. 480-483, 2018.
[21] R. Baly et al., "Multi-source multi-class fake news detection," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 726-732, 2018.
[22] A. Hanselowski et al., "A retrospective analysis of the fake news challenge stance detection task," in *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 1859-1874, 2018.
[23] S. R. Kumar and T. M. Sebastian, "Comparative study on fake news detection using machine learning algorithms," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 890-894, 2019.
[24] Z. Zhao, P. Resnick, and Q. Mei, "Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 1395-1405, 2015.

PLAGIARISM REPORT

211-35-713

ORIGINALITY REPORT

24%	19%	18%	15%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Midlands State University Student Paper	2%
2	aclanthology.org Internet Source	2%
3	dokumen.pub Internet Source	1%
4	Submitted to Coventry University Student Paper	1%
5	arxiv.org Internet Source	1%
6	S.P. Jani, M. Adam Khan. "Applications of AI in Smart Technologies and Manufacturing", CRC Press, 2025 Publication	1%

LIBRARY CLEARANCE